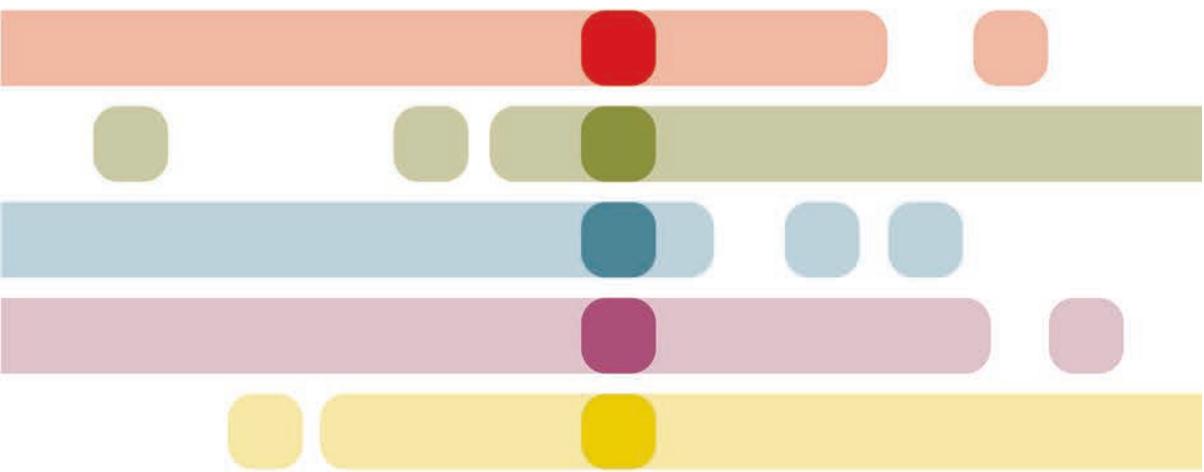




Società Italiana di Ricerca Didattica

Giornale Italiano della Ricerca Educativa
Italian Journal of Educational Research

anno XII
numero speciale
Settembre 2019





Giornale Italiano della Ricerca Educativa
Italian Journal of educational Research

RIVISTA SEMESTRALE

anno XII – numero speciale SETTEMBRE 2019

Guest Editors: Luciano Cecconi, Roberto Trinchero

Direttore | Editor in chief

PIETRO LUCISANO | Sapienza Università di Roma

Condirettori | Co-editors

LOREDANA PERLA | Università degli Studi di Bari "Aldo Moro"

ACHILLE M. NOTTI | Università degli Studi di Salerno

Comitato Scientifico | Editorial Board

GIOVANNI BONAIUTI – University of Cagliari (Italy)

PILAR MARTINEZ CLARES – University of Murcia (Spain)

EAN-MARIE DE KETELE – Université Catholique de Lovanio (Belgium)

LORETTA FABBRI – University of Siena (Italy)

ETTORE FELISATTI – University of Padua (Italy)

LUCIANO GALLIANI – University of Padua (Italy)

MARIA LUCIA GIOVANNINI – University of Bologna (Italy)

MARIA LUISA IAVARONE – Parthenope University of Naples (Italy)

ALESSANDRA LA MARCA – University of Palermo (Italy)

PATRIZIA MAGNOLER – University of Macerata (Italy)

ANTONIO MARZANO – University of Salerno (Italy)

GIOVANNI MORETTI – University of Rome 3 (Italy)

ELISABETTA NIGRIS – University of Milano-Bicocca (Italy)

ANITA NORLUND – University of Borås (Sweden)

ACHILLE M. NOTTI – University of Salerno (Italy)

ROSSELLA SANTAGATA – University of California, Irvine (USA)

JAAP SCHEERENS - University of Twente (Netherlands)

MARÍA JOSÉ MARTÍNEZ SEGURA – University of Murcia (Spain)

VITALY VALDIMIROVIC RUBTZOVA – Moscow City University (Russia)

ROBERTO TRINCHERO – University of Turin (Italy)

IRA VANNINI – University of Bologna (Italy)

BYUNG JUN YI – Pusan National University, Busan (South Korea)

Comitato editoriale | Editorial management

CRISTIANA DE SANTIS | Sapienza Università di Roma

DANIA MALERBA | Sapienza Università di Roma

ROSA VEGLIANTE | Università degli Studi di Salerno

ARIANNA LODOVICA MORINI | Università degli Studi Roma TRE

Note per gli Autori | Notes to the Authors**Per l'invio delle proposte:**

<https://ojs.pensamultimedia.it/index.php/sird/about/submissions>

For management and submission of proposals:

<https://ojs.pensamultimedia.it/index.php/sird/about/submissions>

Consultazione numeri rivista

<http://ojs.pensamultimedia.it/index.php/sird>

Codice ISSN 2038-9744 (testo on line)

Registrazione Tribunale di Bologna n. 8088 del 22 giugno 2010

Finito di stampare: Settembre 2019

Editing e stampa

Pensa MultiMedia Editore s.r.l. - Via A. Maria Caprioli, 8 - 73100 Lecce - tel. 0832.230435
www.pensamultimedia.it - info@pensamultimedia.it

Progetto grafico copertina

Valentina Sansò

Obiettivi e finalità | Aims and scopes

Il **Giornale Italiano della Ricerca Educativa**, organo ufficiale della **Società Italiana di Ricerca Didattica (SIRD)**, è dedicato alle metodologie della ricerca educativa e alla ricerca valutativa in educazione.

Le aree di ricerca riguardano: lo sviluppo dei curricula, la formazione degli insegnanti, l'istruzione scolastica, universitaria e professionale, l'organizzazione e progettazione didattica, le tecnologie educative e l'e-learning, le didattiche disciplinari, la didattica per l'educazione inclusiva, le metodologie per la formazione continua, la docimologia, la valutazione e la certificazione delle competenze, la valutazione dei processi formativi, la valutazione e qualità dei sistemi formativi.

La rivista è rivolta a ricercatori, educatori, formatori e insegnanti; pubblica lavori di ricerca empirica originali, casi studio ed esperienze, studi critici e sistematici, insieme ad editoriali e brevi report relativi ai recenti sviluppi nei settori. L'obiettivo è diffondere la cultura scientifica e metodologica, incoraggiare il dibattito e stimolare nuova ricerca.

The **Italian Journal of Educational Research**, promoted by the **Italian Society of Educational Research**, is devoted to Methodologies of Educational Research and Evaluation Research in Education.

Research fields refer to: curriculum development, teacher training, school education, higher education and vocational education and training, instructional management and design, educational technology and e-learning, subject teaching, inclusive education, lifelong learning methodologies, competences evaluation and certification, docimology, students assessment, school evaluation, teacher appraisal, system evaluation and quality.

The journal serves the interest of researchers, educators, trainers and teachers, and publishes original empirical research works, case studies, systematic and critical reviews, along with editorials and brief reports, covering recent developments in the field. The journal aims are to share the scientific and methodological culture, to encourage debate and to stimulate new research.

Comitato di referaggio | Referees Committee

Il Comitato di Revisori include studiosi di riconosciuta competenza italiani e stranieri. Responsabili della procedura di referaggio sono il direttore e il condirettore della rivista.

The Referees Committee includes well-respected Italian and foreign researchers. The referral process is under the responsibility of the Journal's Editor in Chief and Co-Editors.

Procedura di referaggio | Referral process

Il Direttore e Condirettore ricevono gli articoli e li forniscono in forma anonima a due revisori, tramite l'uso di un'area riservata nel sito della SIRD (www.sird.it), i quali compilano la scheda di valutazione direttamente via web entro i termini stabiliti. Sono accettati solo gli articoli per i quali entrambi i revisori esprimono un parere positivo. I giudizi dei revisori sono comunicati agli Autori, assieme a indicazioni per l'eventuale revisione, con richiesta di apportare i cambiamenti indicati. Gli articoli non modificati secondo le indicazioni dei revisori non sono pubblicati.

Editor in chief and co-editor collect the papers and make them available anonymously to two referees, using a reserved area on the SIRD website (www.sird.it), who are able to fulfill the evaluation grid on the web before the deadline. Only articles for which both referees express a positive judgment are accepted. The referees evaluations are communicated to the authors, including guidelines for eventual changes with request to adjust their submissions according to the referees suggestions. Articles not modified in accordance with the referees guidelines are not accepted.

INDICE

- 8 **INTRODUZIONE**
a cura di **LUCIANO CECCONI, ROBERTO TRINCHERO**
- 13 **PAOLO MARIA FERRI**
MOOC, digital university teaching and Learning analytics. Opportunities and Perspectives | **MOOC, didattica universitaria digitale e Learning analytics. Opportunità e prospettive**
- 27 **KATIA SANNICANDRO, ANNAMARIA DE SANTIS, CLAUDIA BELLINI, TOMMASO MINERVA**
Analysis of completion and dropout rates in EduOpen MOOCs | **Analisi di tassi di completamento e abbandono nei MOOC di EduOpen**
- 43 **LEONARDO CAPORARELLO, FEDERICA CIRULLI, MANUELA MILANI**
Design of a Learning Analytics framework proposal in academic context | **Proposta di un framework per i Learning Analytics nel contesto accademico**
- 56 **ROSA VEGLIANTE, MARTA DE ANGELIS**
E-tutoring in online training courses | **L'e-tutoring nei percorsi di formazione online**
- 69 **MICHELE BALDASSARRE, VALERIA TAMBORRA**
How can we make it happen? From description to foresight of online learning environments | **Come possiamo farlo succedere? Dalla descrizione alla previsione negli ambienti di apprendimento online**
- 89 **ALBERTO FORNASARI**
Learning Analytics for distance learning in Lifewide Learning contexts. A proposal to analyze the reporting of an e-learning course for Sudanese university students | **Learning Analytics per la formazione a distanza in contesti di Lifewide Learning. Una proposta di analisi della reportistica di un corso e-learning per studenti universitari sudanesi.**
- 109 **ANNA DIPACE, LUCIA MARIA BORRELLI, SARA PERRELLA, KATIA CAPOSENO**
Student demographics and MOOC selection on Eduopen. An exploratory study on UniFg MOOCs | **Dati demografici degli studenti e selezione dei MOOC su Eduopen. Uno studio esplorativo sui MOOC erogati da UniFg**

- 125 **GIOVANNI MORETTI, ARIANNA GIULIANI, ARIANNA LODOVICA MORINI**
The valorisation of the Learning Analytics to promote assessment for learning: the case of the Master in Leadership and Management in Education of Roma Tre University | Valorizzare i Learning Analytics per promuovere la valutazione per l'apprendimento. Il caso del Master in Leadership e Management in Educazione dell'Università Roma Tre
- 140 **DANIELA ROBASTO**
Learning Analytics for monitoring and analyzing student's careers at University | Learning Analytics per il monitoraggio e l'analisi delle carriere degli studenti universitari.
- 156 **GIACOMO NALLI, LEONARDO MOSTARDA, ANDREA PERALI, SEBASTIANO PILATI, DANIELA AMENDOLA**
Application of machine learning to the learning analytics of the Moodle platform to create heterogeneous groups in on-line courses | Applicazione del machine learning ai learning analytics della piattaforma Moodle per creare gruppi eterogenei nei corsi on-line
- 174 **CATERINA BEMBICH**
Social Network Learning Analytics: identification of students at risk of early school leaving | Social Network Learning Analytics: identificazione degli studenti a rischio di abbandono scolastico
- 187 **VALENTINA DELLA GALA**
Analytics dei testi riflessivi scritti dai docenti neoassunti nel portfolio digitale | Newly Qualified Teachers' reflective writing: analytics and preliminary results
- 205 **GIUSI ANTONIA TOTO**
Learning Analytics and special learning needs: a possible combination | Learning Analytics e bisogni speciali di apprendimento: un connubio possibile.
- 217 **LUISA PANDOLFI**
Promote educational success: what is the contribution of Learning Analytics to educational research? | Promuovere il successo formativo: quale il contributo del Learning Analytics per la ricerca educativa?

Ricerca, dati e democrazia



Il sistema educativo italiano è dotato, per sua fortuna, di un alto grado di resilienza. Nonostante gli urti che negli ultimi cinquanta anni hanno scosso la sua struttura il sistema sembra sopravvivere. Una resilienza che gli ha permesso di resistere, oltre agli urti delle riforme (che quasi sempre si pretendevano epocali), anche all'effetto erosivo di fattori quali, ad esempio, una spesa per l'istruzione inadeguata (in particolare per quella terziaria), un sistema di reclutamento che, di sanatoria in sanatoria, ha impoverito il profilo culturale dei docenti assunti, l'assenza di un sistema di formazione continua che promuovesse lo sviluppo del personale in servizio. Una erosione che nel tempo ha prodotto un corpo docente con un'età media tra le più alte del mondo, elemento che da solo ben rappresenta l'incapacità del sistema di rinnovarsi nel corso del tempo. Ebbene, per effetto di questi urti e di queste erosioni (e di tanto altro) il nostro sistema educativo oggi dovrebbe essere al collasso ma sappiamo che così non è. Il sistema è vivo anche se non sta affatto bene, come dimostrano i risultati delle indagini comparative internazionali e le più recenti rilevazioni nazionali.

Da un punto di vista più vicino alla ricerca che alla polemica politica e/o mediatica sarebbe più utile interrogarsi su questa miracolosa sopravvivenza che non su questioni che recentemente si sono imposte al centro del dibattito come la "distruzione della scuola" o la "scuola che non educa più". Senza trascurare l'effetto cumulativo prodotto nel lungo periodo dai fattori erosivi prima ricordati, che improvvisamente potrebbero provocare frane e smottamenti, ciò che oggi è urgente è fermarsi ad osservare lo stato di salute di una organizzazione complessa, di analizzarne componenti, processi e prodotti, e sulla base delle informazioni e dei dati raccolti tentare una diagnosi.

Ma chi osserva il sistema e, soprattutto, come lo fa? L'attenzione dei media al sistema educativo si risveglia puntualmente alla fine e all'inizio

di ogni anno scolastico e poi tende inesorabilmente a cadere come le foglie in autunno, con il risultato che la conoscenza del sistema educativo rimane al livello superficiale dei servizi di giornali e tv e a quello tanto istantaneo quanto evanescente dei social media. In tutto questo, sia chiaro, non c'è nulla di strano, i media fanno il loro mestiere che, come è noto, non è quello lungo e paziente della ricerca ma quello immediato, sintetico e ad effetto della notizia. Se parliamo di osservazione sistematica e di analisi approfondite il soggetto da chiamare in causa è un altro: la ricerca, in particolare quella educativa. A questo proposito è lecito chiedersi: la ricerca educativa, nell'ultimo mezzo secolo, in quale misura ha contribuito a produrre nuova conoscenza? A ben vedere la conoscenza prodotta, per quantità e qualità, non è stata di grande rilevanza per la piena comprensione del sistema educativo e, soprattutto, dei suoi problemi. Inoltre, il prestigio e l'autorevolezza dei "pedagogisti" nella società e in particolare nel mondo della politica sono diminuiti e questo ha indebolito notevolmente la loro capacità di influire sulle politiche educative.

Non solo, sappiamo che la ricerca ha bisogno di essere orientata e messa nelle condizioni di operare e questo è un compito della politica. Purtroppo negli ultimi decenni le amministrazioni che si sono succedute alla guida del Paese, al di là della retorica elettorale, non hanno garantito, né alla ricerca né all'istruzione, le condizioni per sostenere e sviluppare il nostro sistema educativo. Si tratta insomma di due debolezze, quella della ricerca educativa e quella della politica, che si confermano e si rinforzano reciprocamente.

Questa mancata sinergia tra ricerca, scuola e politica, tra educazione e democrazia non ci deve preoccupare soltanto per lo stato di salute dell'educazione ma anche per quello della nostra democrazia. Come ci ha insegnato Dewey l'una non può fare a meno dell'altra. Prendiamo come esempio il problema attualissimo dell'integrazione e dell'inclusione sociale, come può una società democratica affrontare tale problema senza una profonda e lunga azione educativa? E come può un sistema educativo essere inclusivo senza l'interazione costante e continua con una società democratica?

Il mondo della ricerca, dunque, risente negativamente di un clima politico che in alcuni casi è più che insensibile nei confronti della cultura e dell'educazione, è addirittura ostile (si pensi al discredito sociale nei confronti della cultura, degli esperti, dei tecnici e delle competenze). Possiamo rimanere in attesa che si affermi un clima politico e sociale più favorevole all'educazione? Evidentemente no. La comunità della ricerca educativa deve continuare con più determinazione a fare la sua parte che consiste in primo luogo nell'assumersi risolutamente il com-





pito di condurre l'osservazione e l'analisi del sistema educativo allo scopo di svilupparlo, e in secondo luogo nell'affermare che questo lavoro deve essere condotto con un atteggiamento scientifico che, come diceva Dewey, consiste nel "desiderio di ricercare, esaminare, discriminare, tracciare conclusioni solo sulla base dell'evidenza, dopo essersi presi la pena di raccogliere tutti i dati possibili".

Ecco, qui veniamo al punto. Nel campo della ricerca educativa vanno potenziate e valorizzate tutte quelle iniziative che hanno un carattere empirico e responsivo, che si basano cioè sulle evidenze acquisite mediante la raccolta di tutti "i dati possibili" allo scopo di rispondere a reali esigenze del mondo dell'educazione. Si tratta di metodi, tecniche e strumenti, talvolta nuovi, che vanno studiati, perfezionati, adattati e sperimentati. I progressi e la grande diffusione delle TIC, anche in campo educativo, hanno portato alla produzione di grandi quantità di dati, i cosiddetti *big data*, che possono aiutare a formare quelle evidenze che danno significato e solidità alla ricerca di tipo scientifico con una rapidità e con una facilità inimmaginabili soltanto venti anni fa. Naturalmente i dati da soli, per quanto numerosi possano essere, non producono nuova conoscenza, hanno bisogno di teorie e metodi che orientino la loro raccolta, la loro analisi e la loro interpretazione. Questo vale anche per i *learning analytics*, cioè i dati prodotti in grandi quantità dagli ambienti di apprendimento digitali, e per le teorie pedagogiche. Questo significa che la ricerca educativa in questo momento deve attivare tutte le interazioni possibili tra la dimensione teorica e quella empirica, tra i *fini* e i *mezzi*. Dare una *intelligenza pedagogica* ai dati è una sfida che ci attende tutti alla prova dei fatti. Una intelligenza, è bene ricordarlo, da cui dipende anche la solidità del legame che si stabilisce tra educazione e democrazia. La resilienza, infatti, può aiutare ma non è la soluzione al problema.

È proprio in questa prospettiva che la SIRD e la SIE-L nel maggio scorso hanno organizzato presso l'Università di Roma La Sapienza il Convegno internazionale "Learning Analytics. Per un dialogo tra pratiche didattiche e ricerca educativa". È stato un successo inaspettato che testimonia l'interesse della comunità nei confronti del tema e che ha convinto le due associazioni a confermare il loro impegno ad approfondire la ricerca in questo settore. Prima di tutto con la pubblicazione sull'*Italian Journal of Educational Research* e sul *Journal of e-Learning and Knowledge Society* di un numero speciale dedicato ai *learning analytics* (da qui in poi LA) e in secondo luogo con la preparazione di un secondo convegno sul rapporto tra raccolta e analisi dei dati e ricerca educativa che si terrà nella primavera del 2020 presso l'Università di Torino.

Come accennato, i contributi del presente numero speciale si focaliz-

zano sui LA come strumento per stimolare il dialogo tra pratiche didattiche e ricerca educativa.

Il contributo di Paolo Maria Ferri si propone l'obiettivo di tratteggiare uno stato dell'arte relativamente al fenomeno della diffusione dei MOOC. Esso si concentra sull'impatto "disruptive" dei MOOCs sistemi universitari tradizionali delle nazioni più sviluppate, così come sul modo di utilizzare la metodologia dei LA per valutare la qualità dei corsi MOOC, l'efficacia e la valenza formativa. Il contributo di Katia Sannicandro, Annamaria De Santis, Claudia Bellini, Tommaso Minerva si concentra anch'esso sui MOOCs e in particolar modo sul loro tasso di completamento che è generalmente inferiore al 10% degli iscritti a causa di fattori quali: reclutamento spontaneo, estrema eterogeneità degli iscritti, processi di autoregolazione, differenze nei profili motivazionali e culturali, concentrandosi in particolare sul caso del portale EduOpen. Il contributo di Leonardo Caporarello, Federica Cirulli e Manuela Milani, si focalizza sui LA nell'ambito dell'istruzione universitaria, analizzando potenzialità e criticità per lo sviluppo di modelli di riferimento, adattabili ai diversi contesti, per la definizione degli obiettivi e per ciò che dovrebbe essere fatto per raggiungerli. Il contributo di Rosa Vegliante e Marta De Angelis affronta il tema dell'e-tutoring approfondendone alcuni degli aspetti fondamentali, dal punto di vista sia teorico sia pratico: strategie di e-tutoring, ricadute formative, vantaggi tratti dalle funzioni svolte dall'e-tutor nel processo di apprendimento. Il contributo di Michele Baldassarre e Valeria Tamborra propone un approccio di LA rivolto allo studio di *event-driven data* provenienti da 18 corsi attivati all'interno di una piattaforma Moodle afferenti a due categorie: corsi di formazione post-lauream e corsi attivati nell'ambito di progetti di ricerca. L'obiettivo è quello di individuare possibili pattern d'interazione da parte degli utenti iscritti alle due categorie di corso al fine di dare indicazioni utili alla progettazione efficace di futuri corsi. Il contributo di Alberto Fornasari propone un modello di analisi dei report di un corso e-learning Moodle per studenti sudanesi attivato dall'Università di Bari nell'ambito del progetto *INSO – Innovation in society: training paths and human capital enhancement* in Sudan. Il fine è quello di trarre dai dati conoscenze utili a un'efficace progettazione di ambienti di apprendimento online in contesti di lifewide learning. Il contributo di Anna Dipace, Lucia Maria Borrelli, Sara Perrella, Katia Caposeno, mira a descrivere la provenienza dei corsisti dei MOOC del progetto Eduopen, in particolare quelli offerti dall'Università di Foggia, mediante uno studio esplorativo sui dati demografici degli studenti iscritti. Il contributo di Giovanni Moretti, Arianna Giuliani e Arianna Lodovica Morini, esplora le possibilità dei LA per la valutazione per l'apprendimento in-





dagando in che modo le funzionalità connesse ad essi possono valorizzarne lo sviluppo nei soggetti in formazione. Il contributo presenta un'indagine svolta analizzando diacronicamente i dati relativi al Master in Leadership e Management in Educazione (Università Roma Tre) in cui i LA sono stati utilizzati nelle fasi di avvio e di monitoraggio allo scopo di predisporre interventi individualizzati. Il contributo di Daniela Robasto presenta le prime tappe di un progetto pilota svolto presso l'Università di Parma, volto ad utilizzare i LA al fine di condurre a scelte consapevoli e strategiche in merito alla calibrazione delle azioni di orientamento in itinere, all'eventuale riprogettazione dei piani di studi dei CdS oggetto di analisi e all'individuazione precoce di insegnamenti "problematici" dal punto di vista dello studente, anche in vista di percorsi di formazione dei docenti inerenti competenze didattiche e docimologiche. Il contributo di Giacomo Nalli, Leonardo Mostarda, Andrea Perali, Sebastiano Pilati, Daniela Amendola, si focalizza sulla formazione dei gruppi di studenti universitari per poter costruire attività collaborative efficaci. Gli autori propongono un software in grado di creare gruppi eterogenei di studenti, utilizzando tecniche di Machine Learning con i dati estratti da Moodle. Il contributo di Caterina Bembich si focalizza sulla Social Network Analysis (SNA) come settore dei LA, presentando un caso di applicazione ad un gruppo di studenti frequentanti una scuola professionale. La ricerca evidenzia un rischio di abbandono maggiore in presenza di reti sociali meno dense e meno coese o di reti sociali formate da studenti a loro volta a rischio abbandono. Il contributo di Valentina della Gala presenta i risultati preliminari di un'analisi, svolta con strumenti di Trattamento Automatico del Linguaggio (TAL), condotta su circa 50.000 testi scritti da docenti neo nominati in ruolo volti a riflettere su due attività didattiche svolte con gli studenti, nel contesto del percorso dell'anno di formazione e prova 2016/17. Il contributo di Giusi Antonia Toto analizza il rapporto fra LA e disabilità (e più in generale i bisogni speciali). L'esigenza metodologica nasce dalla necessità di raccogliere una grossa mole di evidenze quantitative mediante strumenti standardizzati per misurare gli apprendimenti nella pedagogia speciale, in particolare per accertare in quale misura l'accessibilità a corsi online contribuisca all'inclusione e al successo formativo degli studenti disabili. Il contributo di Luisa Pandolfi esplora alcuni punti di intersezione tra le strategie di intervento e di contrasto all'abbandono scolastico, la ricerca educativa e le attuali sfide poste dai LA, al fine di tracciare possibili direzioni di analisi e di indagine empirica per migliorare i processi di apprendimento e insegnamento, le pratiche ed i contesti educativi.

MOOC, didattica universitaria digitale e Learning analytics. Opportunità e prospettive

MOOC, digital university teaching and Learning analytics. Opportunities and Perspectives

Paolo Maria Ferri

University of Milano Bicocca,
Department of Human Sciences for Training "Riccardo Massa", paolo.ferri@unimib.it

This article has the aim to outline the current state of the art regarding the worldwide diffusion of MOOCs phenomenon. It is focused on the "disruptive" impact of MOOCs on the traditional University System in the most developed countries nations and on the way of using Learning Analytics to evaluate MOOCs quality, effectiveness and the learning value. The adoption of MOOCs as course material and the growing of MOOCs based degrees in many universities is changing dramatically the way of teaching and learning inside the Accademia. The article outlines also some methodological approach to use Learning Analytics to evaluate MOOCs quality, effectiveness and the educational value. In the first part we present, a review of the most recent studies on the MOOC phenomenon. In the second part we will deal with the characteristics and peculiarities of the MOOC phenomenon, with particular attention to the increasing number of graduated degrees with MOOC or Blended Learning format. In the third part we will analyze how the use of Learning Analytics can increase the quality of learning and teaching through the MOOC.

Keywords: MOOC; Learning technologies; University; Long Life learning; Learning Analytics

Questo articolo si propone l'obiettivo di tratteggiare uno stato dell'arte relativamente al fenomeno della diffusione dei MOOC. Si concentra sull'impatto "disruptive" dei MOOCs sistemi universitari tradizionali delle nazioni più sviluppate, così come sul modo di utilizzare la metodologia dei Learning Analytics per valutare la qualità dei corsi MOOC, l'efficacia e la valenza formativa. L'adozione dei MOOCs come materiali didattici e la crescita delle lauree basate su percorsi di MOOC sta cambiando drasticamente, in molte università, il modo di insegnare e apprendere. L'articolo delinea anche un approccio metodologico all'uso di Learning Analytics per valutare la qualità, l'efficacia e il valore didattico dei MOOCs. Nella prima parte verrà proposta una review dei più recenti studi relativi al fenomeno dei MOOC. Nella seconda parte ci si occuperà di tratteggiare le caratteristiche odierne del fenomeno MOOC, con una particolare attenzione al moltiplicarsi recente dell'accredimento di lauree "in formato" MOOC. Si tratta di un fenomeno che presenta caratteristiche altamente "disruptive" rispetto alle attuali metodologie e forme istituzionali dell'istruzione universitaria. Nella terza parte si analizzerà in che modo l'utilizzo dei Learning Analytics possa incrementare la qualità dei MOOC e permetterne una valutazione più efficiente e sostenibile.

Parole chiave: MOOC; Tecnologie dell'apprendimento; Università; Long Life Learning; Learning Analytics



1. La storia del successo dei MOOC dalle origini ad oggi: una ricostruzione bibliografica

1.1 *La nascita del fenomeno MOOC*



Negli Stati Uniti e negli altri paesi sviluppati a partire dal 2008 ha preso forma un nuovo modello di didattica universitaria che può essere considerata come un'evoluzione radicale dell'interno eco-sistema della formazione universitaria. A partire dal 2008 l'innovazione digitale, infatti, si è diffusa in maniera virale in tutte le più prestigiose Università mondiali grazie al fenomeno dei *Massive Open Online Courses*. Essi offrono, rispetto ai tradizionali corsi *blended*, percorsi di formazione aperti e disponibili on-line, spesso in forma gratuita, pensati per una formazione a distanza che coinvolga un numero elevato di utenti (Porter, 2015). Gli utenti possono essere sia studenti universitari sia professionisti in formazione ma anche persone interessate ad approfondire un determinato tema in modo "non formale". Si tratta di corsi, generalmente basati su *Virtual Learning Environment* (VLE) che contengono video, materiali e attività interattive (esercitazioni, casi studio, simulazioni) on-line sostenute dai corsisti con il supporto di tutor (*tutored, proctored*) o autonomamente (*self paced*). Possiamo distinguere alcune fasi del "ciclo di vita" dei MOOC analizzando il loro tasso di diffusione nel mondo. Sulla base dei dati dell'indagine annuale di *Class Central MOOC Report* e di altri autorevoli report di ricerca (Horigan, 2016), possiamo notare come nel 2016 si possa considerare conclusa la fase pionieristica della parabola di questa tecnologia dell'istruzione, e come tra il 2016 e 2019 si sia assistito all'esplosione globale del fenomeno (Sanchez-Gordon, Luján-Mora, 2018). Tra il 2019 e il 2020 sempre secondo le previsioni di *Class Central* e *Pew Internet & American Life Project* si può invece prevedere un assestamento nella crescita e un'evoluzione "a sistema" del fenomeno MOOC.

1.2 *La fase di esplosione virale dei MOOC*

Analizziamo sinteticamente la fase di esplosione virale e il progressivo consolidarsi delle piattaforme internazionali di MOOC hanno creato e che oggi "gestiscono" e "governano" il mercato di questa innovativa modalità di trasmettere, erogare, diffondere e certificare le competenze e il sapere. Secondo i dati citati tra il 2016 e il 2019 gli studenti universitari e i professionisti che hanno seguito un Mooc sono passati da dieci milioni a più di cento milioni. Anche le principali piattaforme

di MOOC sono cresciute esponenzialmente in notorietà e autorevolezza.

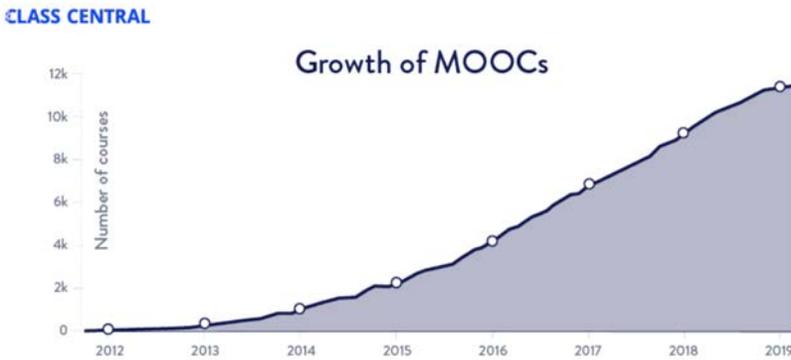


Fig.1: La crescita dei MOOC dal 2012 ad oggi Fonte Class Central 2018

Se passiamo poi ad analizzare l'offerta delle piattaforme, in maggioranza si tratta di corsi universitari erogati dai grandi consorzi statunitensi e da università non statunitensi che hanno stretto accordi di varia tipologia con questi consorzi. Ad esempio in Italia il Politecnico di Milano e l'Università Bocconi sfruttano questa opportunità con il consorzio Coursera, l'Università degli Studi Federico II di Napoli con EdX, mentre altre università si sono riunite in una rete autonoma è il caso del Network Eduopen (<https://learn.eduopen.org/>) promosso dall'Università degli Studi di Modena e Reggio Emilia, che raccoglie 17 Atenei. Eduopen ospita 247 MOOC, prodotti dalle 17 Università consorziate e ha più di 55.000 iscritti.

Consideriamo, però, la storia delle prime cinque piattaforme internazionali. *Coursera* è il più grande aggregatore di MOOC al Mondo: è un consorzio privato fondato da alcuni docenti dell'Università di Stanford in California (Mamgain, Sharma, Goyal, 2014). Anche *EdX* trova le sue origini al MIT e ad Harvard (Mamgain, Sharma, Goyal, 2014). A Stanford nasce il quarto consorzio mondiale di MOOC, *Udacity* (Rothe, Steier, 2017). I due soggetti non statunitensi che si trovano tra le prime cinque piattaforme sono la cinese *XuetangX*, fondata dall'Università di Tsinghua che ha stipulato un accordo di cooperazione con *EdX* (Yong, Guochang, Jianping, 2018) e il Consorzio *Future Learn*, britannico, fondato dalla Open University che al maggio 2018 ha incluso 143 partner britannici e internazionali, inclusi partner non universitari (Chua, Tagg, Sharples, Rienties, 2017).



1.3. Le caratteristiche innovative del fenomeno e le tipologie di MOOC

I MOOC hanno prefigurato fin dalla loro prima apparizione un nuovo modello di fruizione, organizzazione e gestione della formazione universitaria e *long life*. Già nel 2012 un articolo del New York Times definiva i MOOC il fenomeno *educational dell'anno* (Pappano, 2012) e negli anni successivi si sono sviluppati un grande mole di studi internazionali su questo fenomeno (Bozkurt et al., 2015; Pomerol e Al., 2015; Kim, 2015; Israel 2015; Suen, 2014). In particolare, i MOOC dal loro avvio sono stati concepiti utilizzati come strumento flessibile di formazione digitale che permette alle università e alle piattaforme di erogazione di perseguire obiettivi multipli oltre all'abbattimento dei costi:



- Aprire all'esterno, anche in chiave di marketing e di formazione permanente, le aule dei corsi curricolari con una forma di fruizione più adatta alle esigenze degli studenti lavoratori e alle necessità di riqualificazione dei *professional* (manager, insegnanti, infermieri, operatori sociali e della comunicazione ecc.);
- Estendere l'offerta formativa adattandosi alle necessità di formazione di un'utenza *professional* che si caratterizza sempre di più per la sua mobilità e per il suo essere perennemente on-line (Floridi, 2017).
- Creare dei percorsi di orientamento e allineamento delle competenze progettati per permettere agli studenti di comprendere a fondo le caratteristiche e la specificità e gli sbocchi professionali di un determinato corso di studi.

Volendo classificare le principali tipologie di MOOC che sono apparse nella prima fase di sviluppo essi possono essere distinti in:

- corsi universitari MOOC on-line (solamente a distanza), molto vantaggiosi da un punto di vista economico, inizialmente corsi spesso *self paced* o con un tutoraggio per lo più organizzativo ma con contenuti di qualità molto elevata. Ad esempio, *Georgia Tech* (una delle migliori Università americane) insieme a *EdX* decise di offrire un corso magistrale nel 2013 in Computer Science (*On line Master in Computer Science - OMCS*) a 6.600 Dollari per anno -circa un decimo del costo dei suoi corsi standard. Si trattò di un'esperienza pionieristica ma di grande successo dal momento che, ad oggi, ha registrato una crescita notevole. L'*On line Master in Computer Science* oggi vanta una comunità studentesca di 6.365 studenti (iscrizioni alla primavera 2018) che rappresentano oltre 100 paesi. Inoltre il curriculum *On Line Master Degree* vanta 28 corsi, con molti altri

in produzione (<http://www.omscs.gatech.edu/explore-oms-cs>). Questa politica, quindi, è stata seguita anche da molte altre università statunitensi delle Ivy League.

- Corsi universitari *blended* in cui i docenti titolari dei corsi usano i MOOCs come supporto didattici multimediali riducendo il numero delle lezioni frontali che devono svolgere integrandole, però, con esercitazioni sul campo tutorate on-line. Questa è la soluzione che impone cambiamenti relativamente contenuti all'organizzazione della didattica e potrebbe consentire di offrire agli studenti lavoratori percorsi personalizzati di apprendimento (Sancassani et al. 2014). Si tratta di una soluzione adottata in molti atenei e cui fa riferimento un ricca letteratura internazionale (Israel, 2015; Brali, Divjak, 2018) in Italia ad esempio molte delle università del consorzio *Eduopen* utilizzano questo modello.
- Corsi di studio che mixano corsi MOOC e in presenza. In questa tipologia gli studenti hanno a disposizione una libreria di MOOC verranno seguiti (sotto la guida dei docenti) oltre ai corsi che vengono invece erogati in forma tradizionale dai docenti interni all'Università o al Dipartimento. Questa soluzione distingue decisamente il ruolo dei docenti dell'Università responsabili di un corso di studi e quelli che producono i MOOC, rendendo i programmi di studi flessibili anno per anno.



1.4 Quali le ragioni del successo dei MOOC e quali i principali trend di sviluppo?

Per rispondere a queste domande è necessario, in primo luogo, notare come queste nuove forme di didattica universitaria si siano sviluppate così rapidamente ed in modo così esteso a livello globale a causa di un insieme di fattori che sono stati ben individuati da Mauro Scalise nel suo paper *MOOC sfide e opportunità* (Scalise, 2018, p. 11). Li esporremo, qui, integrandoli e liberamente interpretandoli. Secondo Scalise i MOOC hanno avuto un'esplosione virale così massiccia in tutto il mondo - ed in particolare nel mondo universitario anglosassone - per le seguenti ragioni:

- La necessità da parte di tutti gli atenei del mondo – in un contesto di rapida globalizzazione e rapida obsolescenza dell'offerta di didattica – di ottimizzare i costi e cercare di attrarre nuovi studenti. Studenti che spesso provengono dal segmento in espansione della *life-long-education*, cioè della riqualificazione dei *professional* che operano in aziende *knowledge intensive*.

- La crescita di nuove forme di apprendimento sociale *bottom-up* legate più in generale al fenomeno dell’“accesso universale” reso possibile dal web e più in specifico alle dinamiche proprie del Web 2.0 e dei social network che hanno generato un “cultura partecipativa” (Jenkins, 2009), tra gli utenti, oggi abituati ad un maggior protagonismo dal basso nella partecipazione ad occasioni formative e di discussione dal basso. E’ da notare, inoltre, che le dinamiche della “cultura partecipativa” dei nativi digitali sono particolarmente coerenti con lo stile di progettazione e di fruizione di molti MOOC, spesso di impronta costruttivista e connettivista. Nei MOOC, ad esempio, vengono lasciati alla discussione tra pari ed al *peer tutoring*, ruoli tradizionalmente ricoperti dai docenti e dai tutor di percorso (Suen, 2014).
- La necessità di adeguare i contenuti e le metodologie anche della formazione universitaria ai mutati stili di apprendimento e ai linguaggi della generazione dei “nativi digitali”; una necessità che si fa sempre più pressante dal momento che oggi entrano all’università i nati nel 2001 e 2002,



È necessario ora chiarire quali sian i principali trend di sviluppo dei MOOC oggi (Shah, 2019; Johnson, 2019). Tra i trend in crescita ne desideriamo evidenziare, in particolare, tre.

L’istituzionalizzazione delle lauree basate su un’offerta MOOC. Si può affermare come il grande cambiamento del 2018-2019 è stato quello della “istituzionalizzazione” dei diplomi basati su MOOC. Nel 2017 negli USA solo sette università annunciavano 15 titoli di studio basati a vario titolo sui MOOC, nel 2018 si sono aggiunte altre 30 università, che hanno lanciato più di 45 lauree e il trend è sicuramente in crescita anche nel 2019. Ad esempio, *Coursera* si è mosso nella direzione di proporre un sistema di crediti per ciascun MOOC frequentato che possa portare al conseguimento di una laurea o oppure ad una certificazione che possa permettere al corsista di dimostrare all’azienda per cui lavora il conseguimento di nuove qualifiche. Un sistema che si avvicina molto a quello europeo degli *Open Badge for education*.

La concentrazione in oligopoli delle piattaforme di distribuzione dei MOOC. E’ un fenomeno molto evidente se si prendono in considerazione i dati del 2018 di elaborati da Reda e Cerr per Classe Central (Kerr, Reda, 2018):

2. Come valutare la qualità dei MOOC: gli analytics e le loro fonti

2.1 *I learning analytics cosa sono e che cosa servono?*

Si tratta di una domanda che non ammette ancora risposte definitive e il dibattito tra gli studiosi è vivace. La discussione riguarda le modalità per valutare il successo formativo e la qualità dei MOOC. In effetti, le interazioni all'interno dei MOOC sono davvero "altre" rispetto a quelle della formazione in presenza, dal momento che nei MOOC le relazioni sono quasi tutte "mediate" dalla comunicazione digitale. Per descrivere il fenomeno potremmo tratteggiarlo come l'esponentiale evoluzione delle interazioni che caratterizzano le pratiche della didattica cooperativa che si svolge all'interno dei *Virtual Learning Environment* (VLE) utilizzati in ambito dei progetti di *Blended learning* scolastici o professionali. Nel caso dei MOOC il docente è per lo più o solamente un "progettista didattico" e un facilitatore dell'apprendimento (Laurillard, 2012). Spesso il docente non svolge nemmeno le funzioni di tutor che sono vicariate da altro personale formato specificamente e/o dal gruppo dei pari. Proprio, per questo, per valutare la qualità dell'apprendimento all'interno dei MOOC diventa strategico il "tracciamento" delle attività del corsista all'interno della piattaforma. La valutazione della qualità dell'apprendimento è divenuta da tempo centrale, si tratta di una variante "educational" della più ampia problematica dei *Big Data*. Il tema dei *Learning Analytics*, e la definizione stessa di questo termine è comparso per la prima volta nella *Call for Papers* della prima conferenza internazionale su *Learning Analytics and Knowledge* (LAK 2011) organizzata dalla *Society for Learning Analytics Research* (Ferguson, 2014) ed è stato ripreso nella seconda conferenza 2012 (Buckingham Shum, Gasevic, Ferguson, 2012), in questo contesto i *Learning Analytics* essi sono stati definiti in questo modo: "Con Learning Analytics ci si riferisce alla misurazione, alla raccolta, all'analisi e alla presentazione dei dati sugli studenti e sui loro contesti, ai fini della comprensione e dell'ottimizzazione dell'apprendimento e degli ambienti in cui ha luogo." (Ferguson, 2012, p.139). Ora si pensi alla mole sterminata di dati che popolano le piattaforme dei principali provider di MOOCs, ma è notevolissima anche quella che si genera all'interno di una singola aula virtuale dei VLE. Se consideriamo i più di cento milioni di utenti di MOOC, nel 2019, e proviamo a immaginare il numero di accessi (log), dati personali, oggetti digitali e materiali consultati, esercitazioni svolte, post nei forum, materiali scaricati e di mail e post con i tutor e i pari si può immaginare le migliaia di Gigabyte di dati a cui facciamo riferimento. Al netto dell'applicazione ancora sperimentale dell'Intelligen-



za Artificiale all'analisi di questi dati (Ahmad et al. , 2018), la comunità dei ricercatori sta ancora cercando di comprendere in che modo analizzare, standardizzare e rendere fruibili una tal quantità di informazioni, certo difficilmente maneggiabile, ma che sicuramente costituisce un miniera per la ricerca e per il miglioramento della qualità dell'offerta di MOOC. A partire dal 2012 gli studi sugli *LA educational* si sono progressivamente discostati dall'approccio dell'*Educational Data mining* (EDM). Rispetto all'approccio EDM essi puntano maggiormente sul ruolo interpretativo dei ricercatori rispetto alla automazione nell'analisi dei dati.. Si mette, cioè, una maggior enfasi sul processo educativo e di apprendimento inteso come un insieme unitario di processi piuttosto che sull'analisi e valutazione dei singoli step di questo processo o delle fasi dell'apprendimento. Il focus degli studi sui LA è, infatti, concentrato sulle possibilità di miglioramento del sistema formativo e sull'*empowerment* delle competenze degli studenti e dei docenti, più che sulla valutazione di questi (Gaeta, Marzano, Miranda, Sandkuhl, 2016; Miranda, Marzano, Lytras, 2017). Conseguentemente i metodi di analisi utilizzati hanno una curvatura maggiormente qualitativa e tra questi possiamo annoverare oltre alla Social Network Analysis, l'analisi dei *Sentiment*, quella del discorso e i modelli di analisi *Sense Making* (Siemens, Barker, 2012).



2.2 I Learning Analytics: cosa indagano?

La comunità dei ricercatori nel campo degli LA, come accennavamo, pone al centro non solo la valutazione o le *performace*, ma studia le relazioni interpersonali che si svolgono all'interno delle piattaforme dei MOOC che vengono intese come "piattaforme sociali". All'interno dei VLE l'analisi dei dati si concentra sulle seguenti tipologia di azioni, interazioni e "oggetti digitali":

- *discourse analytics* : il linguaggio è il primo strumento per la costruzione e lo sviluppo dell'apprendimento per cui l'analisi del discorso e dei contenuti testuali postati all'interno delle piattaforme è di grande importanza;
- *content analytics*: sono studiati, cioè, i contenuti generati dagli utenti all'interno delle piattaforme del Web 2.0 e le interazioni che essi producono, così come le interazioni dei corsisti con gli oggetti predisposti all'interno della piattaforma dai docenti;
- *disposition analytics*: cioè quegli indicatori per lo più qualitativi ma anche quantitativi che possono misurare la motivazione intrinseca

all'apprendimento del corsista ma anche del docente e dei tutor attraverso i loro log, le loro interazioni e lo *scaffolding* e la valutazione anche motivazionali che esprimono.

- i *context analytics*, i fattori di contesto: dove, come, quante volte e a che ora docenti e studenti entrano in contatto con la piattaforma che ospita il MOOC e tra loro, con quali device e con quali connessioni. Ad esempio la fruizione dei percorsi di apprendimento in mobilità attraverso smartphone o tablet o altri device che possono cambiare in maniera significativa le modalità di accesso ai saperi.

2.3 Il ciclo di vita dei Learning Analytics e i suoi limiti

Ora proviamo a delineare il “ciclo di vita” degli LA utilizzando le riflessioni su questo tema di Mohammad Khalil e Martin Ebner (2015). Proviamo, cioè, a descrivere opportunità e problematiche legate all'utilizzo dei *Learning Analytics* nella valutazione dell'efficacia della formazione dei MOOC, secondo lo schema dei due autori presentato qui sotto.

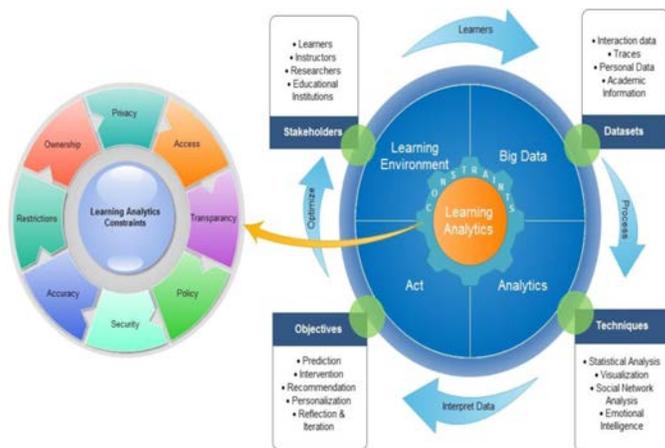


Fig. 3. Opportunità e criticità degli LA (Khalil, Ebner, 2015, p. 2)

Il flusso di creazione e utilizzo dei MOOC potrebbe essere il seguente: a. Studenti, docenti, ricercatori ed istituzioni educative rappresentano gli stakeholders delle piattaforme MOOC. Le istituzioni educative (docenti, ricercatori e tecnici) progettano e generano l'offerta formativa che viene allestita all'interno del VLE. Gli studenti insieme a loro danno vita al MOOC, attraverso la loro iscrizione e la condivi-

sione della loro identità digitale. b. Insieme (docenti, tutor e studenti) generano i set di dati che abbiamo descritto più sopra: informazioni per la fruizione del MOOC o accademiche, contenuti (testi, video, podcast, presentazioni, info-grafiche, bibliografie), interazioni, percorsi di navigazione, dati personali, valutazioni individuali e di gruppo. Questi dati vengono sottoposti a differenti tipologie di metodologie di indagine, da quelle statistiche e quelle di clima, a quelle motivazionali, alla *social network analysis*, in questo modo la progettazione dei successivi MOOC sia ottimizzata in maniera incrementale. I risultati delle analisi permettono, cioè, di migliorare l'offerta formativa sulla base di un insieme di analisi statistiche e qualitative che prevedono la cooperazione e non la concorrenza tra i differenti *stakeholders*. d. Le analisi si traducono, poi, in interventi, raccomandazioni, linee guida e ridefinizioni dell'offerta formativa che permettono un miglioramento sia di quest'ultima che degli out-come formativi degli e per gli studenti. Tutto questo non può che avvenire, però, tenendo conto dei molti vincoli legati sostanzialmente alla *privacy* (GDPR) e all'utilizzo dei dati personali dei corsisti, perché nessuno si augura che scoppino casi simili a quello di *Cambridge Analytica* anche all'interno dei grandi Consorzi MOOC. I vincoli legati all'utilizzo degli LA riguardano, anche le norme, nazionali e internazionali, sulla proprietà intellettuale, la sicurezza dei dati stessi e la trasparenza nel loro utilizzo.



2.4 Conclusioni

Questa analisi di secondo livello sulla tematica dei MOOC e dei Learning Analytics ci ha permesso di provare a ricostruire come – anche nella didattica universitaria “aumentata” digitalmente – il “mega-cambiamento” introdotto dalle tecnologie stia trasformando le regole del gioco e della competizione internazionale tra gli Atenei, questo avviene almeno a quattro livelli.

- In primo luogo, lo sviluppo tumultuoso dei MOOC e delle forme di didattica digitale è dovuto alla progressiva trasformazione degli stili di fruizione mediale, di comunicazione e di apprendimento dei giovani “nativi digitali” che sono nati in un mondo già abbondantemente *wired*.
- In secondo luogo, la possibilità di utilizzare piattaforme interattive per l'apprendimento ha progressivamente globalizzato lo scenario della formazione universitaria abbattendone nel contempo i costi. C'è da chiedersi quanto a lungo potranno resistere i sistemi nazio-

nali universitari della formazione superiore. Il problema è rilevante perché è evidentemente come l'attrattiva di lauree conseguite in una grande e prestigiosa università internazionale – anche se “aumentata digitalmente” – può rappresentare per lo studente di un paese come l'Italia una grande tentazione soprattutto in vista delle opportunità che può offrire sul mercato del lavoro.

- La considerazione precedente è confermata dai dati che abbiamo presentato relativi al progressivo istituzionalizzarsi dei MOOC. Negli ultimi due anni si è assistito ad una notevole crescita dei Corsi di Laurea erogati in questa modalità, con il progressivo accreditamento anche di corsi di formazione continua e Master anche in accordo con grandi imprese multinazionali.
- È sempre più diffusa anche la transizione al digitale dei cosiddetti corsi di “formazione permanente”. Una volta limitati ad un ristretta élite di lavoratori della conoscenza o di professionisti, oggi i percorsi di *long life learning* non possono che rivolgersi a tutta la platea dei lavoratori e i MOOC, anche utilizzati come supporto per corsi Blended, sono la soluzione più economica ed efficiente in questo mercato.
- I Learning Analytics, cioè le tracce delle interazioni on-line tra persone (docenti, studenti, tecnici) e quelle con i materiali predisposti negli ambienti virtuali per la l'apprendimento e gli scambi sociali tra pari, tutor e docenti, possono costituire un enorme messe di dati che se correttamente utilizzati possono costituire percorso di monitoraggio e miglioramento continuo nelle qualità dell'insegnamento.



Riferimenti bibliografici

- Ahmad R., Kamarul A. J., Bushrah B., Noraisikin S., Miftachul H., Andino M. (2018). Demystifying Learning Analytics in Personalised Learning. *International Journal of Engineering & Technology*, 7 (3), 1124-1129.
- Bozkurt A., Akgun-Ozbek E., Onrat-Yilmazer S., Erdogdu E., Ucar H., Guler E., Sezgin S., Karadeniz A., Sen N., Goksel-Canbek N., Dincer G.D., Ari S., Aydin C.H. (2015). Trends in Distance Education Research: A Content Analysis of Journals 2009- 2013. *International Review of Research in Open and Distributed Learning*, 16 (1), 330-363.
- Brali A., Divjak D. (2018). Integrating MOOCs in traditionally taught courses: achieving learning outcomes with blended learning. *International Journal of Educational Technology in Higher Education*, 15, 1-13.
- Buckingham Shum S., Gasevic D., Ferguson R. (Eds.). (2012). *LAK '12: Proceedings of the 2nd International Conference on Learning Analytics and Knowledge*.

Vancouver, BC, Canada - April 29 - May 02, 2012. New York, NY, USA: ACM.

- Crisp G., Oliver B. (2019). Re-imagining graduate achievement and employability. In J. Higgs et al. (Eds.), *In Education for Employability: Learning for Future Possibilities* (Vol. 1, pp. 73-82). Leiden: Brill Sense.
- Chua S., Tagg C., Sharples M., Rienties B. (2017). Discussion Analytics: Identifying Conversations and Social Learners in FutureLearn MOOCs. In Vigenini Wang, Yuan, Paquette, Urrutia, Le'on (eds.), *MOOC analytics: live dashboards, post-hoc analytics and the long-term effects* (pp. 36-62). CEURWS.org.
- Ferguson R. (2014). Learning analytics: fattori trainanti, sviluppi e sfide. *TD Tecnologie Didattiche*, 22(3), 138-147.
- Floridi M. (2017). *La quarta rivoluzione. Come l'infosfera sta trasformando il mondo*. Milano: Raffaello Cortina.
- Gaeta M., Marzano A., Miranda S., Sandkuhl K. (2016). The competence management to improve the learning engagement. *Journal of Ambient Intelligence and Humanized Computing*, 6, 151-164.
- Horrigan J.P. (2016). *Lifelong Learning and Technology*. *Pew Internet & America Life*. Disponibile al sito, <http://www.pewinternet.org/2016/03/22/lifelong-learning-and-technology/>.
- Israel M.J. (2015). Effectiveness of integrating MOOCs in traditional classrooms for undergraduate students. *International Review of Research in Open and Distributed Learning*, 16 (5), 102-118.
- Jenkins H. (2009). *Confronting the Challenges of Participatory Culture. Media Education for the 21st Century*. Cambridge: MIT Press.
- Johnson S. (2019). *Much Ado About MOOCs: Where Are We in the Evolution of Online Courses?* Disponibile al sito: <https://www.edsurge.com/news/2019-02-26-much-ado-about-MOOCs-where-are-we-in-the-evolution-of-online-courses>.
- Kerr R., Reda V. (2018). *MOOCs as Institutional Internationalization Strategy First Italian courses on edX platform*. Madrid: Proceedings of EMOOCs 2019: Work in Progress Papers of the Research, Experience and Business Tracks (pp. 94-97).
- Khalil M., Ebner M. (2015). Learning Analytics: Principles and Constraints. In *Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications* (pp. 1326-1336). Chesapeake, VA: AACE.
- Kim P. (2015). *Massive Open Online Courses: The MOOC Revolution*. London: Routledge.
- Laudrillard D. (2012). *Teaching as a Design Science Building Pedagogical Patterns for Learning and Technology*. New York: Routledge.
- Lord M. (2019). Show what they Know. *ASEE Prints*, 28 (7), 34-37.
- Mamgain N., Sharma A., Goyal P. (2014). Learner's perspective on video-viewing features offered by MOOC providers: Coursera and edX, Proceedings of IEEE International Conference on MOOC, Innovation and Technology in Education (MITE), pp. 331-336.
- Miranda S., Marzano A., Lytras D. (2017). A research initiative on the construction of innovative environments for teaching and learning. Montessori and



- Munari based psycho-pedagogical insights in computers and human behavior for the 'new school'. *Computers in Human Behavior*, 66, 282-290.
- Pappano L. (2012). The Year of the MOOC. *The New York Times*. Disponibile al sito <https://www.nytimes.com/2012/11/04/education/edlife/massive-open-online-courses-are-multiplying-at-a-rapid-pace.html?pagewanted=all>
- Pomerol J. Ch., Epelboin Y., Thoury C. (2015). *MOOCS: Design, Use and Business Models*. New York: ISTE Wiley.
- Porter S. (2015). *To MOOC or Not to MOOC: How Can Online Learning Help to Build the Future of Higher Education?* Waltham, MA: Chandos Publishing.
- Rothe H., Steier F. (2017). Shaping the Boundaries of a Service Ecosystem: The Case of Udacity. In *Proceedings of the 50th Hawaii International Conference on Sys(1), 1-4. tem Sciences*. Hawaii International Conference on System Sciences.
- Sancassani S., Casiraghi D., Corti P., Trentinaglia N. (2014). MOOC, OER e l'approccio "flipped classroom": due case study di transizione in ambito scolastico e aziendale. *Form@are*, 14, (1), 1-4.
- Sanchez-Gordon S., Luján-Mora, S. (2018). *Lifecycle for MOOC Development and Management*. Harshley, NY.: IGI Publishing.
- Scalise M. (2018). *MOOC sfide e opportunità*. Piano Nazionale Università digitale. Disponibile al sito, https://www.crui.it/images/allegati/pubblicazioni/2015-/MOOC_2015.pdf.
- Shah D. (2019). *Year of MOOC-based Degrees: A Review of MOOC Stats and Trends in 2018, Class Central Report*. Disponibile al sito <https://www.classcentral.com/report/MOOCs-stats-and-trends-2018/>
- Siemens G., Barker R. (2012). Learning analytics and educational data mining: towards communication and collaboration. In S. Buckingham, D. Gasevic, R. Ferguson (Eds.), *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge*. New York: ACM.
- Suen H.K. (2014). Peer assessment for massive open online courses (MOOCS). *The International Review of Research in Open and Distributed Learning*, 15 (3).
- Yong L., Guochang Z., Jianping Li (2018). *Comparing the Chinese University MOOC Platform to the Three Major MOOC Players*. Proceedings of the 2nd International Conference on Computer Science and Application Engineering. New York: ACM.



Analysis of completion and dropout rates in EduOpen MOOCs

Analisi di tassi di completamento e abbandono nei MOOC di EduOpen

Katia Sannicandro

University of Modena and Reggio Emilia (Italy), katia.sannicandro@unimore.it

Annamaria De Santis

University of Modena and Reggio Emilia (Italy), annamaria.desantis@unimore.it

Claudia Bellini

University of Modena and Reggio Emilia (Italy), claudia.bellini@unimore.it

Tommaso Minerva

University of Modena and Reggio Emilia (Italy), tommaso.minerva@unimore.it

The completion rate of massive open online courses (MOOCs) is generally less than 10% of participants. This is due to several factors, many of which cannot be eliminated: spontaneous enrolment, participants' extreme heterogeneity, self-regulated processes and differences in motivational and cultural profiles. One of the factors that can affect the rate of completing a MOOC is the modality of delivery. The active presence of the teacher and of other support figures in MOOCs, even where criticality is linked to the number of students and the management of the dynamics present in the online learning environment, can qualitatively and quantitatively affect both the levels of interaction and participation of the users and the completion percentages of the course itself. The MOOCs published on the EduOpen Portal provide two specific methods of use: self-paced and tutoring. The choice of modality, which is defined in the design phase, "impacts" the structure and timing of the course itself, its learning objectives and the types of teaching resources. Consequently, the levels of interaction and evaluation processes are also "calibrated" in relation to the "presence or absence" of support figures in the online environment. The contribution, starting from the first data generated by the Learning Analytics system of the Portal, focuses on analysis of the percentage of the completion/ dropout rate recorded for the entire group of MOOCs published in relation to the delivery methods defined in the design phase of the various courses. In July 2019 there were 247 courses in the catalogue with more than 55,000 users. The final objective of the analysis is to include in the guidelines for the design of a MOOC the results of this first study.

Keywords: MOOCs; learning analytics; self-regulated processes; instructional design; dropout rate

Il tasso di completamento di MOOCs è generalmente inferiore al 10% degli iscritti. Questo a causa di diversi fattori, molti non eliminabili, quali: reclutamento spontaneo, estrema eterogeneità degli iscritti, processi di autoregolazione, differenze nei profili motivazionali e culturali. Uno dei fattori che può incidere sul tasso di completamento di un MOOC è rappresentato dalla modalità di erogazione. La presenza attiva del docente e di altre figure di supporto in corsi MOOCs, se pur con le evidenti criticità legate alla numerosità degli studenti e alla gestione delle dinamiche presenti dall'ambiente di apprendimento online,



può incidere (qualitativamente e quantitativamente) sia sui livelli di interazione e partecipazione degli utenti sia sulle percentuali di completamento del corso stesso. I MOOCs pubblicati sul Portale EduOpen prevedono nello specifico due modalità di fruizione: autoapprendimento e tutorata. La scelta della modalità - definita in fase progettuale - "im-patta" sulla struttura e sulle tempistiche stesse del corso, sugli obiettivi di apprendimento e sulla tipologia delle risorse didattiche. Di conseguenza, i livelli di interazione e i processi di valutazione sono "calibrati" anche in relazione "alla presenza o all'assenza" di figure di supporto nell'ambiente online. Il contributo, a partire dai primi dati generati dal sistema di Learning Analytics del portale, si focalizza sull'analisi delle percentuali di completamento/tasso di abbandono registrate sull'intero insieme di MOOCs pubblicati in relazione alle modalità di erogazione definite nella fase di progettazione dei vari corsi. A luglio 2019 i corsi presenti nel catalogo sono 247 con un numero di utenti superiore a 55000 utenti. L'obiettivo finale dell'analisi e quello di includere nelle linee guida alla progettazione dei MOOCs i risultati emersi da questa prima ricerca.

Parole chiave: MOOCs; learning analytics; processi di autoregolazione; progettazione didattica; tassi di abbandono.



1. Introduction

Recent growing academic interest in massive open online courses (MOOCs) – described as a disruptive technology that challenges traditional educational models (Bozkurt et al., 2017; Yu et al., 2017) – is linked to their ability to influence and increase the spread of higher education provision (Jung, Lee, 2018), to foster the flexibility of the learning process and to provide access to “disadvantaged” students.

However, to be successful in a MOOC environment, greater literacy (and not merely digital) is often required (Jordan, 2014). As described in the latest report published by the International Council for Open and Distance Education (ICDE) entitled *Global quality in online, open, flexible and technology enhanced education*, strengths and weaknesses related to the spread of technology-enhanced learning and digital learning environments coexist. Concerning the European context, what emerges from the report can be summarised in three points:

1. “Within European universities, digital learning environments maintain a strong presence and there seems to be more acceptance related to the value of learning in these modalities.
2. The development of blended and online learning does not always appear to be developed through a systematic approach. Instead development may rely on the interest and commitment of individuals resulting in slow and limited implementation.
3. There is a need to build competence and expertise in blended and online learning design by offering professional development on re-

levant topics. However, there may be challenges within academic environments where the culture does not encourage innovation” (p. 8).

Common points, both in Europe and internationally, are linked to the quality and construction of a “quality framework”, to professional development in terms of strengthening the skills of teachers and students, and finally to the social perception linked to the relationship between distance and traditional learning (ICDE, 2018).

With respect to this scenario, the spread of MOOCs has contributed to a wider “awareness and acceptance of the added value of blended and online education” (ICDE, 2018, p. 32). However, as anticipated, there are still forms of “stigma” and critical issues related to the “quality” of content and educational design, and in particular to high rates of abandonment. In many cases, the completion rate of MOOCs is generally less than 10% of participants (Jordan, 2014, 2015; Onah et al., 2014; Bozkurt et al., 2017) due to several interrelated factors, many of which cannot be eliminated, for example spontaneous recruitment, the extreme heterogeneity of learners, self-regulated processes and differences in motivational and cultural profiles. One of the factors that may affect the completion rate of a MOOC is how it is delivered. The active presence of the teacher and other support figures in MOOCs, even where the criticalities are linked to the number of students and the management of the dynamics present in the online learning environment, can influence both the levels of interaction and participation of users and the completion rates of the course itself.

Numerous studies have investigated the relationship between tutoring systems, student motivation, support tools/resources present in training courses and training success (Tait, 2003; Loizzo et al., 2017; Khalil, 2014). One example is the Open University, which since its inception in 1969 has established a system of support/tutoring for students: a personal tutor follows and works with a group of students not exceeding 25 people (Tait, 2003). Other studies and research underline the importance of the teacher’s presence in relation to significant effects recorded in terms of “learning commitment”; however, “to increase learners’ participation in MOOCs, instructor-centered learning activities should focus on supporting learners with feedback and having the learning contents well organized” (Jung, Lee, 2018, p. 19). This result is also explained by voluntary participation in these courses, “if learners do not experience presence during the course, their participation and involvement become lower, and the probability of dropping out would be higher” (p. 19). Jung and Lee (2018) define the concept of teaching presence “as the degree to which learners perceive that instructors fa-



ilitate learning by designing and organizing content and supporting them” (p. 11).

The relationship between dropout levels and “support” offered to students has an impact not only in relation to the learning processes, but also regarding the possibility of strengthening the student’s sense of self-efficacy, self-esteem and motivation (Kizilcec & Schneider, 2015), as well as the levels of completion of planned training activities (Tait, 2003).

Owing to the numerous variables involved, if the first experiments related to the development of models and intelligent tutoring systems (even in open mode) that provide for the presence of integrated systems by which to structure teaching materials, with the possibility of receiving feedback from the teacher and other support figures, research in recent years has focused on the development of “predictive” tools that exploit recent systems of learning analytics (LA) with the ability to act in “real time” and with direct repercussions on the “process” of designing the courses.



2. Instructional Design of the EduOpen MOOCs

Numerous researches have deepened the link between course design and dropout rates (Yousef et al., 2014; Margaryan et al., 2015; Kim, 2016). According to Yousef and colleagues (2014), despite a large number of criteria available for successful design, not all models can be used in the context of MOOCs due to their unique characteristics.

According to Margaryan and colleagues (2015), if the learning experiences are fundamental, the quality of the didactic design of a course is also a critical indicator and prerequisite of the potential of the course with respect to the effectiveness of learning in terms of course completion. A “unique format” in the design of MOOCs can negatively affect the completion of courses and the learning process owing to the different backgrounds of students (Onah et al., 2014).

As far as the design and production of EduOpen¹ MOOCs is concerned, the starting point is the “EduOpen Guidelines”. The workflow for producing an EduOpen MOOC can be simplified in the following actions (EduOpen Guidelines):

¹ EduOpen Portal (<https://learn.eduopen.org>).

1. Presentation of the general project (title only, general description, objectives, certification structure and duration estimate);
2. Educational Macro-Design;
3. Educational Microdesign;
4. Recording of video content;
5. Production of teaching materials;
6. Structuring of the course and/or pathways on the EduOpen Portal;
7. Technical validation and quality standards;
8. Validation by the working group;
9. Publication of the course and/or pathways on the EduOpen Portal;
10. Delivery of the course and/or pathways (course planner);
11. Ex-post evaluation of the course.

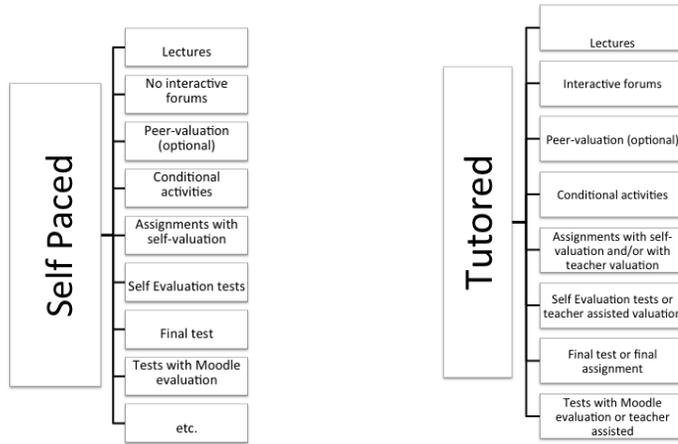
With the drafting of the guidelines for the design of MOOCs and the development of the online learning environment – the result of the joint effort of the entire network of EduOpen – we set ourselves the goal of supporting teachers and other professionals involved in the design of training content while also facilitating the process of learning and act on self-regulated learning. Good design is essential to “accompany” and “orient” the user in the choice of courses and towards the completion of the training course undertaken. This last aspect increases its importance when the MOOC offer opens the user up to certified courses linked to academic courses and which may involve the issue of university credits (CFU).

The process of hybridisation between “formal” training paths and open and “informal” environments involves a rethinking of previous models, as described by García-Peñalvo (2018): “it is necessary to design specific technological frameworks for the MOOC context to take advantage of the massification, diversity, and multiculturalism they present; generate new pedagogical approaches” (p. 1020). Previous research (De Santis et al., 2019) has shown that EduOpen MOOC participants pursue different learning objectives, as “EduOpen learners are to be found in a personal training needs and curiosity/interest in the topics of the courses” (p. 363), as confirmed by other studies that show that MOOCs participants not only have different academic objectives, but present extremely heterogeneous age groups and educational levels (García-Peñalvo et al., 2018). Some studies have focused attention precisely on the link and the ways in which the “instructional conditions influence the prediction of academic success” (Gašević et al., 2016, p. 70). The same author suggests “that it is imperative for learning analytics research to account for the diverse ways technology is adopted



and applied in course-specific contexts” (Ivi, p. 81), with a particular attention for learning design.

With respect to this complex scenario, the choice made at the design stage of the course delivery method cannot be random, because it “impacts” on the structure and timing of the course itself, the learning objectives and the type of teaching resources. Consequently, the levels of interaction and evaluation processes are also “calibrated” in relation to the “presence or absence” of support figures in the online environment and in close relation to the life cycle of MOOCs (Fig. 1).



Tutored LifeCycle

T0: Enrollment start date
 T1: Course open date
 T2: Active Tutoring start date
 T3: Active Tutoring stop date
 T4: Soft tutoring or SelfPaced start date depending on instructional design
 T5: Enrollment stop date
 T6: Course close date
 T7: Next edition Course Open date

Self Paced LifeCycle

T0: Enrollment start date
 T1: Course start date
 T5: Enrollment close date
 T6: Course close date >> Archived;

Fig. 1: Course life cycle

The design guidelines also apply to the development and design of the pathways, a sequence of courses that define a single set of training objectives and ending with a capstone course (Fig. 2). The capstone

(final course of the pathway) includes the final activities associated with the entire path, for example the final evaluation of all the individual MOOCs that make up the pathway, which may also be linked to Master's and advanced courses and so forth. Specifically, the EduOpen MOOCs provide two modes of use: self-paced and tutoring.

The screenshot shows the interface for a pathway on the University of Modena and Reggio Emilia's EduOpen platform. The main title is "Enabling and rehabilitating approach to sensory disabilities" under the "Nano Master" title. A progress indicator shows 0% completion. A green button labeled "In corso" (In progress) with "Iscriviti adesso" (Enroll now) is visible. The course details include: Language: Italian; Category: Medicine and Health; Credits: 150; Objective: Formation of Teachers, Lifelong Learning; Frequency: Free; Enrollment fee: 2,000 EUR. A list of courses is shown under the "Corsi" tab, including "Apprendimenti scolastici e disabilità sensoriali", "Neuropsicologia dello sviluppo", "Didattica speciale e approccio alla LIS", "Approccio al trattamento riabilitativo delle disabilità uditive e visive", and "[CAPSTONE] Approccio abilitativo e riabilitativo alle disabilità sensoriali". Each course entry includes a thumbnail, the university name, course title, duration, start date, and status.

Fig. 2: Example of the structure of a pathway on EduOpen

In the tutoring modality, the courses provide for a more structured temporal scan of the training activities, and the course life cycle provides in most courses for the indication of a pre-established date for the conclusion of the activities and a closing date of the MOOCs (Fig. 3). After the closing of the course, it is possible for the students to access the didactic resources, but it is not possible to obtain the certificate of participation and consequently consider the course completed. In addition, in this mode the presence of support figures (the same teacher of the course or tutor) with whom one can interact and ask questions related

to the topics covered in the MOOCs is offered. In the self-paced mode there are neither tutors nor the possibility to interact or ask questions to the teacher of the course, but there are still discussion forums (news forums, thematic forums and so on) in which participants can interact with one another or receive communication from the referents of the technical support of the Portal (example alerts on the closing date of the course, etc.). Moreover, in this mode the course calendar (Fig. 3) may not include at the time of opening the MOOCs indication of a default closing date of the course and training activities.

34

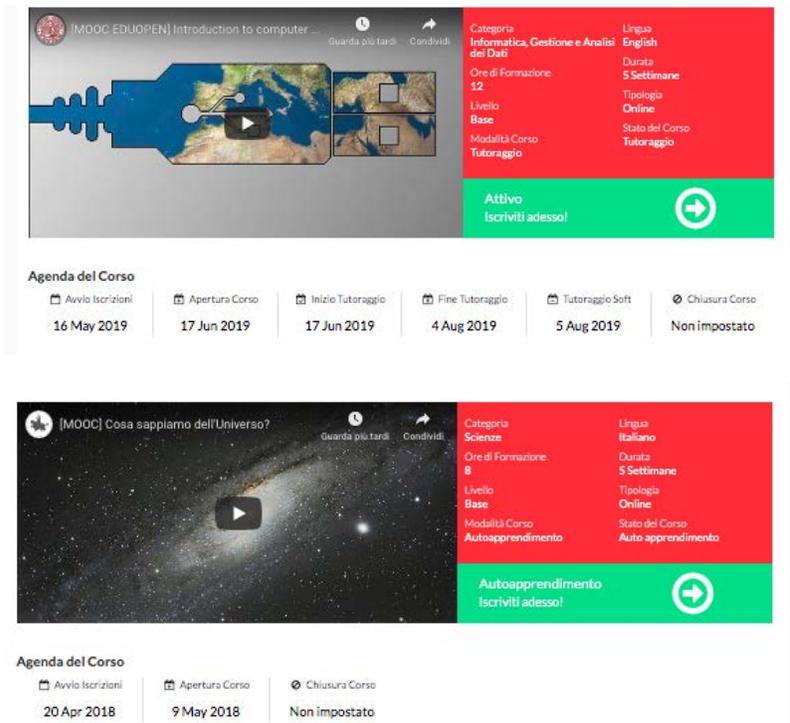


Fig. 3: Example of a course schedule in tutoring mode and in self-paced mode

These are central elements in the design process of the MOOCs, from which this research, analysing course delivery methods, was developed.

3. Materials and Methods

This empirical study, realised from data produced by the learning analytics system of the EduOpen Portal, focused on analysing the percentage of the completion/abandonment rate recorded on the whole set

of MOOCs published on EduOpen, differentiated on the basis of the delivery mode defined in the design phase. In July 2019 there were 247 courses in the catalogue with more than 55,000 users. This contribution answers the following research question:

- can the way in which the courses are delivered – self-paced and tutoring – influence the levels of completion of the courses?

The purpose of the research is to determine any differences in the completion rates of the courses with respect to the mode of delivery chosen. The results of the research can be integrated into the process of instructional design of MOOCs on EduOpen.

3.1. Data Set

Our data set is composed of 195 MOOCs published on the EduOpen Portal, selected from the entire set of published courses (247). Of the total number of courses, 66 are part of one or more pathways, while the number of active pathways is 30. Compared to the categories present on EduOpen (arts and humanities, computer and data sciences, health and pharmacology, science, social science, technology/design and engineering) the highest percentage of courses falls into the social science category (42%) (Fig. 4).

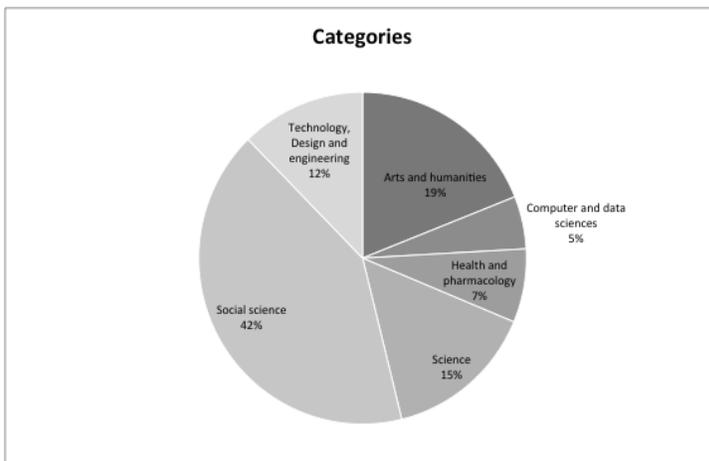
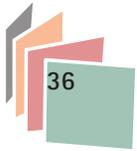


Fig. 4: Categories used on the EduOpen Portal

Data on the number of enrolments and completions were collected through the LA system and produced by the LMS of the EduOpen Portal.

The total number of MOOCs published excludes “capstone” courses, whose life cycles were incomplete with respect to the dates envisaged for their delivery (for example, courses running for fewer weeks than those envisaged for the completion of training activities, as indicated in the presentation form of the MOOCs). In fact, the choice of courses took into account the life cycles of the MOOCs linked to the mode of delivery of the same, as described above. In addition, courses closely linked to university courses whose levels of completion are by nature high have been excluded. The data collected and associated with the individual courses are related, in addition to the mode of delivery, to:

- the state of the course, which may differ from the mode of supply indicated by the date of the MOOCs’ opening (Fig.3). In this case we can find courses in self-paced, tutoring, archived, in closing, etc.;
- hours of training and duration (indicated in weeks);
- the course category, the reference university and indication of the possible pathway to which the MOOCs are linked;
- the opening date of the course, the start/end date of the tutoring (if any), the closing date of the enrolment, the closing date of the course (if stability);
- the number of users enrolled in the course and the number of users who completed it (obtaining the certificate of participation).



For self-paced courses the average training hours associated with each individual course is equal to 13.45 hours and the average duration in the week is equal to 4.5 hours; while in the tutoring courses the average duration in training hours is equal to 16 hours and the average duration in the week is, as for self-paced courses, equal to 4.5 hours.

The selected sample is made up of 93 courses in “Self-paced” mode (S), 82 courses in “Tutoring” mode (T) and 20 courses in “Undefined Type” (U). The latter includes MOOCs that for reasons related to the course agenda and the change in delivery mode (e.g. the change from Tutoring mode to Self-paced) it was not possible to associate with one of the two modes indicated (S or T).

The full data set and R Markdown file are available as an attachment.

3.2 Method

To verify the differences between the three distributions related to the three groups identified within the sample, we used the tools of descriptive statistics, identifying the mean, standard deviation and quartiles.

We focused in particular on the calculated averages and, in order to verify if the differences between them were statistically significant, we found the possible overlapping of the confidence intervals for each of the three groups; the normality of the distributions was verified through the Shapiro-Wilk test. The t-test was used to verify the reliability of our assumptions.

4. Analysis and Findings

We started from the description of the three course samples, calculating the central trend measurements for each (Tab. 1) to check if the course completion trend in the three subsamples exhibited any differences.

As shown in Table 2, the average of group S is slightly higher than that of the other two groups, which differ by a few tenths. The boxplots (Fig. 4) show a significant superposition.



Course status	N courses	Average % completion	DEV.ST
Self-paced (S)	93	25.75	13.01
Tutoring (T)	82	22.83	11.52
Undefined type (U)	20	22.22	9.13

Tab. 1: Average and standard deviation of the percentages of completion in the three subsamples

Course status	N courses	MIN	1st Qu.	Median	3st Qu.	MAX
Self-paced (S)	93	2.03	16.19	23.84	34.34	56.00
Tutoring (T)	82	2.42	14.96	22.39	29.28	50.67
Undefined type (U)	20	6.46	15.66	22.14	27.97	40.20

Tab. 2: Quartiles in the three groups

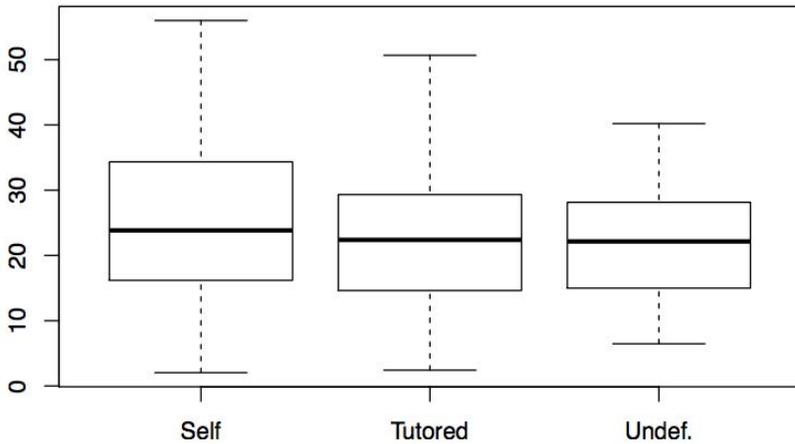


Fig. 5: Boxplot of the three samples of courses



We sought to evaluate the confidence intervals to check for differences or overlaps.

First, we needed to test the normality distribution of the samples referred to the completion rate (Rate). We used the Shapiro-Wilk normality test (H_0 : distribution is normal). In the three cases we attained a value of $p > 0.05$, and so at 95% we could not refuse the NULL hypothesis and hence considered the three distributions as normal.

We moved on to evaluate confidence intervals for the means in the three samples (if CIs were not disjointed we gained some indication that the means were not statistically different). As can be seen in Tab. 4 and Fig. 5, the three CIs were not disjointed, so there was evidence that the three means were not statistically different.

	inf	sup
Confidence intervals for the mean of the completion rate in tutored courses	20.34	25.32
Confidence intervals for the mean of the completion rate in self-paced courses	23.10	28.39
Confidence intervals for the mean of the completion rate in undefined courses	18.22	26.22

Tab. 3: Confidence interval

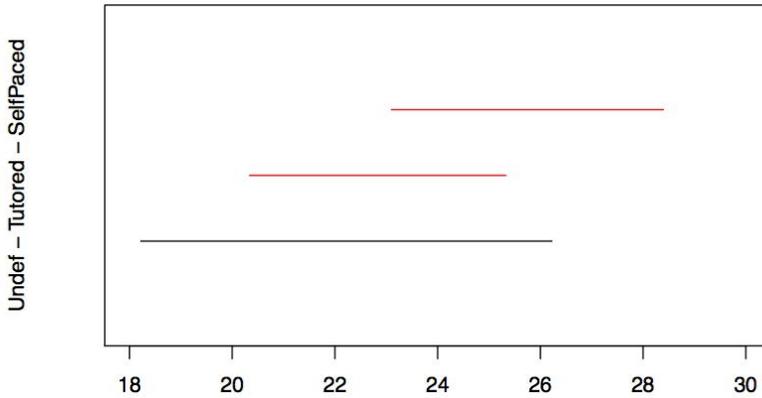


Fig. 6: Confidence intervals

Lastly, we evaluated the *t-test* for the NULL hypothesis “Difference is generated by casuality” against the H1 “Difference is generated by the tutoring”. We compared only tutored courses versus self-paced courses.

The *p-value* was $0.12 > 0.05$, so we could not refuse the NULL hypothesis. This implied the two means were not statistically different and thus we could conclude that tutored courses or self-paced courses showed statistically similar behaviour in terms of completion rate. We applied the same methodology appending undefined courses to tutored (case 1) and self-paced (case 2) courses, obtaining the same results.

The data analysis revealed that tutoring a course does not increase its completion rate.

5. Conclusions

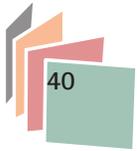
The aim of the research was to determine the differences in the completion rates of courses in MOOCs with their delivery methods. The results of the study showed that the completion rates of tutoring courses do not increase compared to self-paced courses.

The next step of the research will consider the instructional design process of EduOpen’s MOOCs focusing on rethinking the tutoring and the interaction modality in the online learning environment. It is possible to identify future areas of research related to the need to investigate the effects of the data emerging from the study, with respect to:

- the didactic design process of the MOOCs, and how to reconsider the tutoring modality in the courses. Does the presence of support



- figures in MOOCs qualitatively and quantitatively affect the levels of user interaction and participation beyond the levels of completion?
- the analysis of the “reasons” for abandonment from a student’s perspective in relation to the processes of self-regulated learning (Crosslin, 2018); in fact, several studies associate to the limited self-regulating abilities of learners one of the possible causes leading to dropout (Maldonado-Mahauad et al., 2018a, 2018b). In a digital learning environment, significant self-regulation capabilities are required, with obvious repercussions on completion rates. As Van Laer and Elen (2017) have highlighted, every learning environment, but especially digital, differs considerably in terms of the technologies and didactic methodologies adopted: just as these environments are created, they *challenge* the processes of self-regulation of learners. It is therefore necessary to move towards different analysis approaches, such as “study of learning analytics to better understand students’ performance in digital environments” (Gil-Jaurena et al., 2018, p. 53).



If the presence of support figures does not affect the completion rates, it is not possible to disregard the importance linked to the presence of the instructor (or other professional figures) with respect to the significant effects recorded in terms of “learning engagement”, the quality of the formative activity and the levels of interaction of the learners. The focus is on perceptions, motivation and learning attitudes (Saadé et al. 2017). If certain conditions – e.g. instructional design and in EduOpen this aspect is linked to the guidelines already mentioned - are met “MOOC participants can and do experience engaged, high quality learning” (Wintrup et al., 2015, p. 4). In agreement with Michele Pellerey (2006) it becomes fundamental to identify analysis and intervention tools to act on self-determination and self-regulation in learning processes, to investigate the role played by motivations in promoting personal, cultural and professional development. These processes are also enhanced by the spread of MOOCs in academia and in formal learning contexts.

As of for this complex educational scenario, the role and support offered by the learning analytics system is not limited only to the collection and monitoring of data – associated in this study with levels of completion and rates of dropout that are still two elements to which instructors look carefully to evaluate the outcome of learning processes and instructional design – but it is the resource from which (re)start to (re)think the processes of personalization of learning environments: in relation to the course and curriculum design of MOOCs (Haras et al., 2017) and to the experience of the learners.

References

- Bozkurt A., Akgun-Ozbek E., Zawacki-Richter O. (2017). Trends and patterns in massive open online courses: Review and content analysis of research on MOOCs (2008-2015). *The International Review of Research in Open and Distributed Learning*, 18(5).
- Crosslin M. (2018). Exploring self-regulated learning choices in a customisable learning pathway MOOC. *Australasian Journal of Educational Technology*, 34(1).
- De Santis A., Sannicandro K., Bellini C., Minerva T. (2019). Reasons for Attending a MOOC: A Survey on EduOpen Learners. In A. Volungeviciene, A. Szucs (eds.), *Connecting through Educational Technology*. 28th EDEN Annual Conference "Connecting through educational technology" (pp. 356-364). Bruges, 16-19 June 2019.
- Gašević D., Dawson S., Rogers T., Gasevic D. (2016). Learning analytics should not promote one size fits all: The effects of instructional conditions in predicting academic success. *The Internet and Higher Education*, 28, 68-84.
- García-Peñalvo F.J., Fidalgo-Blanco Á., Sein-Echaluce M.L. (2018). An adaptive hybrid MOOC model: Disrupting the MOOC concept in higher education. *Telematics and Informatics*, 35(4), 1018-1030.
- Gil-Jaurena I., Dominguez Figaredo D., Theeraroungchaisri A., Yamada T. (2018). 'EdX Insights' metrics from a socio-constructivist pedagogical perspective. In A. Volungeviciene, A. Szucs (Eds.), *EDEN 2018 Annual Conference: Exploring the micro, meso and macro - Navigating between dimensions in the digital learning landscape* (pp. 53-60).
- Haras C., Taylor S.C., Sorcinelli M.D., von Hoene L. (2017). *Institutional Commitment to teaching excellence: Assessing the Impacts*. Retrieved from <https://www.acenet.edu/Documents/Institutional-Commitment-to-Teaching-Excellence.pdf>
- International Council for Open and Distance Education (ICDE) (2018). *Global quality in online, open, flexible and technology enhanced education*. Retrieved from <https://www.icde.org/knowledge-hub/report-global-quality-in-online-education>
- Jordan K. (2014). Initial trends in enrolment and completion of massive open online courses. *The International Review of Research in Open and Distributed Learning*, 15(1).
- Jordan K. (2015). Massive open online course completion rates revisited: Assessment, length and attrition. *The International Review of Research in Open and Distributed Learning*, 16(3).
- Jung Y., Lee J. (2018). Learning engagement and persistence in massive open online courses (MOOCs). *Computers & Education*, 122, 9-22.
- Khalil H., Ebner M. (2014). MOOCs completion rates and possible methods to improve retention. A literature review. In *EdMedia+ Innovate Learning* (pp. 1305-1313). Association for the Advancement of Computing in Education (AACE).
- Kim S.W. (2016) MOOCs in higher education. In D. Cvetkovic (ed.), *Virtual Learning* (pp. 121-135). InTech.



- Kizilcec R.F., Schneider E. (2015). Motivation as a lens to understand online learners: Toward data-driven design with the OLEI scale. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 22(2), 6.
- Loizzo J., Ertmer P.A., Watson W.R., Watson S.L. (2017). Adult MOOC learners as self-directed: Perceptions of motivation, success, and completion. *Online Learning*, 21(2).
- Maldonado-Mahauad J., Pérez-Sanagustín M., Kizilcec R.F., Morales N., Muñoz-Gama J. (2018a). Mining theory-based patterns from big data: Identifying self-regulated learning strategies in Massive Open Online Courses. *Computers in Human Behavior*, 80, 179-196.
- Maldonado-Mahauad J., Pérez-Sanagustín M., Moreno-Marcos P.M., Alario-Hoyos C., Muñoz-Merino P. J., Delgado-Kloos C. (2018b). Predicting learners' success in a self-paced MOOC through sequence patterns of self-regulated learning. *European Conference on Technology Enhanced Learning* (pp. 355-369). Springer, Cham.
- Margaryan A., Bianco M., Littlejohn A. (2015). Instructional quality of massive open online courses (MOOCs). *Computers & Education*, 80, 77-83.
- Onah D.F., Sinclair J., Boyatt R. (2014). Dropout rates of massive open online courses: Behavioural patterns. *EDULEARN14 proceedings*, 1, 5825-5834.
- Pellerey M. (2006). *Dirigere il proprio apprendimento: autodeterminazione e autoregolazione nei processi di apprendimento*. Brescia: La Scuola.
- Saadé R.G., He X., Kira D. (2007). Exploring dimensions to online learning. *Computers in human behavior*, 23(4), 1721-1739.
- Tait A. (2003). Reflections on student support in open and distance learning [Editorial]. *International Review of Research in Open and Distance Learning*, 4(1).
- Van Laer S., Elen J. (2017). In search of attributes that support self-regulation in blended learning environments. *Education and Information Technologies*, 22(4), 1395-1454.
- Wintrup J., Wakefield K., Davis H.C. (2015). Engaged learning in MOOCs: a study using the UK Engagement Survey. Retrieved from https://eprints.soton.ac.uk/373640/1/HEA_engaged-learning-in-MOOCs.pdf
- Yousef A.M.F., Chatti M. A., Schroeder U., Wosnitza M. (2014). What drives a successful MOOC? An empirical examination of criteria to assure design quality of MOOCs. *2014 IEEE 14th International Conference on Advanced Learning Technologies* (pp. 44-48).
- Yu H., Miao C., Leung C., White T.J. (2017). Towards AI-powered personalization in MOOC learning. *Science of Learning*, 2(1), 15.

Design of a Learning Analytics framework proposal in academic context

Proposta di un framework per i Learning Analytics nel contesto accademico

Leonardo Caporarello

Bocconi University, and SDA Bocconi School of Management, leonardo.caporarello@unibocconi.it

Federica Cirulli

Bocconi University, federica.cirulli@unibocconi.it

Manuela Milani

Bocconi University, manuela.milani@unibocconi.it

In the last decades, learning analytics (LA) have been considered as one of the key emerging trend in higher education (Vassakis et al., 2018) and they have attracted a lot of attention among researchers and designers for their potential to address some of the major challenges in the academic sector (Bach, 2010). Despite the broad interest and implementation of LA processes, there remain numerous issues regarding the availability of developed and adaptable-to-the-context reference framework across Universities (Gašević et al., 2015). Furthermore, scholars and practitioners have approached LA from a range of perspectives: it is necessary to define not only the aims of what could be achieved using LA but also what should be done to attain it. Generally, LA are considered as measurement, collection, analysis and reporting of learners' data and their contexts, to optimize learning and the environments in which it occurs (Khalil and Ebner, 2016). Starting from this definition, and using an action research method, we propose an "extended" LA framework that puts learner and instructor at the centre of it.

Keywords: Learning Analytics; framework; personalized learning

Negli ultimi decenni, i learning analytics (LA) sono stati interpretati come uno dei trend di maggiore importanza nell'ambito dell'istruzione universitaria (Vassakis et al., 2018) e hanno attratto l'attenzione di ricercatori e instructional designer per il loro potenziale nell'affrontare alcune delle principali sfide all'interno delle istituzioni accademiche (Bach, 2010). Nonostante l'interesse generale e l'implementazione di processi di LA, nelle università permangono numerosi problemi relativi alla disponibilità e allo sviluppo di modelli di riferimento adattabili ai diversi contesti (Gašević et al., 2015).

Inoltre, ricercatori e instructional designer hanno indagato i LA da una serie di prospettive: è necessario definire non solo gli obiettivi che si possono raggiungere usando i LA, ma anche ciò che dovrebbe essere fatto per raggiungere tali obiettivi. Generalmente, i LA sono considerati come misura, raccolta, analisi e reporting dei dati degli studenti e dei loro contesti, per ottimizzare l'apprendimento e gli ambienti in cui esso si verifica (Khalil and Ebner, 2016). Partendo da questa definizione e adottando il metodo della ricerca-azione, il paper propone un framework di LA "esteso" che metta al centro lo studente e il docente.

Parole Chiave: Learning Analytics; framework; Apprendimento personalizzato

This paper is jointly written by all authors. However, Leonardo Caporarello wrote paragraphs 3, 5.3 and 6; Federica Cirulli wrote paragraphs 1 and 2; Manuela Milani wrote paragraphs 4, 5.1 and 5.2.

1. Introduction

A learning analytics (LA) intervention is considered as the surrounding group of actions through which analytic tools, data, and reports are taken up and used. It refers to a soft technology that involves human processes (Arthur, 2009; Hlupic et al., 2002). To date, most research and development in LA has focused on the principal issues of data collection, management, processing and display. However, as we enter a stage in which LA are rapidly being rolled out for more general use, the design of a framework for LA becomes a critical element in supporting the effective implementation of these tools in the academic domain (Wise, 2014).

Nevertheless, short attention has been paid to adopt analytic models from operations to improve students and instructors' "direct" involvement even though their contribution seems to have had a large impact on their success (Bach, 2010).

To come into line with this research's flow, we conducted a systemic literature review of LA. Then, we propose an extended framework for a successful LA design and management that put the learner and instructor at the centre of the process.

In the following sections, we depict first a rationalization of LA, then we describe different components concerning LA. Secondly, we present our research method in which the process to define our framework design proposal is illustrated. Results, discussions and conclusions are drawn in the last part of the paper. Implications for practitioners are also considered.

2. Learning Analytics: definitions, data, models and obstacles

A LA intervention is described as the surrounding frame of activities through which analytic tools and data are taken up and adopted; it is a soft technology that involves the orchestration of different procedures. For this reason, existing research on LA focuses on different elements.

Firstly, some scholars explored LA definitions and their possible stakeholders (Campbell and Oblinger, 2007). For example, LA are intended as the measurement and reporting of data about learners, and have been advocated as a support tool for instructor regulation of collaborative learning or they have been considered as statistical techniques and predictive modelling to help faculty and advisors in



determining which students may face academic difficulty, allowing interventions to help them in succeeding (Campbell and Oblinger, 2007). In these definitions, learners are usually considered “passive actors” or, mainly, “sources of data” even if there is a growing learners’ demand for active participation to enhance their learning (Gašević et al., 2015).

Secondly, research has pondered elements that must be taken into consideration when conducting LA studies. In such regard, Greller and Drachsler (2012, p. 43) define six elements that must be considered “to ensure appropriate exploitation of LA in an educationally beneficial way”: stakeholders, objectives, data, instruments, external limitations (i.e. ethical, legal, managerial/organizational) and internal limitations (i.e. the lack of experts in analytics projects).

Thirdly, more recent analytics models about LA have been implemented by numerous scholars (Clow, 2013; Greller and Drachsler, 2012; de Freitas et al., 2015; Buckingham et al., 2012). The Learning Analytics Model (LAM) (de Freitas et al., 2015) is one of the most frequently used and defines how information should be tracked, aggregated and reported in a Learning Analytics System (LAS). Furthermore, Greller and Drachsler (2012) proposed a generic model intended as a guide in setting up services within an educational institution. In particular, they put in evidence the challenges of the soft dimensions of LA like ethics and the need for instructors to develop competences in interacting with data. Others are mainly oriented on tracking behaviours, persistence, achievement (Macfadyen and Dawson, 2010), participatory and peer learning (Clow and Makriyannis, 2011) and social LA (Buckingham et al., 2012).

Nevertheless, these studies continue to derive from traditional models that basically consider learners as “passive actors” and instructor the one who collects data on learners’ activities (Siemens, 2013). Moreover, to date, the conceptualisations of these frameworks aspire to be completely realised, and LA implementations across higher education organizations are typically immature with limited ability to demonstrate manifest impact (Colvin et al, 2015).

Finally, obstacles in LA are often related to data sources and management: big sets of data and data collected from diverse sources, with distinctive standards and from users with different levels of access, reveal an important challenge presented by incorporating data analyses into strategic planning. Data are abundant and usually easy to extract, but they need to be turned into useful information (Van Barneveld et al., 2012), that is the crucial issue when designing an LA institutional strategy.



2.1 *Learning Analytics in the academic context*

There are different opinions about the main meanings of academic analytics. Some scholars (U. bin Mat et al., 2013; Aljohani and Davis, 2012; Goldestain and Katz, 2005) consider academic analytics as a new way of applying business intelligence in academic domains to provide data with the emphasis being on institutional, regional, and international levels; others (Huda et al., 2017; Maseleno et al., 2018) thought academic analytics as mainly focused on the improvement of organizational processes, workflows, resource allocation through the “use” of learner, academic, and institutional data. Likewise, academic analytics clarify the role of LA at the institutional, administrative and policy making levels (Aljohani, 2012).

Thanks to a well-defined process of LA, Universities could use specific set of data to develop decision making and resource allocation, recognize at-risk learners and areas of concern, they can get a better insight into their strengths and weaknesses, they can drill down on causes of complex challenges, and they can generate and try different academic strategies (Marks et al., 2016).

To reach these “potential aims”, the actors involved (i.e. instructors, learners, designers) need to become further familiar with issues related to the use of LA, so most part of institutions continue to encourage training and innovations in this field (Avella et al., 2016).

3. Learning analytics: insights on the literature gap

As it is evident, no complete agreement exists on definitions, elements to be considered, obstacles to face and models to be adopted. The lack of a reference architecture causes troubles, both for institutional collaboration, setting an agenda and for the teaching and learning improvement.

The charge to Universities is to determine what data can support the improvement of learning and teaching, what actors have to be involved and at what level (Mattingly, et al, 2012). Furthermore, there are few considerations about learners and instructor as “active” participants in the whole LA processes (Gašević et al., 2015).

Additionally, research shows that typically instructor is not viewed as a user and as an analyst at the same time, specifically as an analyst who goes deep into collected data to approach educational problems (Siemens, 2013). On the contrary, if instructors come through the mapping of data (i.e. methods and number of students, type of stu-



dents or level etc.), they would dynamically enter the framework being able to reflect on data that concern them directly or indirectly.

For these reasons, we strongly support the hypothesis of the creation of a structure of LA in which instructor and student's role are "integrated" in it. Learners will be viewed not only as the "source" of data, but they will be part of an interactive approach to explore their views and ideas about LA (Piotrkowicz, 2017). Thanks to LA outputs, learners can be encouraged to take personal responsibility for their personal situations - making use of the feedback available about what they are doing, and making proper decisions about support. Additionally, the instructor need to be more engaged and supported in the process and his/her role need to be more specified to give him/her authentic opportunities for reflection and reaction (Van Leeuwen, 2014).

4. Methodology

The research method adopted for the scope of this paper is the action research, which aims at "producing practical knowledge that is useful to people in the everyday conduct of their lives" (Eden, and Huxham, p. 238), through a continuous cycle of developing and elaborating theory from practice. Subjects and researchers are jointly responsible for developing and evaluating theory to guarantee that the results of the research help to solve a real challenge of the subjects and reflect the knowledge created through the participative process (Caporarello et al, 2020). Action research has been chosen because it aims to contribute both to the practical challenges of people in an immediate problematic situation and to further the goal simultaneously (Gilmore et al., 1996).

In our specific case, the challenge was related to the need of providing a tool based on data and information in LA processes enabling learner and instructor to have an active role.

Broadly speaking, our research entails phases of groundwork, intervention and theory testing and development. At this stage of the research, we are in the intervention phase.

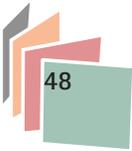
In the groundwork phase, we identified the research gap analysing previous studies on LA: we search for best practices, conducting obviously first of all a literature review. Specifically, a structured documents retrieval process has been realized by launching on Scopus database the search terms "Learning Analytics", which have been cross-referenced (AND search) with "University", "learners", "definitions", "obstacles" and "models". The first search showed more 26.400 results after 2018. As our domain of analysis of LA is the higher education, then we ap-



plied the criteria more pertinent for academic domain (i.e., academic analytics, analytics and universities' course and programmes) and a short list of about 100 relevant papers emerged.

In the intervention phase, we organized focus groups with small groups of students held at Bocconi University. The aim of these meetings was to design a LA framework viewed as the best fitting tool able to reach the challenge. A choice of heterogeneity has been made for the creation of the groups. During the focus groups, an instructional designer acted as moderator and an academic developer transcribed the answers. Then, a team of academic developers put together the focus group outcomes, gathered the data and developed the framework.

In the theory testing and development phases, we will include the implementation of our framework within a real University context. Being the research method cyclic, theory testing and development phases will imply revision and modification of previous phases' results.



5. Framework proposal

As described in the previous paragraph, the action research method has been adopted to design the framework here presented (Fig. 1).

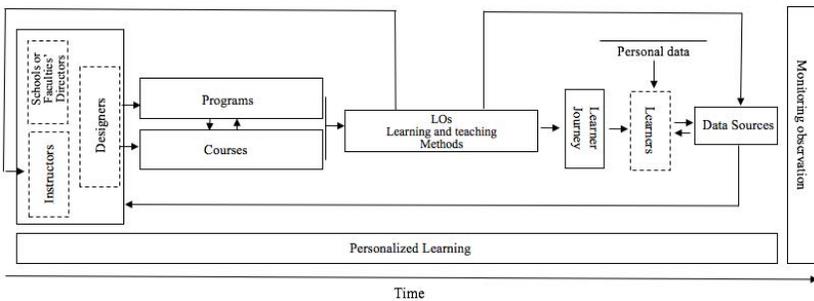


Fig. 1: LA Framework proposal

The list of elements in the framework is not intended to be exhaustive and can be extended on a case-by-case basis. More in detail, this framework identifies and maps each potential element of the LA in the academic context. The framework gives a picture of these elements considering a timeline that highlights the iterative nature of the entire process.

5.1 *Actors in the framework: the role of the instructor*

This framework is organized as follows. In the left part of figure 1, the actors potentially involved in the LA processes are illustrated: instructors, Schools or Faculties' directors and designers.

More in detail, the label "designers" refers to all different actors that can have responsibilities in all phases of courses' design. The educational actors in the framework are indicated in the dashed boundary boxes.

Instructors and Schools or Faculties' directors are in relation with the other components of the framework as described later on. Specifically, Schools or Faculties' directors and instructors interact to guarantee the coherence between programs and courses (i.e. in terms of outcomes, teaching and assessment methods) and to ensure the best possible learning experience.

Thanks to analytics, instructors can reflect on data that concern them directly (i.e. teaching methods) or not directly (i.e. students' performance). Regarding the data concerning the instructors directly, they can act as "designer" and then re-evaluate the goodness of their course design in a longitudinal perspective. Regarding the data concerning the instructors not directly, they can look at the students who perform worse and, therefore, intervene in good time.



5.2 *The learning experience: the role of the learner*

Then, moving from the left to the right, there is the part related to the teaching and learning experience that is articulated in programs and courses.

Of course, both programs and courses are designed starting from learning outcomes (LOs) identification and teaching and learning methods definition, according to the constructive alignment principle that emphasize the importance of the direct alignment among teaching activities, assessment and learning outcomes (Biggs, 2003). In fact, the next element of the framework is represented by the set – when available - of LOs of the programs as well as those of the courses. Between these two groups of LOs a relation exists: LOs of the programs refers – and should be coherent - with the LOs of the courses. In fact, ideally single courses' LOs should be integrated with the LOs attaining the whole program. Moreover, the LOs are - or should be - the basis of design and teaching methods choices. So, LA can be an also an excellent opportunity to verify the achievements of the outcomes themselves at

different layers: the single course learning outcomes as well as the program's learning outcomes. Of course, this can be possible only thanks to a precise articulation of the LOs. Nevertheless, over time LA can potentially support the analysis of the outcomes and their reformulation according to information on a larger scale optimizing the learning journey that will be defined later on. So, LA permit to highlight some discrepancies between expected and collected data. At this stage, the competence of designers is needed in order to identify at which level the redesign should be realized in terms of outcomes, methods or assessment.

In such perspective, analytics provide a great opportunity for design reflection: outcomes and methods can be more easily evaluated and discussed by whoever acts as a designer of the single course as well as of the degree courses.

After the space devoted to LOs and learning and teaching methods, it is possible to note the learner journey box.

Learner journey provides data (such as class or work groups participation or other data according to the used LMS tracking capabilities) and could be connected with other personal data of the learners and, eventually, with individual data. Learners, being at the same time providers and users of the data, could - and should - be part of the framework at different layers. Depending on whether the learner has access to data relating to her/his peer group of the same course or attend the same programs, she/he has the opportunity to check how she/he is placed with respect to the target group. This is in line with what affirmed by Kruse and Pongsajapan (2012) who proposed a "student-centric" approach, as opposed to an "intervention-centric" approach to LA. This suggests that student should be considered "as co-interpreter of their own data and perhaps even as participants in the identification and gathering of those data" (pp. 4-5). In addition, if the student is given the opportunity to identify the types of peers with whom to deal, the opportunity to consciously plan learning path increases.

Sources are the last element of our framework. Examples of available sources can be the attendance data, exams grade, data from LMS that are often managed by different professionals belonging to diverse units. Connecting these available datasets can facilitate the development of mash-up applications that can lead to more learner-oriented services and therefore improved personalisation (Greller and Drachsler, 2012).

Concluding, this framework allows us to put in evidence that the same information available to the learner is also the one available to the designer (i.e. instructor and/or Schools or Faculties' directors), who



obviously has at her/his disposal the aggregated data, as well as the data of the single learner and can intervene directly on him/her to prevent any drop out or difficulty (Greller and Drachsler, 2012).

Obviously, the learner lives a learning experience that derives from the choices made by the designer.

5.3 *Monitoring observation*

Shifting from the single course edition to diverse editions means to add a timeline dimension in the framework. LAs process results regarding different editions of the same program/course can be compared. Schools or Faculties' directors evaluate LA data to understand if and how it is possible to improve the programs' organization. These kinds of analyses can be conducted on two layers: the first one concerns the whole program's performances; the second one regards single courses.

Even the instructor analysis is done on two layers. The first one regarding the general performance of the course; the second one conducted at single lecture's layer to understand the single lecture method effectiveness as well as what are the more difficult topics to be comprehended by learners.

These two layers are part of an iterative process that develop itself along a specific timeline as showed in Fig.1. If the process is well organized and the data well structured, different goals could be reached:

- from the point of view of the learners, a real commitment and a personalized learning experience;
- from the point of view of all the designers involved the great opportunity to evaluate and eventually re-align the course or program's design.

What represents an added value of a well-structured LA frame is the fact that instructor can be pushed to think deeply about different learners' need or preferences, so that he/she can better approach "new" learners on the bases of all meaningful data available.

6. Discussions and conclusions

LA involve relatively long time processes, using data from various institutional layers (for example, courses, programs, etc.) to inform decisions about future (Wise, 2014).



In this paper we propose an extended and efficacious LA design and management framework in which the learner and instructor are at the centre of the process. Our methodology can be considered appropriate because the action research involved different LA profiles. These profiles included not only technical professionals but also educational experts as well as students.

Based on our review of the theory background, what emerged is the necessity to design an architecture where:

- all actors are involved at different and well-defined layers;
- data are integrated from multiple sources to improve the accuracy of a learner profile and the subsequent personalization;
- instructor is both viewer and analyser of data resulting from LA;
- learner is both provider and viewer of data;
- sources of data are clearly indicated.



The proposed framework represents a first contribution - to be further developed in successive research - to rethink the concept of LA itself and the interactions between different LA layers with the intention of better defining actors involved, and then, learner's profile and her/his "personalized" learning path. Moreover, thanks to the proposed framework, instructors can periodically improve their teaching from the point of view of both quality and efficiency. In this sense, our work offers a significant contribution in the LA research's area.

This reflection sets the ground for novel investigations on how LA sustain personalized learning experiences through customized recommendations (Siemens, 2010).

Additionally, in the near future, context - and so the framework itself - could be enlarged with the inclusion of data from other sources like mobile devices, physical data from supervision meetings and game environments in addition to the usage of university resources such as libraries as well as learners' preferences, might result in a more complete learner and instructor profiles (Baalsrud-Hauge et al., 2014).

References

- Aljohani N, & Davis H. (2012). *Learning analytics in mobile and ubiquitous learning environments*. In Proceedings of the 11th World Conference on Mobile and Contextual Learning.
- Arthur W. B. (2009). *The nature of technology: What it is and how it evolves*. New York, NY: Free Press.

- Avella J. T., Kebritchi M., Nunn S. G., & Kanai T. (2016). Learning analytics methods, benefits, and challenges in higher education: A systematic literature review. *Online Learning*, 20(2), pp. 133-149.
- Baalsrud Hauge J., Fernández-Manjón B., Bellotti F., Fiucci G., Padrón-Nápoles C., Nadolski R., & Westera W. (2014). *Implications of learning analytics for serious game design*. In Proceedings of the IEEE 14th International Conference on Advanced Learning Technologies Los Alamitos, California (pp. 230-232).
- Bach C. (2010). *Learning analytics: Targeting instruction, curricula and student support*. Office of the Provost: Drexel University.
- Balancing Between Overview and Overload. *Journal of learning Analytics*, 2(2), pp. 138-162.
- Biggs J.B. (2003). *Teaching for quality learning at university* (2nd ed.). Maidenhead: Open University Press.
- Buckingham S., & Ferguson R. (2012). Social learning analytics. *Educational Technology and Science*, 15(3), pp. 3-26.
- Campbell J.P., & Oblinger D.G. (2007). *Academic analytics*. Retrieved July 15th, 2019 EDUCAUSE. Retrieved July 12th, 2019 from <http://net.educause.edu/ir/library/pdf/PUB6101.pdf>
- Caporarello L., Cirulli F., & Manzoni B. (2020). Designing a Self-regulated Online Learning Course Using Innovative Methods: A Case Study. In R. Gennari et al. (eds), *Methodologies and Intelligent Systems for Technology Enhanced Learning, 9th International Conference. MIS4TEL 2019*. Advances in Intelligent Systems and Computing, 1007. Springer, Cham. https://doi.org/10.1007/978-3-030-23990-9_15
- Clow D. (2013). An overview of learning analytics. *Teaching in Higher Education*, 18(6), pp. 683-695
- Clow D., & Makriyannis E. (2011). *iSpot analysed: Participatory learning and reputation*. In Proceedings of the 1st International Conference on Learning Analytics and Knowledge. Retrieved July 10th, 2019 from <http://dl.acm.org/citation.cfm?id=2090121&CFID=82269174&CFTOKEN=35344405>
- Colvin C., Wade A., Dawson S., Gasevic D., Buckingham Shum S., Nelson K., & Fisher J. (2015). *Student retention and learning analytics: A snapshot of Australian practices and a framework for advancement*. Draft final report.
- de Freitas S., Gibson D., Du Plessis C., Halloran P., Ambrose M., Dunwell I. & Arnab S. (2015). Foundations of dynamic learning analytics: Using university student data to increase retention. *British Journal of Educational Technology*, 46(6), pp. 1175-1188.
- Gašević D., Dawson S., & Siemens G. (2015). Let's not forget: Learning analytics are about learning. *TechTrends*, 59(1), pp. 64-71.
- Goldstein P.J., Katz R.N. (2005). *Academic Analytics: The Use of Management Information and Technology in Higher Education - Key Findings*, Boulder. Center for Applied Research: CO: EDUCAUSE.
- Greller W., & Drachsler H. (2012). Translating learning into numbers: A generic framework for learning analytics. *Educational Technology and Society*, 19(3), pp. 42-57.
- Eden C., & Huxham C. (1996). Action research for the study of organisation. In S. Clegg, C. Hardy & W. Nord (eds.), *The handbook of organisation studies* (pp. 526-542). Sage Publications: London.



- Gilmore T., Krantz J., & Ramirez R. (1986). Action Based Modes of Inquiry and the Host Researcher Relationship. *Consultation*, 5(3), pp. 160-176.
- Hlupic V., Pouloudi A., & Rzevski G. (2002). Towards an integrated approach to knowledge management: 'Hard', 'soft' and 'abstract' issues. *Knowledge and Process Management*, 9(2), pp. 90-102.
- Huda M., Jasmi K.A., Embong W.H., Safar J., Mohamad A.M., Mohamed A.K., Muhamad N.H., Alas Y., & Rahman S.K. (2017). Nurturing Compassion-Based Empathy: Innovative Approach in Higher Education. In M. Badea, & M. Suditu (Eds.), *Violence Prevention and Safety Promotion in Higher Education Settings* (pp. 154-173). Hershey, PA: IGI Global.
- Khalil M., & Ebner M. (2016). *Learning Analytics in MOOCs: Can Data Improve Students Retention and Learning?* In Proceedings of EdMedia 2016--World Conference on Educational Media and Technology (pp. 581-588). Vancouver, BC, Canada: Association for the Advancement of Computing in Education (AACE). Retrieved June 20th, 2019 from <https://www.learntechlib.org/primary/p/173003/>.
- Kruse A., & Pongsajapan R. (2012). *Student-centered learning analytics*. Retrieved July 11, 2019 from <https://cndls.georgetown.edu/m/documents/thoughtpaper-krusepongsajapan.pdf>
- Macfadyen L., & Dawson S. (2010). Mining LMS data to develop an "early warning system" for educators: A proof of concept. *Computers and Education*, 54(2), pp. 588-599.
- Marks A., AL-Ali M. & Rietsema K. (2016). Learning Management Systems: A Shift Toward Learning and Academic Analytics. *International Journal of Emerging Technologies in Learning (ijET)*, 11, pp. 77-82.
- Maseleno A., Sabani N., Huda M., Ahmad R., Jasmi K.A., & Basiron B. (2018). Demystifying learning analytics in personalised learning, *International Journal of Engineering & Technology (UAE)*, 7(3), pp. 1124-1129.
- Mattingley K.D., Rice M.C., & Berge Z.L. (2012). Learning analytics as a tool for closing the assessment loop in higher education. *Knowledge Management & E-Learning: An International Journal*, 4(3), pp. 236-247.
- Piotrkowicz A., Dimitrova V., Treasure-Jones T., Smithies A., Harkin P., Kirby J., & Roberts T. (2017). *Quantified self analytics tools for self-regulated learning with mypal*. In Proceedings of the 7th Workshop on Awareness and Reflection in Technology Enhanced Learning co-located with the 12th European Conference on Technology Enhanced Learning (EC-TEL 2017). Vol. 1997, CEUR Workshop Proceedings. Retrieved June, 23th from https://eprints.soton.ac.uk/342971/1/mlearn2012_Full_Paper-_Aljohani_and_Davis.pdf
- Siemens G. (2010). *What are learning analytics*. *ELEARNSPACE: Learning, networks, knowledge, technology, community*. Retrieved June 12th, from <http://www.elearnspace.org/blog/2010/08/25/what-are-learning-analytics/>
- Siemens G. (2013). Learning Analytics: The Emergence of a Discipline. *American Behavioural Scientist*, (August). Retrieved June 21th, 2019 from 10.1177/0002764213498851
- U. bin Mat, N. Buniyamin, P.M. Arsad, R. Kassim (2013). An overview of using academic analytics to predict and improve students' achievement: A proposed proactive intelligent intervention- In Engineering Education (ICEED), 2013 IEEE 5th Conference on, IEEE, pp. 126-130.

Van Barneveld A., Arnold K.E., & Campbell J.P. (2012). Analytics in Higher Education: Establishing a Common Language. *EDUCAUSE ELI. Paper 1*, pp. 1-11.

Van Leeuwen A., Janssen J., Erkens G., & Brekelmans M. (2014). Supporting teachers in guiding collaborating students: Effects of learning analytics in CSCL. *Computers & Education*, 79, 28-39. Retrieved, June 21th, 2019 from <http://dx.doi.org/10.1016/j.compedu.2014.07.007>

Vassakis K., Petrakis E., & Kopanakis I. (2018). Big data analytics: applications, prospects and challenges, *Mobile Big Data* (pp. 3-20). Berlin: Springer.

Wise A.F. (2014). *Designing pedagogical interventions to support student use of learning analytics*. In Proceedings of the 4th International Conference on Learning Analytics And Knowledge (pp. 203-211). ACM.



L'e-tutoring nei percorsi di formazione online

E-tutoring in online training courses

Rosa Vegliante

Department of Human, Philosophical and Education Sciences, University of Salerno, rvegliante@unisa.it

Marta De Angelis

Department of Human, Philosophical and Education Sciences, University of Salerno, mdeangelis@unisa.it

Starting from a diachronic analysis regarding the passage from distance education to e-learning, the paper aims to analyze the theme of e-tutoring both from a theoretical and a practical point of view. In drawing up a summary of the e-tutoring strategies and possible educational consequences, the advantages derived from the functions performed by the e-tutor in the learning process emerge, which derive from different styles of e-tutoring. To different pedagogical reference paradigms (behaviorist, constructivist or socio-constructivist), in fact, correspond dissimilar ways of designing e-learning environments, based on instructor centered or learning team centered teaching models. From the analysis carried out, therefore, the crucial role of the e-facilitator emerges, as it greatly affects the training success of the participants, reducing the spatial distance of the online environments and limiting possible e-dropouts.

Keywords: e-tutoring; open distance learning; feedback; learning, Learning Analytics.

Partendo da un'analisi diacronica che ripercorre il passaggio dalla formazione a distanza all'e-learning, il contributo affronta il tema dell'e-tutoring approfondendone alcuni degli aspetti fondamentali, sia dal punto di vista teorico che pratico. Nel tracciare un quadro di sintesi inerente le strategie di e-tutoring e le eventuali ricadute formative, ciò che emerge sono i vantaggi tratti dalle funzioni svolte dall'e-tutor nel processo di apprendimento, i quali derivano dai differenti stili di e-tutoring. A paradigmi pedagogici di riferimento differenti (comportamentista, costruttivista o socio-costruttivista), infatti, corrispondono diverse modalità di progettazione degli ambienti di e-learning, basate su modelli didattici instructor centered o learning team centered. Dall'analisi effettuata, dunque, viene rimarcato il cruciale ruolo dell'e-facilitator, in quanto questa figura incide notevolmente sul successo formativo dei partecipanti, riducendo la distanza spaziale degli ambienti online e arginando possibili e-dropout.

Parole chiave: tutorato online; open distance learning; feedback; apprendimento, Learning Analytics.

Il contributo rappresenta il risultato di un lavoro congiunto delle autrici, tuttavia Marta De Angelis ha scritto i paragrafi 1 e 3 e Rosa Vegliante il paragrafo 2.



1. Introduzione

L'avvento della società dell'informazione e, con essa, dell'economia digitale segna l'ingresso dell'*e-learning* nell'ambito educativo. Per comprendere tale fenomeno, è necessario evidenziare il passaggio dalla formazione a distanza (FaD) all'*e-learning*, contraddistinto dall'affermazione delle tecnologie della comunicazione. Secondo Keegan (1994) un sistema di formazione a distanza si definisce tale per una serie di elementi: la mancata continuità spaziale e temporale tra il docente e il discente; la presenza di un'organizzazione didattica ben strutturata nella quale le interazioni personali e la comunicazione bidirezionale, tra docente/tutor e discente, sono mediate dalla tecnologia; possibili incontri occasionali e un approccio metodologico analogo a quello industriale. Nipper (1989) individua tre generazioni nei sistemi FaD: la prima, basata sulla corrispondenza di materiali prevalentemente cartacei; la seconda, dominata dall'uso integrato di materiali multimediali e media specifici e la terza caratterizzata dalla formazione in rete (*online education* o *e-learning*). Si è passati dalle tecniche a stampa e la posta tradizionale ai mezzi di comunicazione di massa, sino all'affermazione delle reti telematiche e del Web. Le prime due generazioni (formazione a distanza tradizionale) rinviano a un processo formativo di tipo estensivo, volto alla distribuzione dei materiali didattici su larga scala, in cui l'interazione docente-discente è bidirezionale ed è assente la comunicazione tra pari. La terza generazione si fonda, invece, su un processo di tipo intensivo, più circoscritto, nel quale viene valorizzata la dimensione sociale che si manifesta nelle interazioni tra i partecipanti in un ambiente di apprendimento collaborativo (Trentin, 2001). La FaD di terza generazione è svincolata da barriere spazio/temporali, combina in sé le differenti modalità di interazione (docente-studenti, studente-docente, studente-studente) dilatando l'apprendimento oltre l'aula.

L'incisività e i vantaggi dell'*e-learning* vengono richiamati nel Rapporto Unesco del 2016 nel quale si sottolinea il valore aggiunto, fornito dalle Tecnologie dell'Informazione e della Comunicazione (TIC) e dall'*Open Distance Learning* (ODL), all'educazione permanente. Con l'affermarsi del Web 2.0, l'utente non è più solo consumatore (*consumer*) di informazioni ma è anche produttore (*prosumer*) di contenuti (Toffler, 1980): si inserisce nel sistema mediale, personalizzando i percorsi e adeguandoli ai tempi e alle modalità a lui più congeniali. Il soggetto/destinatario dei processi di apprendimento ha la possibilità di agire mediante l'utilizzo di piattaforme, *Learning Management System* (LMS) che, oltre a erogare contenuti, *Learning Content* (LC), permettono di interfacciarsi in modalità orizzontale, *peer to peer*, in comunità



virtuali, paragonabili a comunità di pratica di derivazione *wengeriana* (Wenger, 1998), nelle quali l'apprendimento è sia individuale che collaborativo.

All'interno di un percorso formativo adeguatamente progettato, in grado di coniugare lo studio individuale con quello di gruppo (Kaye, 1994), si colloca il *tutoring online*, ovvero l'insieme delle attività e delle strategie adottate nella gestione degli ambienti *online*, che consentono di passare dalla relazione "uno a uno" all'interazione "molti a molti", riducendone la distanza spazio-temporale.

Muovendo da tali presupposti, il contributo mira a delineare gli apporti forniti da studi e ricerche, nazionali e internazionali, in riferimento all'eventuale relazione esistente tra le caratteristiche relazionali, organizzative e metodologico-didattiche, tipiche delle attività di *tutorship online*, e le ricadute sull'apprendimento.



2. e-Tutoring e Learning Analytics

In generale, con il termine *tutoring o tutorship* (dal latino, *tutus* = sicuro) si suole indicare una relazione duale tra un soggetto in formazione e un esperto, con il compito di facilitare la conversione delle informazioni e delle conoscenze in *know-how* (Quaglino, 2005).

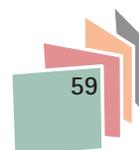
Alla base dell'azione di *tutoring* vi sono teorie pedagogiche che affondano le proprie radici nel costruttivismo sociale e che si riferiscono, in particolare, al concetto di *scaffolding* (Wood, Bruner & Ross, 1976) e a quello di *zona di sviluppo prossimale* (Vygotskij, 1978) nella sua duplice accezione affettivo-motivazionale e cognitiva, nonché alle potenzialità derivanti dalle azioni di sostegno nel processo di acquisizione di specifiche conoscenze e abilità (Rogers, 1969). Il tutor, in qualità di mentore, dispone le proprie conoscenze e competenze per favorire il processo di apprendimento, sostituendo, almeno in parte, "la funzione istruttiva centrata sull'insegnante [...] con una serie di interazioni più complesse, che portano in primo piano quelle allievo-allievo" (Calvani & Rotta, 1999, p. 135). Di contro, il tutor *online* aggiunge alle competenze disciplinari, didattico-comunicative quelle tecnologico-gestionali, specializzandosi in quattro aree specifiche: *scaffolding*, clima, comunicazione e organizzazione (Berge, 1995).

La figura dell'*e-tutor* è centrale negli ambienti di apprendimento tecnologici, in quanto consente di operare il passaggio da uno stile di insegnamento centralizzato e direttivo a un modello formativo basato principalmente sulla partecipazione del soggetto in formazione. Un ambiente tecnologico di apprendimento, infatti, prevede la disponibili-

lità, da parte dei diversi attori coinvolti, ad agire con gli artefatti simbolici, culturali e tecnologici (Mantovani, 1995). È una struttura reticolare mediata da oggetti, compiti, relazioni, attività e pratiche che ri/specificano l'apprendimento, le cui linee di indirizzo adottano i principi comportamentisti, cognitivisti e costruttivisti. Le attività progettate sono protese ad attivare processi di assimilazione e accomodamento (Piaget, 1947) e a favorire apprendimenti significativi (Ausubel, 1968; Novak, 2001).

In riferimento alle strategie utilizzate e al livello di interattività richiesto, emerge un processo di autoapprendimento basato su una fruizione individuale dei materiali e un apprendimento in rete o collaborativo che enfatizza le interazioni tra i partecipanti (Trentin, 1998; 1999; 2006). La strategia di apprendimento collaborativo è favorita dall'*e-tutor* nei momenti dedicati alla discussione, nell'organizzazione del lavoro di gruppo, nelle esercitazioni e nelle interazioni. Trentin (2001) sottolinea come tale modalità, diretta da un soggetto esperto, vada differenziata dall'apprendimento collaborativo mutuato o reciproco, tipico di una comunità di pratica, che si sostanzia principalmente nell'assistenza reciproca e nella condivisione della conoscenza, finalizzata al raggiungimento di uno specifico obiettivo formativo.

Nella FaD di terza generazione, la partecipazione attiva dei soggetti nel percorso formativo e le esperienze *peer to peer* sono gli elementi alla base di una *Virtual Learning Community* (Calvani & Rotta, 1999). Il sapere si costruisce in un ambiente generato dalla connettività, la Rete, e si definisce per essere partecipativo e astratto, situato e analitico, pratico e riflessivo, individuale e sociale, determinando un apprendimento permanente (Greco, 2017) centrato sulla relazione discente-tutor-discenti, secondo un'ottica ricorsiva. In tal senso, la Rete viene concepita come *medium collaborativo* (Trentin, 2001), uno spazio che induce il singolo a interagire con i membri della comunità e che connette i contenuti formali a esperienze nuove e complesse, ancorate al quotidiano. All'interno dello scenario presentato, si profila l'identità dell'*e-tutor*, dotata di ruoli e funzioni definiti che, se per alcuni aspetti potrebbe richiamare la figura del docente classico, per altri si colloca agli antipodi rispetto alle attività e alle mansioni da assolvere. Un tratto specifico risiede nella gestione della comunicazione sincrona e asincrona, prerogativa essenziale della formazione in rete, che ingloba l'insieme delle posizioni assunte nelle interazioni tra i partecipanti e le variegate forme di messaggistica a supporto delle attività iniziali, in itinere e conclusive (messaggi di apertura, di aiuto, di facilitazione, feedback), riducendo sempre più la distanza sociale tra gli attori coinvolti (Trentin, 1998; 1999; 2001; 2004).



Le modalità di intervento dell'*e-tutor* sono strettamente condizionate dai paradigmi di riferimento di derivazione costruttivista, interazionista o comportamentista, ai quali corrispondono un tutor-animatore, uno *scaffolder* o un istruttore (Galliani & De Waal, 2005a; 2005b), così come dalla progettazione degli ambienti di *e-learning*, centrati sulla trasmissione delle informazioni o sulla costruzione della conoscenza in modalità individuale e/o collaborativa (Charlier, 2000; Strijbos & Martens, 2001).

Questo aspetto viene esplicitato da Mason (1998) nei tre modelli di strutturazione dei corsi e nelle relative azioni praticate dall'*e-tutor*. Nello specifico l'Autore mostra come il ruolo cambi a seconda che si tratti di un corso *content + support, wrap around* o basato su un modello integrato. Il primo modello, altamente strutturato, rinvia a una modalità di autoapprendimento in cui il ruolo del tutor è marginale e viene contattato dallo studente in base alle sue esigenze. Ne scaturisce un percorso di tipo erogativo, focalizzato principalmente sui materiali didattici. Il secondo modello si basa su un apprendimento non strutturato, si caratterizza per una maggiore flessibilità e per contenere una varietà di contenuti da fruire. Il terzo modello è espressione di un apprendimento collaborativo, un'evoluzione della tipologia *wrap around*, dominato dall'interazione tra pari e con il tutor, nel quale i contenuti non sono già dati ma si costruiscono negli interscambi comunicativi, in attività collaborative e cooperative (Calvani & Rotta, 2000; Rotta, 2002; Khan, 2004).

Un'ulteriore distinzione è attribuibile al modello di insegnamento-apprendimento utilizzato, il quale può essere *instructor centered* o *learning team centered*. Nel primo caso la conoscenza è parcellizzata, al fine di renderne più facile l'assimilazione; nel secondo, invece, gli apprendimenti si costruiscono in situazione e sono altamente interconnessi fra loro. La *tutorship* si colloca nel mezzo delle due polarità: nel percorso formativo si alternano momenti predeterminati, rigidi e definiti, ad altri aperti e dinamici, che si costruiscono nel farsi esperienziale. In questa prospettiva, secondo quanto evidenziato dalla letteratura di settore (Rotta & Ranieri, 2005; Rivoltella, 2006), richiamando il modello di Salmon (2000; 2002), è possibile dettagliare le fasi dell'attività di *e-tutoring* in: *e-moderating, access and motivation, socialization, information sharing, knowledge building e development* delle community.

In riferimento alle fasi e ai momenti della formazione (iniziale, intermedio e finale), l'*e-tutor* svolge la funzione *tecnologica*, nel momento iniziale; *sociale*, in itinere; *organizzativo-strutturale*, nella fase di avvio (pianificazione); *valutativa*, durante tutto il processo formativo (Rotta & Ranieri, 2005; Rivoltella, 2006). In questa *escalation*, lo *scaffolding*



subisce una riduzione sostanziale in fase di *fading*, scomparendo in maniera graduale man mano che si avvia uno sviluppo autonomo da parte delle community.

L'*e-tutor* funge così da *scaffolder* sociale, cognitivo, affettivo e tecnologico (Berge, 1995), deputato alla predisposizione di uno spazio d'apprendimento, favorevole alle interazioni e al raggiungimento degli obiettivi formativi. In tale compito si riconosce la capacità di gestire la comunicazione, in modalità sincrona o asincrona, orientando gli utenti nelle differenti conversazioni, monitorando i forum e invogliando alla partecipazione (Shepherd, 2000).

La rassegna narrativa di studi e ricerche, nazionali e internazionali, ha permesso di tracciare un quadro di sintesi sulle attività di *tutoring* nella formazione *e-learning*, cogliendone i punti di forza e le eventuali criticità. I principali quesiti-obiettivi, che hanno orientato la ricerca, sono stati i seguenti:

- L'*e-tutoring* riduce la distanza spaziale tra i partecipanti al processo formativo?
- La *tutorship online* migliora gli apprendimenti?
- Quali strategie di *e-tutoring* sono efficaci nella didattica online?

La risposta al primo quesito viene fornita da alcuni lavori nei quali si riconosce l'importanza di creare momenti dedicati ad attività *e-moderated*, mediante la gestione e la condivisione delle informazioni, traducibili in forme di sostegno cognitivo agli utenti (Kim & Gil, 2007) e azioni di monitoraggio (VanLehn, 2011). Più è efficace lo spazio dedicato alla comunicazione, più il livello di partecipazione dei corsisti risulta regolare (Vanin & Castelli, 2009). Questo è quanto riportato nello studio di Gregori, Zhang, Galván-Fernández e de Asís Fernández-Navarro (2018), nel quale sono state verificate le modalità di *tutorship online* che fronteggiano l'*e-dropout* e/o il mancato completamento dei corsi nei MOOCs (formali, convenzionali e professionali). Tra le tipologie esaminate (interazione *peer to peer*, interazione legata ai contenuti e interazione esperto-discente), l'interazione esperto-discente, oltre ad essere la più impegnativa, risulta significativa ai fini dell'apprendimento. Le piattaforme *Learning Management System* (LMS) rappresentano uno spazio privilegiato per la didattica *online*, in quanto consentono di implementare, al di là dell'aula, pratiche inerenti l'autovalutazione e la co-valutazione, così come di tracciare i tempi spesi e le risorse fruite dai corsisti. In tali azioni, l'*e-tutor* è chiamato, altresì, a gestire le chat e i forum di discussione, al fine di controllare l'apprendimento e gli eventuali progressi del discente. Nei LMS ogni studente può scegliere



e selezionare autonomamente i materiali didattici e può inoltre regolare i modi e i ritmi del proprio percorso (apprendimento autoregolato). Questa possibilità di gestione autonoma, che accresce la motivazione aumentando la partecipazione online degli utenti, va tuttavia sostenuta dalla presenza di un *e-tutor* (Li & Tsai, 2017). Nel supportare gli studenti nel processo di apprendimento, l'*e-tutor* riduce la distanza fisica che si crea negli ambienti di e-learning (Richardson et al, 2015), attraverso una serie di azioni quali: valutazioni diagnostiche, *feedback*, supporto nella gestione delle conoscenze, monitoraggio delle comunicazioni, costruzione di compiti personalizzati.

Nell'operare una mappatura relativa agli stili di *tutorship*, Chae e Shin (2016), pongono l'*e-tutoring* in relazione al coinvolgimento, al rendimento dello studente e al livello di soddisfazione. Nello specifico, si fa riferimento all'*e-tutor* nella veste di esperto, facilitatore, guida e amministratore, dimostrando come ogni stile vada adeguato al relativo ambiente di apprendimento. Tale analisi fornisce la risposta al secondo interrogativo e in buona parte anche al terzo.

Va rimarcato che non tutti gli stili sono validi ed efficienti in termini di costi, tempi e benefici, ma ciò che accomuna le differenti modalità di intervento è il feedback, manifestato mediante i commenti, le pratiche di supporto, in modalità testuale, visiva, audio o mista, piuttosto che le azioni di indirizzo e controllo (Miller, 2012; Kelly, 2014).

La ricerca *evidence based* (Hattie & Timperley, 2007; Hattie, 2009; 2012) corrobora questi risultati. Essa mette in evidenza l'influenza, nella didattica *online*, di azioni significative che enfatizzano l'uso di un feedback continuo, quale efficace strategia d'intervento nel processo di insegnamento-apprendimento. Una particolare attenzione è attribuita al feedback personalizzato e individualizzato, che indirizza lo studente nel percorso online. Non tutti i feedback sono adeguati, quelli validi consentono di esplicitare meglio i criteri valutativi attivando così veri e propri processi metacognitivi. Si parla, pertanto, di *formative feedback* nel momento in cui si colma il divario tra la prestazione effettiva e quella idealizzata, e ciò avviene quando il tutor guida e orienta l'allievo nella risoluzione del problema e non fornisce direttamente la risposta. Il feedback rappresenta quel valido supporto cognitivo che agisce in profondità e che consente di correggere o eliminare gli errori, innescando meccanismi di autoregolazione (Rosenshine, 2009; Robasto & Trincherò, 2015). In tal modo, lo studente è in grado di riflettere sugli errori commessi attraverso momenti di *self verbalization* che innescano azioni metacognitive e auto-valutative.

Nel ricercare le strategie di *e-tutoring* più efficaci nella didattica online (cfr. il terzo quesito), è rilevante il ruolo di facilitatore, adottato



principalmente nei forum di discussione, in quanto stimola la partecipazione dei soggetti e funge da sostegno all'apprendimento (Phirangee, Epp & Hewitt, 2016). La dimensione sociale e affettivo-relazionale, che emerge nei gruppi di discussione, va ad implementare e ad incidere sui processi cognitivi, garantendo in tal modo un apprendimento più duraturo rispetto a quello che si sviluppa in una didattica di tipo frontale (Biasi, D'Aloise & Longo, 2013). Questo risultato prova il valore aggiunto delle comunità online nella formazione di una identità di gruppo, nella quale si negoziano e condividono le informazioni. Dunque, la presenza sociale, nella comunicazione mediata dalla tecnologia, è una condizione necessaria per il successo formativo, nonché una leva strategica che consente di rimarcare il coinvolgimento attivo e il livello di soddisfazione degli studenti (Phirangee, Epp & Hewitt, 2016).

Nella *systematic review*, condotta da Copaci e Rusu (2015), gli studi analizzati hanno permesso di ricostruire un quadro generale dell'*e-tutoring* sulla base dei seguenti criteri di ammissibilità: studenti di età superiore ai 18 anni e attività di tutorato online nell'istruzione superiore, comprese tra il 2010 e il 2015. In quegli anni, e in rimando ai principali criteri indicati, le ricerche hanno evidenziato la relazione esistente tra apprendimento interattivo e sviluppo di competenze professionali; miglioramento di abilità metacognitive, cognitive, sociali e processo di acquisizione delle conoscenze. Dalla rassegna emerge come nei percorsi accademici sia *l'e-tutoring* che il *peer tutoring* favoriscano il coinvolgimento e la motivazione ad apprendere negli studenti. Dal punto di vista tecnologico, la *systematic review* riconosce la centralità delle piattaforme nella gestione dei programmi di tutoraggio online, ma bisognerebbe affinare l'accessibilità per renderle quanto più possibile *user-friendly* e maggiormente centrate sull'utente, così da fornire molteplici opportunità nell'estrapolazione e nell'analisi dei dati. Riprendendo il modello quadripartito di Berge (1995) (manageriale, tecnico, pedagogico e sociale), Martin, Wang e Sadaf (2018) individuano delle strategie efficaci nella *tutorship* online, quali ad esempio: fornire risposte tempestive e personalizzate, essere presente nei forum oppure avvalersi di video interattivi. Nel tentativo di circoscrivere *l'e-tutorship* in base ai quesiti sopra esposti, studi accreditati (Ghislandi, Calidoni, Falcinelli & Scurati, 2008; Van de Vord & Pogue, 2012; Ghislandi, Margiotta & Raffaghelli, 2014) affrontano la problematica del tutorato a distanza in relazione ai tempi della didattica *online* per operare un confronto rispetto alla didattica in presenza. Il tutorato *online*, a differenza di quello in presenza, prevede un impegno maggiore che supera il periodo speso nella parte attiva del corso (Rusu, Copaci & Soo, 2015). Milani, Raffaghelli e Ghislandi (2017) giungono alla conclusione che l'impegno



richiesto nella didattica online, difficilmente quantificabile, risulta maggiore a fronte di quello adottato nell'organizzazione e strutturazione delle attività in presenza. Non a caso, gli autori sottolineano l'assenza, in letteratura, di lavori empirici che analizzano la percezione soggettiva degli *e-tutor* circa il tempo impiegato nella progettazione e nella realizzazione di un percorso *online*.

I partecipanti alle attività in e-learning tendono ad abbandonare i percorsi online (Levy, 2007; Perry et al. 2008; Willging, Johnson, 2009; Bawa, 2016) e il miglioramento dell'esperienza degli studenti per ridurre i tassi di abbandono è l'obiettivo principale di tutti gli studi complementari alla formazione a distanza. Attraverso i *Learning Analytics* è possibile studiare la partecipazione degli utenti e identificare i motivi principali degli abbandoni o le possibili cause degli insuccessi. I *Learning Analytics* non sono dunque un fine, ma il mezzo per analizzare i dati raccolti durante le attività di formazione e delineare delle correlazioni con le situazioni di "disagio": noia, rischio abbandono, scarsa partecipazione e, di conseguenza, scarsi risultati. Parte del ruolo degli *e-tutor* è, quindi, perfettamente in linea con i *Learning Analytics*: identificare delle situazioni a rischio per poter intervenire prima che sia troppo tardi. I *Learning Analytics*, di conseguenza, potrebbero rappresentare un ausilio efficace al tutoring online, lasciando agli *e-tutor* i compiti di individuare delle strategie di intervento che supportino i processi formativi e migliorino l'efficacia della formazione.



3. Riflessioni conclusive

Con lo sviluppo dell'*e-learning*, la formazione viene sempre più co-gestita in ambienti tecnologicamente integrati che se, per alcuni aspetti, valorizzano le potenzialità della rete, per altri perdono i vantaggi derivati da una didattica *faccia-a-faccia*. In generale, le attività di *tutorship* spostano l'asse concettuale da una relazione diadica asimmetrica a un sistema relazionale dominato da strategie di supporto tra un formatore e un formando.

I numerosi studi riportati sottolineano il valore aggiunto dell'*e-tutoring* nei percorsi formativi erogati *online*; a livello didattico viene avvalorata la rilevanza della dimensione socio-relazionale, soprattutto nelle comunicazioni in modalità asincrona, nelle quali la qualità delle interazioni accresce la motivazione ad apprendere degli studenti, riducendo gli abbandoni. Allo stesso modo, l'utilizzo dei *Learning Analytics*, attraverso l'analisi dei dati raccolti durante le attività di formazione, potrebbe favorire l'efficacia dei percorsi diminuendo il tas-

so di abbandoni. Gli stili della *tutorship online*, inoltre, sono condizionati dalle strategie educative adottate e dalla tipologia di ambiente utilizzato per l'erogazione dei percorsi formativi (*Learning and Content Management System* o *Computer Supported Collaborative Learning*).

Dalla rassegna della letteratura emerge un quadro generale dell'*e-tutoring* contraddistinto dalle differenti modalità di intervento che coniugano il virtuale e il reale, il *tutor* e i *tutee*, la dimensione sociale e quella individuale. Le attività di *tutoring online* sono determinanti nelle comunità virtuali nelle quali la pratica si snoda attraverso la reificazione e la condivisione dei significati e, in tale processo, la presenza dell'*e-tutor* è garanzia di una maggiore partecipazione degli utenti.

L'*e-tutor* padroneggia gli strumenti della rete facilitandone l'accesso; codifica e decodifica il messaggio così da adattarlo al destinatario; gestisce le conversazioni e riduce la distanza tra sé e l'altro; organizza e struttura le proprie attività e quelle dell'ambiente virtuale coordinando i differenti gruppi. L'*e-tutor*, inoltre, supera i limiti derivanti dalla separazione spazio-temporale tra studente e docente mediante gli scambi comunicativi che avvengono nei *learning space*, attraverso gruppi di discussione, chat o forum e intensifica la relazione che si instaura tra le istituzioni, gli utenti e il *team* di cui è parte. Per tale ragione è auspicabile investire in una adeguata formazione dell'*e-tutor* (Rossi, Giannadrea & Magnoler, 2007; Lin & Yang, 2013), soprattutto se si considera la molteplicità dei ruoli ad esso associati rispetto alle finalità dei percorsi attivati sottolineando la necessità di promuovere un approccio personalizzato e/o collettivo, in grado di combinare elementi conoscitivi e innovativi, nel quale è decisiva la presenza attiva e regolatoria dell'*e-tutor* nel fornire *feedback*.



Riferimenti bibliografici

- Ausubel D.P. (1968). *Educational Psychology: A Cognitive View*. New York: Holt, Rinehart & Winston.
- Bawa P. (2016). Retention in Online Courses: Exploring Issues and Solutions. A Literature Review. *SAGE Open*, 6 (1), 1-11.
- Berge, Z. L. (1995). Facilitating computer conferencing: Recommendations from the field. *Educational Technology-Saddle Brook NJ-*, 35, 22-22.
- Biasi V., D'Aloise D., & Longo S. (2013). Componenti psicologiche del ruolo del tutor scientifico nell'apprendimento on-line. *ECPS. Journal of Educational, Cultural and Psychological Studies*, 7, 143-159.
- Calvani A., & Rotta M. (1999). *Comunicazione e apprendimento in Internet: didattica costruttivista in rete*. Trento: Erickson.
- Calvani A., & Rotta M. (2000). *Fare formazione in Internet. Manuale di didattica online*. Trento: Erickson.

- Chae S. E., & Shin J. (2016). Tutoring styles that encourage learner satisfaction, academic engagement, and achievement in an online environment. *Interactive Learning Environments*, 24 (6), 1371-1385.
- Charlier B. (2000). *Apprendre à distance: tuteur, un rôle indispensable?* Conférence au CUNIC, Centre Universitaire de Charleroi, 1er septembre 2000.
- Copaci I. A., & Rusu A. S. (2015). A Profile Outline of Higher Education E-Tutoring Programs for the Digital-Native Student Literature Review. *Procedia - Social and Behavioral Sciences*, 209, 145-153.
- Galliani L., & De Waal P. (2005a). *Learning face-to-face, in action and on-line: integrated model of lifelong learning*, in EDEN Conference Proceedings, European Distance and e-Learning Network, Helsinki.
- Galliani L., & De Waal P. (2005b). *Verso un nuovo modello didattico per la formazione degli e-tutor*, in Atti del II Congresso Nazionale della Società Italiana di e-Learning (SI e-L).
- Ghislandi P., Calidoni P., Falcinelli F., & Scurati, C. (2008). E-university: a cross-case study in four Italian universities. *British Journal of Educational Technology*, 39(3), 443-455.
- Ghislandi P., Margiotta U., & Raffaghelli J.E. (Eds.).(2014). Scholarship of teaching and learning: per una didattica universitaria di qualità. *Formazione & Insegnamento*. European Journal of Research on Education and Teaching, 12 (1), 1-289.
- Greco G. (2017). L'apprendimento nell'era della connettività: una riflessione sociologica al confine tra comunicazione ed educazione. In C.M. Scarcelli, R. Stella (Eds.), *Digital literacy e giovani. Strumenti per comprendere, misurare e intervenire* (21-31). Milano: Franco Angeli.
- Gregori E.B., Zhang J., Galván-Fernández C., & de Asís Fernández-Navarro F. (2018). Learner support in MOOCs: Identifying variables linked to completion. *Computers & Education*, 122, 153-168.
- Hattie J.A.C., & Timperley H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81-112.
- Hattie J.A.C. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. London: Routledge.
- Hattie J.A.C. (2012). *Visible learning for teachers*. London, UK: Routledge.
- Kaye A.R. (1994). Co-Learn: an ISDN-based multimedia environment for collaborative learning. In R.D. Mason, P.D Bacsich (Eds.). *ISDN Applications in education and training* (179-200). London: The Institute of Electrical Engineers.
- Keegan D. (1994). *Principi di istruzione a distanza*. Firenze: La Nuova Italia.
- Kelly R. (2014). Five things online students want from faculty, Luzar, J. Reprinted from, What online learners want. *Online Classroom*, 13(2013), 1-2.
- Khan B. (2004). *E-learning: progettazione e gestione*. Trento: Erickson.
- Kim J., & Gil Y. (2007). Incorporating tutoring principles into interactive knowledge acquisition. *International Journal of Human-Computer Studies*, 65(10), 852-872.
- Levy Y. (2007). Comparing dropouts and persistence in e-learning courses. *Computers & Education*, 48, 185-204.
- Li L.Y., & Tsai C.C. (2017). Accessing online learning material: Quantitative behavior patterns and their effects on motivation and learning performance. *Computers & Education*, 114, 286-297.

- Lin W.C., & Yang S. C. (2013). Exploring the roles of Google.doc and peer e-tutors in English writing. *English Teaching: Practice and Critique*, 12(1), 79-90.
- Mantovani G. (1995). *Comunicazione e identità*. Bologna: Il Mulino.
- Martin F., Wang C., & Sadaf A. (2018). Student perception of helpfulness on facilitation strategies that enhance instructor presence, connectedness, engagement and learning in online courses. *The Internet and Higher Education*, 37, 52-65.
- Mason, R. (1998). Models of online courses. *ALN magazine*, 2(2), 1-10.
- Milani M., Raffaghelli J.E., & Ghislandi P. (2017). Fuori orario. Il tempo docente nella didattica online. *Italian Journal of Educational Technology*, 25(3), 35-54.
- Miller J. M. (2012). *Finding what works online: Online course features that encourage engagement, completion, and success*. Northridge: California State University.
- Nipper S. (1989). Third generation distance learning and computer conferencing. In R.D. Mason, A.R. Kaye (Eds.), *Mindweave: Communication, computers and distance education* (63-73). Oxford: Pergamon Press.
- Novak J.D. (2001). *L'apprendimento significativo: le mappe concettuali per creare e usare la conoscenza*. Trento: Erickson.
- Perry, B., Boman, J., Care, W. D., Edwards, M., & Park, C. (2008). Why do students withdraw from online graduate nursing and health studies education? *The Journal of Educators Online* 5(1).
- Phirangee K., Epp C.D., & Hewitt J. (2016). Exploring the relationships between facilitation methods, students' sense of community and their online behaviours. Special issue on online learning analytics. *Online Learning Journal*, 20 (2), 134-154.
- Piaget J. (1947). *La psychologie de l'intelligence*. Paris: Colin.
- Quaglino G.P. (2005). *Fare formazione. I fondamenti della formazione e i nuovi traguardi*. Milano: Raffaello Cortina.
- Richardson J.C., Koehler A.A., Besser E.D., Caskurlu S., Lim J. & Mueller C.M. (2015). Conceptualizing and investigating instructor presence in online learning environments The International Review of Research. *Open and Distributed Learning*, 16 (3), 256-297.
- Rivoltella P.C. (Ed.).(2006). *E-Tutor. Profilo, metodi, strumenti*. Roma: Carocci Faber.
- Robasto D., & Trincherò R. (Eds.).(2015). *Strategie per pensare. Attività evidence based per migliorare la didattica e gli apprendimenti in aula*. Milano: Franco Angeli.
- Rogers C. (1969). *Freedom to Learn: A View of What Education Might Become*. (1st ed.) Columbus, Ohio: Charles Merrill.
- Rosenshine B. (2009). The empirical support for direct instruction. In S. Tobias, T.M. Duffy (Eds.), *Constructivist instruction. Success or failure?* (201-220). London: Routledge.
- Rossi P.C., Giannandrea L., & Magnoler P. (2007). Tempi e spazi per la formazione: un modello per l'on line. *Querty*, 2 (1), 31-46.
- Rotta M. (2002), Il tutor online: tipologie. *Form@re, Open Journal per la formazione in rete*, 8, <http://formare.erickson.it/wordpress/it/2002/il-tutor-online-tipologie-2/>.



- Rotta M., & Ranieri M. (2005). *E-tutor: identità e competenze. Un profilo professionale per l'e-learning*. Trento: Erickson.
- Rusu A. S., Copaci I. A., & Soos A. (2015). The impact of service-learning on improving students' teacher training: Testing the efficiency of a tutoring program in increasing future teachers' civic attitudes, skills and self-efficacy. *Procedia-Social and behavioral sciences*, 203, 75-83.
- Salmon G. (2000). *E-moderating: The key to teaching and learning online*. London: Kogan Page.
- Salmon G. (2002). *E-tivities: the key to active online learning*. London: Routledge Falmer.
- Srijbos J. W., & Martens R. L. (2001). Group-based learning: Dynamic interaction in groups. In P. Dillenbourg, A. Eurelings, K. Hakkarainen (Eds.). *European perspectives on computer-supported collaborative learning: proceedings of the 1st European conference on computer-supported collaborative learning* (569-576). Maastricht: Maastricht University.
- Shepherd C. (2000). *Competence Frameworks – Online Tutors*. London: Institute of IT Training.
- Toffler A. (1980). *The Third Wave*. New York: William Morrow.
- Trentin G. (1998). *Insegnare e Apprendere in Rete*. Bologna: Zanichelli.
- Trentin G. (1999). Qualità nella formazione a distanza. Una analisi del rapporto tra qualità, interattività e costi nella formazione a distanza. *TD*, 16 (1), 10-23.
- Trentin G. (2001). *Dalla formazione a distanza all'apprendimento in rete*. Milano: Franco Angeli.
- Trentin G. (2004). Networked Collaborative Learning in the Study of Modern History and Literature. *Computers and the Humanities*, 38, 299-315.
- Trentin G. (2006). Apprendimento collaborativo in rete e didattica universitaria: i ritorni di tipo educativo. *TD - Rivista di Tecnologie Didattiche*, 38, 5-11.
- Van de Vord R., & Pogue K. (2012). Teaching time investment: Does online really take more time than face-to-face? *The International Review of Research. Open and Distributed Learning*, 13(3), 132-146.
- VanLehn K. (2011). The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. *Educational Psychology*, 46(4), 197-221.
- Vanin L., & Castelli S. (2009). Gli interventi del tutor in forum di discussione online. Da un modello teorico agli aspetti applicative. *Qwerty*, 4(2), 140-159.
- Vygotskij L.S. (1978). *Mind and Society. The development of higher psychological processes*. Cambridge, Ma: Harvard University Press.
- Wenger E. (1998). *Community of practice*. Cambridge, UK: Cambridge University press.
- Willging, P. A., & Johnson, S. D. (2009). Factors that influence students' decision to dropout of online courses. *Journal of Asynchronous Learning Networks*, 13(3), 115-127.
- Wood D.J., Bruner J. S., & Ross G. (1976). The Role of Tutoring in Problem Solving. *Journal of Child Psychiatry and Psychology*, 17, 89-100.

Come possiamo farlo succedere?

Dalla descrizione alla previsione negli ambienti di apprendimento on line

How can we make it happen?

From description to foresight of online learning environments

Michele Baldassarre

University of Bari "Aldo Moro", michele.baldassarre@uniba.it

Valeria Tamborra

University of Foggia, valeria.tamborra@unifg.it

This paper proposes a Learning Analytics approach aimed at the study of event-driven data coming from 18 courses activated within a Moodle platform concerning two categories: post-graduate training courses and courses activated in the context of larger research projects. The objective is to identify possible patterns of interaction on the part of users enrolled in the two course categories in order to give useful indications for an effective Learning Design of future courses.

The context of higher education, in particular the post-graduate one, is the field in which research on the field can find more fertile ground: training is often provided in e-learning mode, consistently with the need to make the training flexibly responding to the needs of an adult user, inserted in a context of lifelong learning.

Keywords: Learning Analytics; Lifelong Learning; Moodle; Learning Design; Predictive Analytics

Il presente contributo propone un approccio di Learning Analytics rivolto allo studio di event-driven data provenienti da 18 corsi attivati all'interno di una piattaforma Moodle afferenti a due categorie: corsi di formazione post-lauream e corsi attivati nell'ambito di progetti di ricerca. L'obiettivo è quello di individuare possibili pattern d'interazione da parte degli utenti iscritti alle due categorie di corso al fine di dare indicazioni utili alla progettazione efficace di futuri corsi.

Il contesto della formazione superiore, in particolare quella post-lauream, è il campo in cui la ricerca sul tema dei Learning Analytics può trovare maggiore terreno fertile: la formazione è sovente erogata in modalità e-learning, coerentemente alla necessità di rendere la didattica flessibilmente rispondente ai bisogni di un'utenza adulta, inserita in un contesto di lifelong learning.

Parole chiave: Learning Analytics; Lifelong Learning; Moodle; Learning Design; Predictive Analytics

Il contributo rappresenta il risultato di un lavoro congiunto degli autori, tuttavia si attribuisce a Michele Baldassarre la scrittura dei paragrafi 1, 2, 3, 5 e a Valeria Tamborra del paragrafo 4



1. Introduzione

Negli ultimi decenni c'è stato un interesse crescente da parte di decisori politici, industrie ed enti pubblici verso lo sviluppo di indicatori della qualità delle pratiche di insegnamento e apprendimento (Bloxham, & Boyd, 2012) che ha alimentato la ricerca nell'ambito della *educational accountability* (Lockyer, Heathcote, & Dawson, 2013). Nell'ambito universitario, ci si avvale a livello istituzionale di indicatori costruiti sulla base dei questionari agli studenti, dei livelli di progressione e *drop-out* (Coates, 2005). L'ampio impiego di tecnologie educative, come i Learning Management Systems (LMS), ha generato un vasto set di dati di apprendimento alternativi e accessibili (Greller, & Drachler, 2012; Pardo, & Kloos, 2012). Nei contesti di formazione online, le interazioni degli studenti con le attività del corso vengono registrate in tempo reale e questi dati digitali possono essere raccolti e analizzati per valutare in modo più efficace l'apprendimento e il grado di coinvolgimento degli studenti. Essi, in sostanza, attraverso opportune tecniche di Learning Analytics che ne permettano l'interpretazione, possono potenzialmente fornire conoscenze utili a prendere decisioni in merito alle future pratiche di insegnamento e apprendimento (Locker, Heathcote, Dawson, 2013).



Nel presente contributo viene proposta un'analisi dei dati di reportistica del LMS Moodle della piattaforma di e-learning del Dipartimento di Scienze della Formazione, Psicologia, Comunicazione dell'Università degli Studi di Bari, con l'obiettivo di valutare alcuni aspetti salienti relativi all'andamento delle attività didattiche al fine di individuare indicazioni utili alla progettazione efficace di futuri corsi.

Il fine, dunque, è quello di esplorare i dati per individuare delle possibili regolarità strutturali che descrivano le modalità di interazione degli utenti in contesti di formazione online differenti: l'educazione formale finalizzata all'acquisizione di un titolo accademico post-lauream e l'educazione formale erogata nell'ambito di corsi di aggiornamento organizzati nell'ambito di progetti attivati per scopi di ricerca. L'individuazione di *pattern* d'interazione in target di studenti differenti consentirebbe di progettare questi corsi in modo specifico e differenziato, sì da stabilire un dialogo tra due discipline: Learning Analytics e Learning Design. Il Learning Design descrive e documenta il modo in cui vengono progettate le attività didattiche fornendo direttive utili all'erogazione di corsi in grado di raggiungere efficacemente gli obiettivi auspicati. Il Learning Design, come *frame* teorico, e i Learning Analytics, come prassi di ricerca, possono agire in sinergia per migliorare la

comprensione e la valutazione dell'andamento delle attività didattiche e dei *learning outcomes* (Lockyer, Heathcote, Dawson, 2013).

2. Learning Analytics nei contesti formali di educazione superiore

I metodi di Learning Analytics implicano la raccolta, analisi e presentazione di dati associati ai comportamenti degli studenti in contesti di apprendimento online. Diversamente dalle tecniche di raccolta dati tradizionali, come i questionari e i focus group, che contano su partecipanti che, al contempo, vogliono fornire un feedback, ricordino e riferiscano in modo accurato eventi passati, le tecniche di Learning Analytics permettono di raccogliere dati su un comportamento specifico e osservabile in tempo reale e in modo automatico (Lockyer, Heathcote, Dawson, 2013).

Questo ambito di ricerca si è sviluppato a partire da contesti diversi:

- *business intelligence*: analisi di dati svolte in ambito aziendale con l'obiettivo di migliorarne il vantaggio competitivo sul mercato (Mitchell & Costello, 2000);
- *web analytics*: sistemi di tracciamento e profilazione degli utenti dei siti web con finalità di marketing mirato (Zaiane, 2001);
- *educational data mining*: ambito di ricerca che si occupa di individuare metodi di analisi esplorativa di dati estratti dagli ambienti di apprendimento basati sul web (educationdatamining.org; Zaiane, 2001; Romero, & Ventura, 2007);
- *social network analysis*: analisi delle reti sociali applicate agli ambienti di apprendimento a distanza con l'obiettivo di esplorare e approfondire il senso di comunità che si stabilisce in questi contesti (Dawson, 2008);

I primi approcci avevano una forte vocazione tecnica volta alla pura analisi dei dati. Nel tempo, l'attenzione si è sempre più spostata dall'individuazione di tecniche di analisi, alla definizione di modelli volti a comprendere e ottimizzare i processi di apprendimento (Ferguson, 2014).

Nel campo delle tecniche di analisi, la letteratura individua le seguenti categorie: modelli descrittivi (Shmueli, 2010), esplicativi, predittivi (Brooks, & Thompson, 2017) e prescrittivi (Rezzani, 2015).

I modelli descrittivi sono finalizzati a rappresentare la struttura dei dati in modo sintetico (Shmueli, 2010). Le analisi descrittive si concentrano sui dati raccolti da strumenti di reportistica; questi dati ven-



gono categorizzati e aggregati al fine di applicare funzioni statistiche descrittive (Rezzani, 2015). Nei modelli esplicativi, l'obiettivo è quello di usare tutte le evidenze disponibili per dare una spiegazione a un certo fenomeno (seguendo un approccio confermativo). Nelle analisi predittive il fine è quello di creare un modello in grado di predire i valori di nuovi dati (Brooks, & Thompson, 2017).

Diversamente dai modelli esplicativi, in quelli descrittivi non c'è dipendenza da una teoria di fondo. Il focus, pertanto, non è posto a livello di costruito, quanto a livello di misurabilità dei fenomeni. Rispetto ai modelli predittivi, invece, i modelli descrittivi non sono finalizzati alla previsione, ma alla descrizione di un determinato contesto (Shmueli, 2010).

Sia le ricerche con scopi predittivi che quelle con scopi esplicativi, partono dai presupposti determinati da analisi esplorative. L'*Exploratory Data Analysis* (EDA) consiste nel riassumere e sintetizzare i dati numericamente e graficamente riducendone la dimensione e preparandoli a una fase successiva maggiormente formale. Anche le analisi EDA assumono forme differenti a seconda dello scopo della ricerca che verrà condotta nelle fasi successive, se di natura predittiva o esplicativa (Shmueli, 2010).

Nella ricerca pedagogica, le analisi predittive si inseriscono in un più ampio contesto di politiche educative orientate all'intervento, in cui le istituzioni si avvalgono di questi modelli per rispondere adattivamente alle necessità contingenti degli studenti. Lo scopo di un modello predittivo è quello di costruire un sistema complesso in grado di descrivere accuratamente i risultati di apprendimento (*learning outcomes*) di ciascuno studente inserito in un determinato contesto di formazione a distanza (Brooks, & Thompson, 2010).

Lockyer, *et al.* (2013), evidenziano che i Learning Analytics possono fornire anche misure più sofisticate del processo di apprendimento degli studenti, che possono guidare i docenti nella progettazione, sviluppo e revisione dei corsi. Rezzani (2015), definisce questo tipo di ricerche "*prescriptive analytics*": un ambito che nasce nel contesto delle analisi predittive ma che, superandole, fornisce regole direttamente applicabili per migliorare un determinato contesto.

Volendo sintetizzare quanto sin ora discusso, si propone di seguito uno schema adattato da Baldassarre (2016).



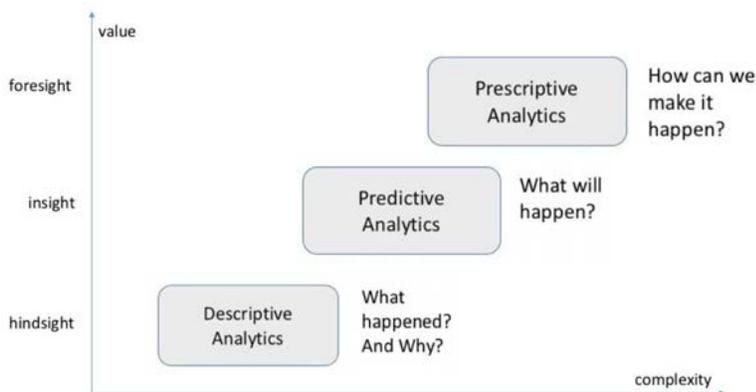


Fig. 1: Dai modelli descrittivi a quelli prescrittivi
(fonte: Baldassarre, 2016; adattato da Daniel, 2016)

Lo schema riassume finalità e risultanze dei modelli di ricerca nel campo dei Learning Analytics, raggruppandoli in Descriptive Analytics (che comprendono i modelli descrittivi e quelli esplicativi), Predictive Analytics (modelli predittivi) e Prescriptive Analytics (modelli prescrittivi) e categorizzandoli rispetto al livello di complessità e al valore dei risultati che permettono di raggiungere.



- Descriptive Analytics (cosa è accaduto e perché?): modelli di analisi a basso livello di complessità che hanno l'obiettivo di spiegare a posteriori un certo fenomeno;
- Predictive Analytics (cosa potrebbe accadere?): modelli di analisi di media complessità che hanno l'obiettivo di prevedere l'andamento di un certo fenomeno;
- Prescriptive Analytics (in che modo possiamo far sì che si verifichi un certo scenario?): modelli di analisi che mirano a individuare orientamenti e indicazioni utili a determinare *learning outcomes* auspicati.

La ricerca che persegue fini prescrittivi stabilisce un dialogo ideale con la scienza del Learning Design fornendo indicazioni utili a progettare corsi di alta qualità in grado di raggiungere efficacemente obiettivi auspicati (Laurillard, 2012). Si individua, pertanto, una relazione circolare tra Learning Design e Learning Analytics: da un lato la progettazione definisce gli obiettivi e i piani educativi dei quali i Learning Analytics possono valutare la qualità in termini di *effectiveness* e *outcomes* del processo di apprendimento, fornendo una prospettiva più olistica sull'impatto delle attività progettate (Lockyer, Heathcote, &

Dawson, 2013); dall'altro, gli esiti delle ricerche nel campo dei Learning Analytics aprono scenari che consentono al Learning Design di ripensare la progettazione dei corsi.

3. Il contesto e la fonte dei dati

Le analisi proposte all'interno del contributo si soffermano sui log-data della piattaforma Moodle del Dipartimento di Scienze della Formazione, Psicologia, Comunicazione dell'Università degli Studi di Bari "Aldo Moro" al fine di individuare possibili *pattern* di interazione tra gli utenti e il LMS che possano fornire informazioni utili a ripensare, per migliorare, l'offerta formativa dei corsi in e-learning.

I log-data possono essere definiti "*event-driven data*" (Alhadad, *et al.*, 2015), ossia dati basati sulle attività degli studenti raccolti da un LMS con il quale gli studenti interagiscono.

La piattaforma e-learning del Dipartimento è così organizzata:

- 5 sezioni
 - Corsi di Laurea Triennale
 - Corsi di Laurea Magistrale
 - Corsi di Formazione Post-Lauream
 - Progetti di Dipartimento
 - Progetti in convenzione
- 34 corsi
- 254 moduli didattici
- 2034 utenti, di cui:
 - 328 utenti che non hanno mai effettuato il log-in (16,13%)
 - 1706 utenti che hanno effettuato almeno un accesso online (definibili "potenzialmente attivi").

Nel Dipartimento l'uso che i docenti fanno dell'e-learning si concentra prevalentemente sui corsi di formazione post-lauream e sui progetti di ricerca/formazione con altri enti convenzionati che si rivolgono comunque a un'utenza adulta già laureata (sovente progetti che prevedono attività di formazione in servizio dei docenti appartenenti agli istituti scolastici partner dei progetti). L'educazione degli adulti si configura come un contesto in cui l'e-learning diviene risorsa fondamentale per un'offerta formativa che, flessibilmente, si adatta alle esigenze dell'utenza. Tuttavia, sebbene le due categorie di utenti cui i corsi si rivolgono siano tra loro assimilabili, perché composte da corsisti adulti e lavoratori, gli scopi di questi due tipi di corsi sono differenti.



I corsi di formazione post-lauream mirano all'acquisizione di un titolo universitario e sono sottoposti a un regolamento che incide sul loro andamento e sulle modalità di acquisizione del titolo, nonché di certificazione ufficiale del superamento di specifici requisiti per la sua acquisizione. Le risorse caricate online sono finalizzate allo studio puntuale e, sovente, i corsisti devono svolgere prove di valutazione intermedie e finali. Inoltre, le attività proposte rispettano l'organizzazione di un piano didattico formalmente elaborato.

I corsi attivati online nell'ambito di progetti svolti in convenzione con altri enti sono, spesso, la soluzione scelta a fronte dell'esigenza di creare uno spazio virtuale di condivisione di materiale di approfondimento, di opinioni ed esperienze, nonché di materiale elaborato dagli utenti per scopi di ricerca. Gli utenti iscritti a questi corsi, generalmente, non hanno versato quote di iscrizione e non acquisiscono un titolo universitario. Anche l'impostazione del corso stesso è generalmente meno focalizzata sul controllo e la certificazione delle attività svolte online dai corsisti, né sull'obbligatorietà delle attività proposte. I livelli di partecipazione di questo tipo di utenza potrebbero, pertanto, essere di natura diversa rispetto a quelli degli utenti iscritti a un corso universitario post-lauream.

Le analisi si sono concentrate, dunque, su una selezione dei corsi attivati all'interno di queste due sezioni della piattaforma. Dalla selezione dei corsi sono stati esclusi quelli aperti ma che non sono stati realmente attivati dai docenti responsabili o che sono risultati inattivi.

La base d'indagine, dunque, è costituita da 18 corsi, equamente suddivisi all'interno delle due sezioni oggetto d'analisi. I dati sono stati raccolti mediante i plugin di Moodle Intelliboard e Analytics Graphs.

4. L'analisi dei dati

L'analisi si è concentrata prevalentemente sui dati di accesso alle risorse al fine di trarne inferenze utili a individuare opportune indicazioni per la riprogettazione degli ambienti di apprendimento.

Si procederà, nel seguito del lavoro, a presentare i dati generali di accesso alle risorse, nonché il numero di attività registrate nei diversi corsi attivi in piattaforma; successivamente si presenteranno, nel dettaglio, i dati di accesso ai forum e alle videolezioni. Sono state escluse da analisi più approfondite le risorse finalizzate alla valutazione, compiti e quiz, perché le analisi preliminari hanno mostrato risultati insufficienti per poter desumere *pattern* comportamentali significativi.

Per una comprensione adeguata di quanto verrà discusso di seguito,



si ritiene necessario specificare il significato di alcuni termini specifici relativi alla lettura della reportistica di Moodle.

- Attività: qualunque tipo di risorsa didattica caricata online; può essere di diversi tipi: forum, file, video-lezione, e-book, chat, quiz, compito, ecc.;
- Accesso/visualizzazione: voce del database dei log-data che indica che l'utente ha cliccato su una determinata risorsa del corso. I LMS registrano automaticamente ciascun click eseguito dagli utenti al suo interno, inserendo questo tipo di azione nella categoria delle visualizzazioni/letture;
- Post in un forum: si definisce "post" un intervento inserito da un utente all'interno di un forum;
- Discussione in un forum: si definisce "discussione" un gruppo di post inseriti da più utenti all'interno di un forum sotto uno specifico topic di discussione. Una discussione contiene come minimo un post, ossia, almeno il primo intervento inserito da un utente all'interno di un forum sotto l'etichetta di un certo argomento scelto dall'utente stesso. Non sempre, dunque, sotto la categoria "discussione" vi sono dei reali confronti tra più utenti. A volte si può trattare di un argomento di discussione aperto da un utente ma che non riceve alcuna risposta.



4.1. *Statistiche generali della piattaforma*

Come detto in premessa, la base dei dati è costituita da 18 corsi suddivisi in 2 macro-sezioni tematiche: corsi inerenti la formazione post-lauream (d'ora in poi "post-lauream") e corsi inerenti progetti di dipartimento in convenzione con enti esterni (d'ora in poi "progetti in convenzione"). Per chiarezza espositiva si elencano gli acronimi dei corsi afferenti alle due sezioni:

- Post-lauream: SM-C4E, CQEPSP, CDD S3, TIC-TFA, DDS, SM-PI, SM-DDEV, DidaLim, CSE;
- Progetti in convenzione: RdC, INSO, MIP, NSNV, PDM, MSV, DCL, IDS, RAV.

Si propone di seguito una prima *overview* dei corsi considerando il rapporto tra il numero di utenti iscritti e di attività registrate all'interno dei singoli corsi al fine di comprendere il livello di *engagement* che ciascun corso ha ottenuto.

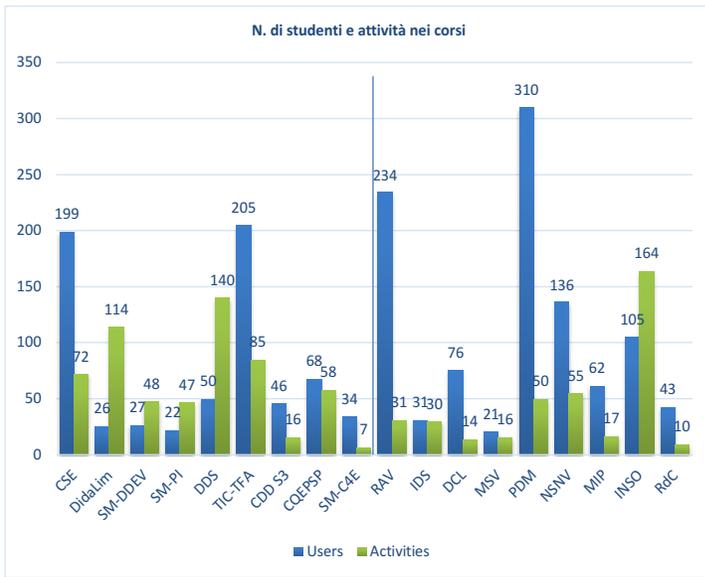


Fig. 2: N. di utenti iscritti e di attività registrate all'interno dei corsi



In 13 dei 18 corsi presi in considerazione il numero di utenti è superiore al numero di attività registrate. Ciò suggerisce un significativo tasso di inattività.

Per comprendere in modo più ampio questo dato si propongono, di seguito, due grafici che mostrano i dati aggregati per macro-sezioni riferiti al numero di utenti iscritti e attività registrate nei corsi, nonché il tasso di partecipazione calcolato dividendo il numero di attività per il numero di utenti registrati.

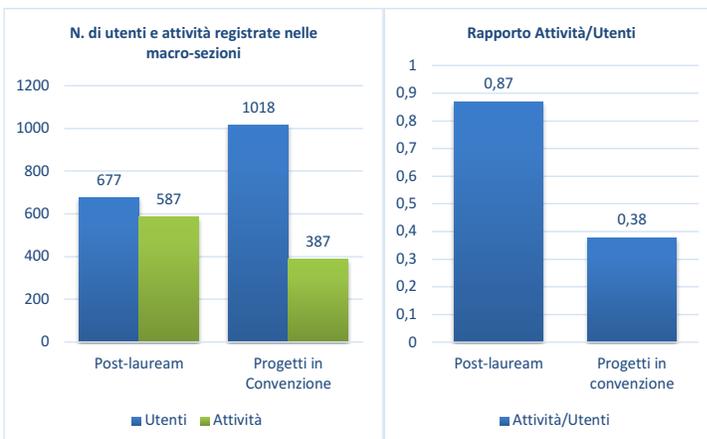


Fig. 3: Rapporto del n. di utenti iscritti e di attività registrate all'interno delle due macro-sezioni

I grafici confermano quanto sostanzialmente rilevato in precedenza, ossia un significativo livello di inattività registrato nella maggior parte dei corsi. Il numero di utenti è generalmente superiore al numero di azioni svolte all'interno dei corsi. In particolare, si registra un maggior livello di inattività all'interno dei corsi attivati nella sezione dei progetti in convenzione.

Entrando nel merito delle risorse proposte all'interno dei corsi, il grafico seguente presenta la distribuzione delle risorse all'interno delle due macro-sezioni considerate per tipologia: chat, forum, book, link, cartelle e files, video, assessment, quiz.

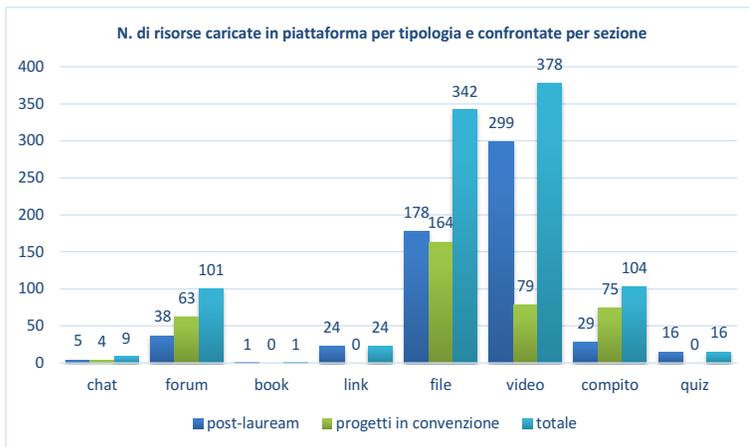


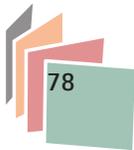
Fig. 4: N. di risorse caricate in piattaforma per tipologia distribuite nelle macro-sezioni

I dati mostrano che le risorse didattiche maggiormente caricate all'interno dei corsi, in generale, sono file e video-lezioni, seguite dai compiti di valutazione e i forum.

All'interno dei corsi attivati nella sezione post-lauream vengono generalmente caricate più risorse e di natura maggiormente differenziata rispetto ai corsi della sezione progetti in convenzione.

Si sottolinea, in particolare, il dato relativo alle risorse finalizzate alla valutazione: compiti e quiz. Si registra una prevalenza di compiti rispetto ai quiz, che, tra l'altro, sono completamente assenti nella sezione progetti in convenzione. In quest'ultima sezione, inoltre, il numero di risorse "compito" è nettamente superiore. Questo dato suggerisce la natura sostanzialmente avalutativa di questi corsi, finalizzati prevalentemente alla condivisione di materiale didattico e l'elaborazione di artefatti principalmente per scopi di ricerca.

Anche il numero dei forum aperti nei corsi dei progetti in conven-



zione è quasi il doppio rispetto a quelli presenti nei corsi post-lauream, avallando l'idea che i corsi attivati nell'ambito di progetti in convenzione nascono prevalentemente con lo scopo di costruire uno spazio virtuale per la condivisione e il confronto.

Per avere un primo dettaglio sui dati di accesso alle risorse, si propone, di seguito, un grafico che mostra le percentuali di utenti che hanno e non hanno visualizzato le risorse caricate nei corsi.

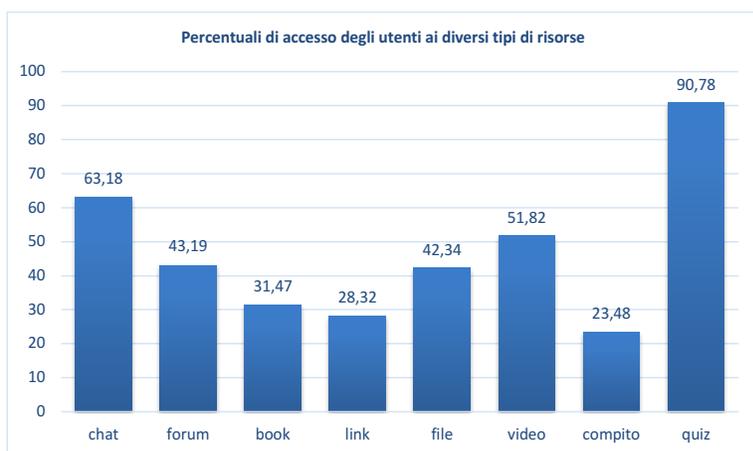
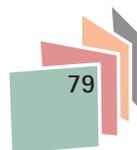


Fig. 5: Percentuali di accesso degli utenti ai diversi tipi di risorse

Il grafico mostra una complessiva tendenza negativa nell'accesso alle risorse. Le principali risorse presenti nei corsi registrano percentuali di visualizzazione inferiori rispetto a quelle relative agli utenti che non le hanno visualizzate. Si richiama, in particolar modo, l'attenzione sulle percentuali di accesso a forum, video, compiti e quiz: i video e i quiz sono stati prevalentemente caricati all'interno dei corsi post-lauream e hanno anche registrato una percentuale maggiore di utenti che vi hanno fatto accesso, rispetto a coloro che non l'hanno fatto; i forum e i compiti, invece, prevalentemente presenti nei corsi dei progetti in convenzione, si registrano percentuali di non-accesso maggiori. Questo dato suggerisce che i corsi post-lauream, poiché finalizzati all'acquisizione di un titolo e con un'utenza di studenti iscritti all'Università, presentano un maggior livello di strutturazione, obbligatorietà e partecipazione rispetto ai corsi dei progetti in convenzione la cui utenza, generalmente, è meno vincolata ad attività strutturate.

Per comprendere meglio quanto questi dati suggeriscono, si propongono di seguito le percentuali di accesso alle risorse, considerate, nel dettaglio, per macro-sezione.



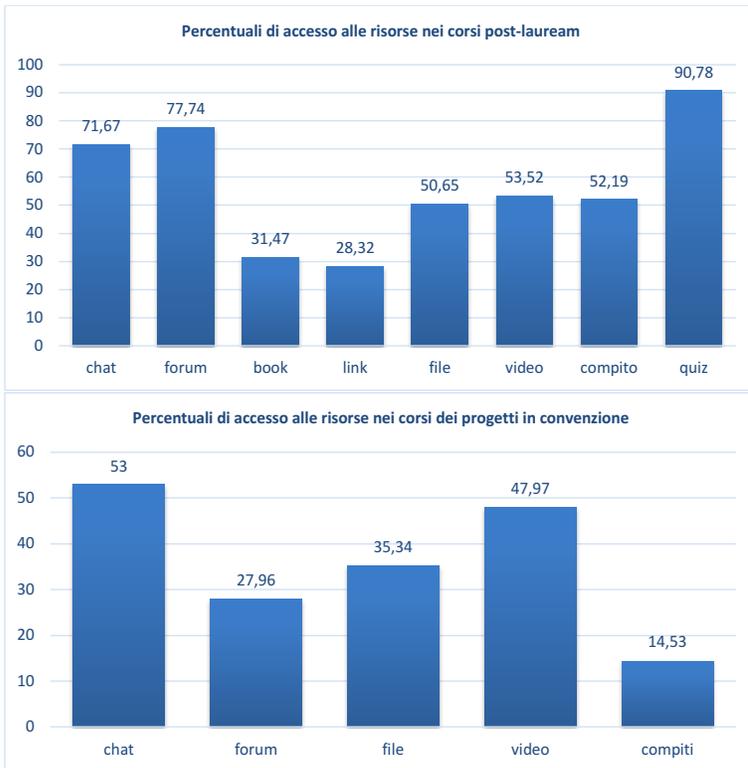


Fig. 6: Percentuali di accesso degli utenti ai diversi tipi di risorse considerati per macro-sezione

I dati confermano la tendenza individuata in precedenza in merito alle differenze di accesso alle risorse da parte degli utenti dei corsi afferenti alle due macro-sezioni. In particolare, nei corsi post-lauream si registra un generalizzato maggiore livello di accesso alle risorse. L'accesso ai quiz, in particolar modo, registra una percentuale che supera il 90% degli utenti che vi hanno fatto accesso. Nei corsi dei progetti in convenzione si evidenzia una tendenza opposta: tutte le risorse, fatta eccezione per le chat, registrano una maggiore percentuale di non-visualizzazione. In particolare, si nota come, nonostante l'elevato numero di risorse "compito" caricate all'interno di questi corsi, l'85,47% degli utenti non vi hanno fatto accesso, confermando un livello più basso di *engagement*.

4.2. Analisi delle interazioni nei forum

I forum aperti complessivamente all'interno della piattaforma sono 101, dei quali, 38 in corsi appartenenti alla sezione post-lauream e 63 in corsi appartenenti alla sezione progetti in convenzione.

Complessivamente sono state aperte 918 discussioni per un totale di 2460 post. In media, dunque, sono stati inseriti 24,36 interventi in ciascun forum e ciascuna discussione ha registrato in media 2,68 post. Si può, pertanto, considerare che ciascuno degli utenti potenzialmente attivi (1706) ha inserito in media 1,44 post all'interno dei forum. Dei 2460 interventi registrati, il 34% è stato inserito dai docenti, il 66% da studenti.

Si propone di seguito un confronto tra il numero di forum e di discussioni avviate nella sezione post-lauream e nella sezione progetti in convenzione, assieme ai livelli di accesso degli utenti ai forum

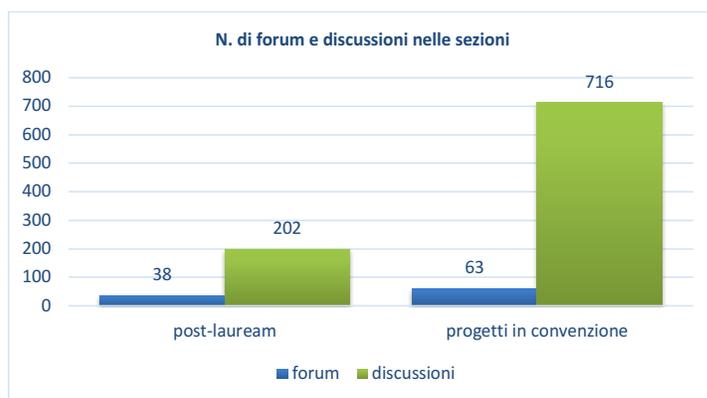


Fig. 7: Numero di forum aperti e discussioni avviate nelle due sezioni della piattaforma.

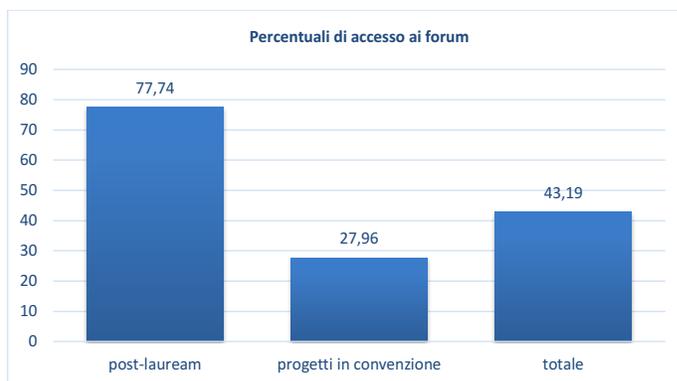


Fig. 8: Percentuali di accesso ai forum

I dati, confermando il maggior numero di forum attivati nei corsi afferenti alla sezione progetti in convenzione, mostrano anche un numero nettamente superiore di discussioni aperte. Tuttavia, si registra, anche, un'alta percentuale di utenti che non hanno mai visualizzato la

risorsa forum nei corsi di questa sezione. Ciò indica, pertanto, un alto numero di discussioni avviate da pochi utenti; discussioni, che però, non sempre sfociano in un confronto perché molte non registrano altri interventi, oltre il primo. Va notato, infine, che i forum messi a disposizione dai docenti dei corsi nella sezione post-lauream sono circa la metà rispetto a quelli disponibili nella sezione progetti in convenzione.

4.3. Accesso e visualizzazione delle video-lezioni

Complessivamente in piattaforma sono state caricate 378 video-lezioni, il 73,81% (279) delle quali in formato *Shareable Content Object Reference Model* (SCORM). Lo SCORM è uno standard che consente, nei contesti di e-learning, di costruire oggetti di apprendimento (tecnicamente definiti *Learning Objects*) caratterizzati da specifiche tecniche che ne consentono l'interoperabilità, ossia il loro uso e riuso in modo indipendente dalla piattaforma che li ospita. Lo standard SCORM, inoltre, conferisce ai *learning objects* metadati che ne consentono la catalogazione, il riuso e il tracciamento.

In prima istanza si presentano i dati relativi alla distribuzione delle video-lezioni nei principali formati utilizzati all'interno delle due macro-sezioni oggetto d'analisi.

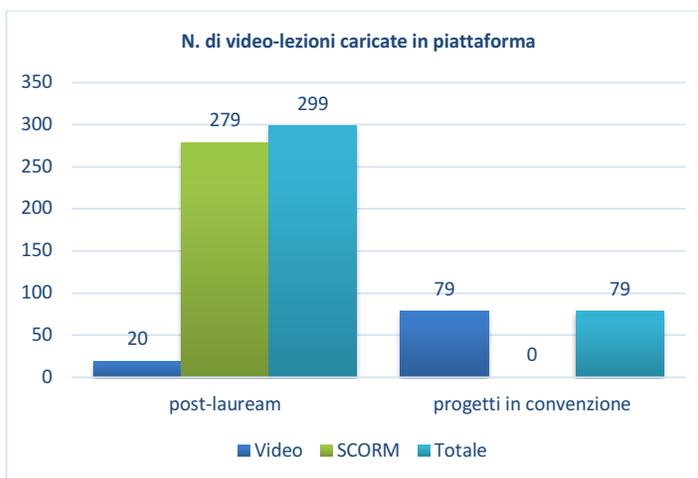


Fig. 9: Numero di video-lezioni caricate nelle due sezioni della piattaforma, distinte per tipologia di formato

Il grafico mostra una differenza piuttosto netta tra i corsi afferenti alla sezione post-lauream e quelli dei progetti in convenzione. Si riscontra, infatti, che il 79,1% (299) delle video-lezioni caricate com-

pletivamente in piattaforma appartiene a corsi afferenti alla sezione post-lauream. Di queste 299 video-lezioni, il 93,31% è stato realizzato in standard SCORM. Le 79 video-lezioni caricate nei corsi afferenti alla sezione progetti in convenzione sono in semplice formato video. Alcune delle analisi successive, pertanto, si concentreranno esclusivamente sulle video-lezioni in formato SCORM perché i dati ottenuti dai video semplici non sono sufficientemente raffinati da consentire particolari approfondimenti oltre al conteggio degli utenti che hanno e che non hanno visualizzato la risorsa.

Prima di procedere a tali analisi, comunque, si propone di seguito un grafico che mostra le percentuali di utenti che hanno visualizzato globalmente le video-lezioni.



Fig. 10: Percentuali di visualizzazione delle video-lezioni

Considerando i dati presentati nei due grafici precedenti si registra, in generale, che più della metà (51,82%) di tutti gli utenti che hanno fatto accesso almeno una volta in piattaforma, ha visualizzato almeno una volta una video-lezione.

Entrando nel merito delle due sezioni, si riscontra, anche per l'accesso alle video-lezioni, una tendenza già registrata in precedenza, ossia una partecipazione più distribuita da parte degli utenti dei corsi post-lauream. I dati di accesso alle video-lezioni nei corsi dei progetti in convenzione mostrano il 52,03% degli utenti di questi corsi che non hanno mai visualizzato queste risorse. Si consideri, inoltre, la netta maggioranza di video-lezioni caricate nei corsi post-lauream rispetto al totale. Questi dati, suggeriscono un uso importante di questa risorsa all'interno dei corsi post-lauream come forma di didattica integrativa



e/o sostitutiva di quella in presenza, elaborata in forme rispondenti alle necessità di impostare la formazione in modo frontale: una modalità d'insegnamento comunemente adoperata per la trasmissione di contenuti teorici secondo un modello trasmissivo. Tale modalità di insegnamento è comunemente adoperata nei contesti formali di istruzione finalizzati all'acquisizione di un titolo, più che in contesti di ricerca-formazione. La video-lezione è la diretta traduzione in e-learning di questo tipo di didattica. I dati, pertanto, risultano coerenti rispetto alla diversa tipologia dei corsi oggetto di analisi.

Per quanto pertiene le modalità di accesso alle video-lezioni si propone, di seguito, un'analisi sul numero e i tempi di visualizzazione. Come anticipato, questi dati si riferiscono esclusivamente alle video-lezioni realizzate in formato SCORM, che sono, rispetto al totale, il 73,81%.

Calcolando il numero di visualizzazioni che ciascuno SCORM ha ottenuto, sono state create 6 categorie di video-lezioni



1. Nessuna visualizzazione: 16 SCORM, pari al 6% del totale;
2. 1 visualizzazione: 29 SCORM (10%);
3. Da 2 a 8 visualizzazioni: 133 SCORM (48%);
4. Da 9 a 15 visualizzazioni: 55 SCORM (20%);
5. Da 16 a 30 visualizzazioni: 9 SCORM (3%);
6. Più di 190 visualizzazioni: 37 SCORM (13%).

Il grafico seguente mostra nel dettaglio i dati sin qui presentati.

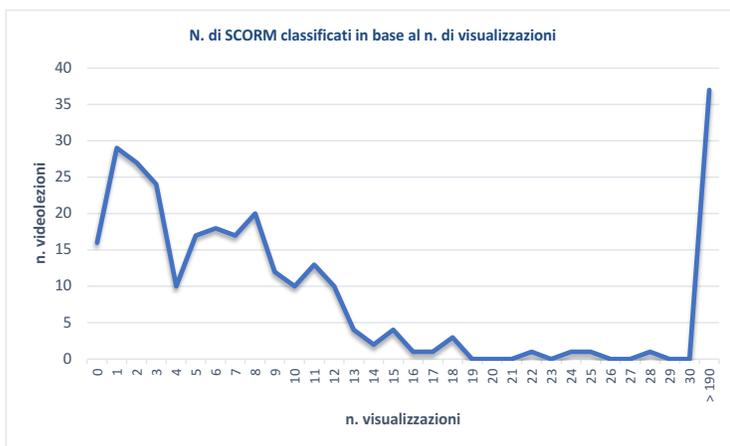


Fig. 11: Numero di SCORM classificati in base al numero di visualizzazioni.

Si noti che la percentuale di SCORM che non sono mai stati visualizzati è relativamente bassa (6% del totale). La maggior parte dei video (48%) sono stati visualizzati da 2 a 8 volte. Segue il 20% visualizzato da 9 a 15 volte. Infine, si registra il 13% dei video che è stato visualizzato da 190 a 270 volte. Vi è, dunque, un netto distacco tra un numero di visualizzazioni medio basso (che si può attestare da 1 a 30 volte) e un gruppo di SCORM che ha registrato un elevato numero di visualizzazioni. Si consideri, comunque, che i 37 SCORM, rappresentanti il 13% del totale, che hanno registrato un numero di visualizzazioni superiore a 190 ciascuna, appartengono a un solo corso in cui si sono registrati livelli importanti di partecipazione da parte degli utenti, risultando essere, pertanto, un outlier rispetto alle tendenze riscontrate a livello generalizzato negli altri corsi.

5. Conclusioni

I dati confermano le assunzioni iniziali (cfr. par. 3) circa la differente natura intrinseca dei corsi: nei corsi di formazione post-lauream, rispetto a quelli dei progetti in convenzione, vi sono meno utenti registrati, più risorse didattiche caricate online e di natura più diversificata; ci sono prove di valutazione sommativa ma meno forum. Il livello di partecipazione in questi corsi è più alto. Le risorse didattiche sono in prevalenza video-lezioni. Sebbene nei corsi dei progetti in convenzione, invece, vi siano più forum, si registrano livelli di partecipazione inferiori rispetto ai corsi post-lauream.

In particolare, dai dati presentati risulta che, riguardo all'uso delle risorse, per i Corsi in Convenzione c'è uno scarto tra quanto è stato progettato e quanto è stato effettivamente rilevato in piattaforma. Infatti, proprio le risorse su cui si è più puntato in sede progettuale sono state quelle meno utilizzate. Questo suggerisce uno scarto tra quanto previsto in sede progettuale e quanto realmente implementato, soprattutto a livello di gestione delle attività con gli utenti. Si ritiene, dunque, che la progettazione delle attività in e-learning all'interno dei corsi blended dovrebbe essere ripensata nella direzione di:

- Un rafforzamento della concertazione delle due modalità di erogazione;
- La riconsiderazione dell'utilità e dell'uso che si intende fare delle risorse online;
- Un uso integrato dei diversi formati;
- Il superamento della separazione tra i due tipi di formazione;



- La costruzione di una continuità circolare all'interno delle attività didattiche.

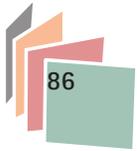
Le attività online dovrebbero incentivare maggiormente il confronto tra pari attraverso un sapiente intervento di tutor specializzati. Esse, inoltre, dovrebbero essere progettate in modo maggiormente strutturato prevedendo lo svolgimento di attività specifiche sui temi proposti all'interno dei materiali di studio e presentate in presenza. Le attività, dunque, dovrebbero essere concertate in modo strettamente interconnesso con quelle in presenza sì da incentivare il lavoro online.

L'interazione nei forum è fondamentale per lo sviluppo della metacognizione e per ridurre il rischio di percezione di isolamento tipico degli ambienti di apprendimento online. Incrementare il livello di interazione all'interno dei forum corrisponde all'obiettivo di aumentare lo spazio dialogico che si instaura nel contesto di apprendimento ed è un aspetto centrale della progettazione didattica dei corsi in e-learning. Nella progettazione di un corso online, pertanto, bisognerebbe riflettere attentamente sulle strategie di costruzione di una comunità di apprendimento e di uno spazio dialogico virtuale che può essere realizzato, in prima istanza, attraverso un'adeguata progettazione della risorsa forum. L'interazione nei forum, infatti, non dovrebbe essere completamente lasciata alla spontanea volontà degli utenti di interagire, ma indirizzata da tutor specializzati.

Il materiale di studio e di approfondimento di tipo testuale dovrebbe offrire sintesi e mappe concettuali al fine di incentivarne l'uso per scopi di studio.

Per incrementare la percezione di controllo sul proprio processo di apprendimento e ridurre, al contempo, la percezione di disorientamento iniziale, si suggerisce di impostare, all'interno del LMS, gli obiettivi del corso e delle scadenze intermedie per il completamento delle attività; un tipo di configurazione disponibile in Moodle che permette, tra l'altro, agli amministratori di accedere a maggiori informazioni circa i livelli di partecipazione degli utenti.

Per quanto concerne i materiali didattici si suggerisce di rendere più omogenei i formati. È, questo, un aspetto centrale delle risorse online quando si lavora nel campo dei Learning Analytics. Sovente i dati non raggiungono un livello sufficiente di raffinatezza proprio a causa della natura tecnica delle risorse didattiche caricate sui LMS; alcune informazioni, proprio per la natura dei formati utilizzati, risultano essere inintelligibili, tale da compromettere la possibilità di svolgere analisi approfondite. Le tecniche di Learning Analytics avanzano, dunque, l'istanza dell'accuratezza dei dati; la sfida è quella di interpretare i dati



ottenuti per lo scopo di valutare il successo educativo di una particolare attività (Dawson, Bkharria, Lockyer, & Heathcote, 2010). L'attenzione ai formati, pertanto, è cruciale nel processo di progettazione didattica di un corso in e-learning. Domande conoscitive in merito ai tempi trascorsi durante lo studio di una singola risorsa didattica, i livelli di completamento delle attività, nonché le modalità di accesso, consentono di avere una comprensione profonda degli stili di apprendimento degli studenti nei contesti e-learning e di progettare, di conseguenza, un'offerta formativa mirata ed efficace.

Riferimenti bibliografici

- Alhadad S., Arnold K., Baron J., Bayer I., Brooks C., Little R.R., Rocchio R.A., Shehata S., & Whitmer J. (2015). The predictive Learning Analytics revolution: Leveraging learning data for student success. *Technical report, EDUCAUSE Center for Analysis and Research*.
- Arnold, K. E. (2010). Signals: Applying academic analytics. *EDUCAUSE Quarterly*, 33(1). Disponibile in: <https://er.educause.edu/articles/2010/3/signals-applying-academic-analytic> [03 luglio 2019].
- Baldassarre, M. (2016). Think big: learning contexts, algorithms and data science. *REM – Research on Education and Media*, 8 (2), pp. 69-83.
- Bloxham S., & Boyd P. (2012). Accountability in grading student work: Securing academic standards in a twenty-first century quality assurance context. *British Educational Research Journal - BR EDUC RES J.* 38, pp. 615-634.
- Breslow L., Pritchard D.E., DeBoer J., Stump G.S., Ho A.D., & Seaton D.T. (2013). Studying learning in the worldwide classroom research into edX's first MOOC. *Research & Practice in Assessment*, 8, pp. 13-25.
- Brooks C., & Thompson C. (2017). Predictive Modelling in Teaching and Learning. In C. Lang, G. Siemens, A. Wise, & D. Gasevic (Eds.), *Handbook of Learning Analytics* (pp. 61-68). Solar.
- Campbell, J.P., & Oblinger, D. (2007). Academic Analytics. *Educause*, pp. 1-20.
- Daniel B.K. (Ed.) (2016). *Big data and learning analytics in higher education: Current theory and practice*. Switzerland: Springer
- Dawson S. (2008). A study of the relationship between student social networks and sense of community. *Educational Technology and Society*, 11(3), pp. 224-238.
- Dawson S., Bkharria A., & Heathcote E. (2010). SNAPP: Realising the affordances of real-time SNA within networked learning environments. In L. Dirckinck-Holmfeld, V. Hodgson, C. Jones, M. de Laat, D. McConnell, & T. Ryberg (Eds.), *Proceedings of the 7th International Conference on Networked Learning* (pp. 125-133). Lancaster, UK: University of Lancaster.
- Dawson S., Bkharria A., Lockyer L., & Heathcote E. (2011). "Seeing" networks: Visualising and evaluating student learning networks. Final Report 2011. *Canberra: Australian Learning and Teaching Council*. Disponibile da: <http://research.uow.edu.au/content/groups/public/@web/@learnnet/documents/doc/uow115678.pdf> [03 luglio 2019]



- De Boer H., Donker A., & Van Der Werf M. (2014). Effects of the attributes of educational interventions on students' academic performance: A meta-analysis. *Review of Educational Research*, 84(4), pp. 509-545.
- Drachler H., & Greller W. (2012). The pulse of Learning Analytics understandings and expectations from the stakeholders. *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge (LAK '12)*, 29 April–2 May 2012, Vancouver, BC, Canada, pp. 120-129.
- Ferguson R. (2012). Learning Analytics: Drivers, developments and challenges. *International Journal of Technology Enhanced Learning*, 4(5/6), pp. 304-3017.
- Fritz J. (2011). Classroom walls that talk: Using online course activity data of successful students to raise self-awareness of underperforming peers. *Internet and Higher Education*, 14(2), pp. 89-97.
- Kop R. (2011). The challenges to connectivist learning on open online networks: Learning experiences during a massive open online course. *International Review of Research on Open and Distance*, 12(3), pp. 19-38.
- Laurillard D. (2012). *Teaching as a Design Science. Building Pedagogical Patterns for Learning and Technology*. New York: Routledge.
- Littlejohn A., & Milligan C. (2015). Designing MOOCs for professional learners: Tools and patterns to encourage self-regulated learning. *eLearning Papers*, 42, pp. 38-45.
- Lockyer L., Heathcote E., & Dawson S. (2013). Informing pedagogical action: Aligning Learning Analytics with learning design. *American Behavioral Scientist*, 57(10), 1439–1459.
- Milligan S., & Griffin P. (2016). Understanding Learning and Learning Design in MOOCs: A Measurement-Based Interpretation. *Journal of Learning Analytics*, 3(2), pp. 88-115.
- Mitchell J., & Costello S. (2000). *International e-VET Market Research Report: A Report on International Market Research for Australian VET Online Products and Services*. Sydney, Australia: John Mitchell & Associates and Education Image.
- Pardo A., & Kloos C.D. (2012). *Stepping out of the box. Towards analytics outside the learning management system*. Paper presented at the 2nd International Conference on Learning Analytics and Knowledge, Vancouver, Canada.
- Rezzani A. (2015). Dalla predictive analytics alla "prescriptive analytics". Disponibile in: <https://www.dataskills.it/dalla-predictive-analytics-alla-prescriptive-analytics/> [13 luglio 2014].
- Romero C., & Ventura S. (2007). Educational data mining: A survey from 1995 to 2005. *Expert Systems with Applications: An International Journal*, 33(1), pp. 135-146.
- Shmueli, G. (2010). To explain or to predict? *Statistical Science*, 25(3), 289–310.
- Veletsianos G. (2013). How Do Learners Experience Open Online Learning? Hybrid Pedagogy. Disponibile in: <http://hybridpedagogy.org/how-do-learners-experience-open-online-learning> [15 maggio 2019].
- Zaïane O. (2001). Web usage mining for a better web-based learning environment. *Proceedings of the 4th IASTED International Conference on Advanced Technology for Education, CATE'01*, pp. 60-64.

Learning Analytics per la formazione a distanza in contesti di Lifewide Learning. Una proposta di analisi della reportistica di un corso e-learning per studenti universitari sudanesi

Learning Analytics for distance learning in Lifewide Learning contexts. A proposal to analyze the reporting of an e-learning course for Sudanese university students

Alberto Fornasari

Department of Education, Psychology, Communication, University of Study of Bari Aldo Moro, alberto.fornasari@uniba.it

Learning analytics is a kaleidoscopic research area with still blurred contours. For a definition of this term it is possible to refer to the call for papers of the First International Conference of Learning Analytics and Knowledge (LAK, 2011) which defines this scientific domain as “the measurement, collection, analysis and reporting of data about learners and their contexts , for purposes of understanding and optimizing learning and the environment in which it occurs” (wikicfp.com). This field of research aims to build analysis models that can make the data produced by a Learning Management System (LMS) sources of information, knowledge and wisdom (Ackoff, 1989) in order to improve the educational offer, the learning environments and learning outcomes of online courses. In the field of higher education, Academic Analytics aims to develop data analysis models to improve the educational offer and the design of online courses for adult users (Laurillard, 2014). This paper proposes a model of analysis of the reports of a Moodle course for Sudanese students activated by the University of Bari in the framework of the INSO project - innovation in society: training paths and human capital enhancement in Sudan. The aim is to draw from the data useful knowledge for an effective design of online learning environments in lifewide learning contexts (Cambridge, 2008).

Keywords: Learning Analytics; Academic Analytics; Lifewide Learning; Moodle; e-Learning

Il *learning analytics* è un ambito di ricerca caleidoscopico dai contorni ancora sfumati. Per una definizione di questo sintagma è possibile fare riferimento alla call for paper della *First International Conference of Learning Analytics and Knowledge* (LAK, 2011) che definisce questo dominio scientifico come “*the measurement, collection, analysis and reporting of data about learners and their contexts, for purposes of understanding and optimizing learning and the environments in which it occurs*” (wikicfp.com). Questo campo d’indagine mira a costruire modelli di analisi che possano rendere i dati prodotti da un *learning management system* (LMS) fonti di informazione, conoscenza e saggezza (Ackoff, 1989) al fine di migliorare l’offerta formativa, gli ambienti di apprendimento e i *learning outcomes* di corsi online. Nel campo dell’*higher education* l’*Academic Analytics* si propone di sviluppare modelli di analisi dei dati con lo scopo di migliorare l’offerta formativa e la progettazione dei corsi online a una specifica utenza adulta (Laurillard, 2014). Il presente contributo propone un modello di analisi dei report di un corso e-learning Moodle per studenti sudanesi attivato dall’Università di Bari nell’ambito del progetto INSO – *innovation in society: training paths and human capital enhancement in Sudan*. Il fine è quello di trarre dai dati conoscenze utili a un’efficace progettazione di ambienti di apprendimento online in contesti di lifewide learning (Cambridge, 2008).

Parole chiave: Learning Analytics; Academic Analytics; Lifewide Learning; Moodle; e-Learning

1. Introduzione



Il presente contributo propone un modello di analisi dei report di un corso e-learning Moodle per studenti sudanesi attivato dall'Università di Bari nell'ambito del progetto INSO – *innovation in society: training paths and human capital enhancement in Sudan* (finanziato su bando del Ministero dell'Interno, Dipartimento per le Libertà Civili e l'Immigrazione all'interno del programma “*Regional Development and Protection Programme for North Africa*”). Il progetto di cooperazione internazionale con il Sudan, della durata di due anni, ha previsto l'attivazione di percorsi formativi realizzati al fine di valorizzare le risorse umane garantendo a 140 giovani studenti universitari sudanesi (più precisamente provenienti da diverse nazioni africane iscritti al Comboni College of Science and Technology di Kartoum) un pieno inserimento nel proprio contesto lavorativo attraverso la costruzione di competenze tecniche specifiche. Obiettivi specifici sono stati quelli di: fornire ai destinatari, giovani sudanesi e rifugiati, il know how inerente i settori del *Computer Science* e dell'*Information Technology*; contribuire ad affiancare i professori universitari nell'ambito della loro azione di orientamento professionale ed i funzionari del Ministero dell'Educazione Superiore e della Ricerca Scientifica del Sudan al fine di creare stabili relazioni tra le università ed i contesti amministrativi ed imprenditoriali locali; promuovere esperienze concrete nel Paese mirate all'inserimento lavorativo degli studenti, attraverso il sostegno delle autorità accademiche. Si è inoltre sostenuta la pratica professionale dei giovani universitari sudanesi e dei rifugiati, sulla base di apposite convenzioni stipulate con le aziende sudanesi, gli uffici governativi e pubblici e le ONG.

Le azioni formative attivate sono state progettate quindi a partire dalle competenze interculturali dei partecipanti, evitando un approccio etnocentrico nella progettazione ed erogazione dei contenuti. La progettualità di un intervento interculturale deve prevedere infatti un impianto pedagogico che costituisca lo sfondo integratore dei diversi percorsi di formazione che favorisca negli studenti partecipanti al progetto oltre alle acquisizioni delle competenze specifiche inerenti le finalità dell'intervento stesso, anche il passaggio da un pensiero autocentrato, a un pensiero idoneo a intrecciare dialetticamente vicino e lontano, particolarità e universalità, per difendere l'identità e l'autonomia intellettuale contrastando la dipendenza e l'omologazione (Pinto Minerva, 2004). Un pensiero, come sostenuto da Sirna Terranova (2003), capace di decentrarsi, allontanarsi dai propri riferimenti cognitivi e valoriali, dirigersi verso quelli di altre culture, per scoprire e comprenderne le differenze e le connessioni; capace, inoltre, di tornare

alla propria cultura arricchito dall'esperienza del confronto e, pertanto, in grado di valutare con maggiore consapevolezza critica la propria specificità.

2. Learning Analytics, lifelong e lifewide learning nei contesti formali di istruzione

2.1 *Lifelong e Lifewide learning*

La riflessione attorno al concetto di Lifelong Learning prende avvio a partire dagli anni '70 fondando il suo significato in due aspetti centrali: l'emancipazione sociale e l'adattamento alle richieste del mercato del lavoro. Sul tema, infatti, Oliva sottolinea che "elevare il livello di aspirazione dei singoli è un impegno che deve vedere coinvolti gli individui, le famiglie, le istituzioni educative, i media, il mondo del lavoro e i decisori pubblici" (2010, p. 14); è altresì rilevante, considerando il fenomeno come parte di un cambiamento sistemico intervenuto a livello sociale, notare che, in un contesto socio-economico improntato alla precarietà, "le traiettorie di vita delle persone sono divenute discontinue e non lineari, i percorsi molto meno prevedibili" (Baldassarre, & Tamborra, 2016, p. 101), mettendo in crisi un sistema educativo fondato sulla metafora del tradizionale "bagaglio culturale" che doveva essere acquisito sostanzialmente negli anni della formazione iniziale, in preparazione alla vita adulta. Piuttosto, il sistema educativo, oggi, guarda l'allievo come un soggetto che potenzialmente resterà in formazione durante tutto l'arco della sua vita. Tra il 1994 e il 2006 la Commissione europea, accogliendo le istanze sociali emergenti cui si è appena fatto cenno, ha definito un quadro concettuale di Lifelong Learning. Il cambiamento di prospettiva sull'educazione, edotto dalla riflessione attorno al Lifelong Learning, ha sollecitato la riflessione pedagogica a definire la forma dell'educazione degli adulti: "progettare interventi educativi e formativi rivolti a soggetti adulti significa innanzi tutto riconoscere che il loro apprendimento è influenzato da un importante bagaglio di esperienze oltre che da aspettative e bisogni educativi molto specifici che li spingono a mettersi in gioco quando entrano in aula" (Limone, 2012, p. VII). In altre parole, implica, ripensare la progettazione didattica in funzione di un'utenza dalle mutate esigenze e che avanza l'istanza di un'offerta formativa flessibile e adattiva. In risposta a queste esigenze le tecnologie digitali hanno mediato il dialogo tra istituzione e fruitori aprendo nuovi scenari per la formazione superiore attraverso il supporto di infrastrutture adeguate per l'erogazione della formazione



a distanza (Mouzakittis, & Tuncay, 2011). Inizialmente “la rete era di supporto all’attività in presenza, oggi lo spazio virtuale ha una sua consistenza, diviene un modo in cui fare esperienze e relazionarsi. In cui creare comunità. [...] l’ambiente on-line [...] diventa il luogo privilegiato dove sperimentare la possibilità di comunicare in rete, di creare una comunità” (Rossi, 2012, p. 43). La progettazione degli ambienti di apprendimento online è oggetto di riflessione e miglioramento continui, cui il Learning Analytics sta offrendo nuovi modelli individuando tecniche di analisi e modelli interpretativi dell’enorme mole di dati prodotta dai LMS che consentono di descrivere le prassi di interazione degli utenti al loro interno e, dunque, di sviluppare modelli che possano prevedere e determinare scenari auspicati.



2.2 Learning Analytics

Il Learning Analytics è un ambito di ricerca caleidoscopico dai contorni ancora sfumati. Per una definizione di questo sintagma è possibile fare riferimento alla call for paper della *First International Conference of Learning Analytics and Knowledge* (LAK, 2011) che definisce questo dominio scientifico come “*the measurement, collection, analysis and reporting of data about learners and their contexts, for purposes of understanding and optimizing learning and the environments in which it occurs*” (wikicfp.com).

Questo campo d’indagine mira a costruire modelli di analisi che possano rendere i dati prodotti da un *Learning Management System* (LMS) fonti di informazione, conoscenza e saggezza (Ackoff, 1989) al fine di migliorare l’offerta formativa, gli ambienti di apprendimento e i *learning outcomes* di corsi online. Si tratta, dunque, di svolgere analisi su dati digitali per generare una conoscenza che possa tradursi in azioni mirate al miglioramento dell’apprendimento (Crick, 2017). Nell’ambito dell’*higher education* l’*Academic Analytics* si propone di sviluppare modelli di analisi dei dati con lo scopo di migliorare l’offerta formativa e la progettazione dei corsi online a una specifica utenza adulta (Laurillard, 2014). Le politiche educative e la letteratura sul tema convergono nel sottolineare l’importanza dei dati e la loro analisi per il miglioramento delle pratiche di insegnamento e apprendimento (Daniel, 2015; Siemens, Dawson, & Lynch, 2013). Colvin *et al.* (2015) hanno sviluppato un modello di “*Strategic Capability*” che descrive il modo in cui le strategie di Learning Analytics possono costituire una risposta situata, multidimensionale e dinamica ad aspetti di natura diversa tra loro correlati, come mostrato nell’immagine seguente che de-

scrive un sistema dinamico di sei componenti fondamentali interagenti con la ricerca nel campo del Learning Analytics: il contesto, la concettualizzazione delle strategie di ricerca di learning analytics, gli stakeholders, la dotazione tecnologica, la strategia istituzionale, la leadership. Gli esiti della ricerca, di rimando, agiscono su queste componenti in una relazione dinamica e ricorsiva. Il modello, inoltre, richiama l'attenzione alla dimensione del tempo che influenza a livello collettivo e individuale le componenti in termini di sviluppo, maturità ed evoluzione definendo un effetto di rotazione delle componenti rispetto al focus della ricerca.

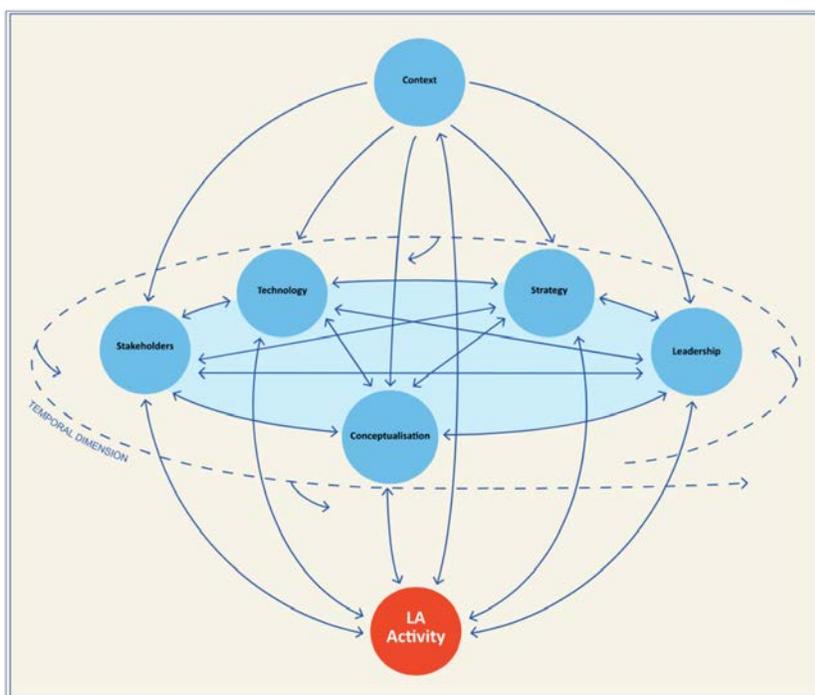


Fig.1: Modello di Strategic Capability (Colvin et al., 2015)

Per quanto pertiene, in modo particolare, della relazione intercorrente tra la ricerca nel campo del Learning Analytics e le infrastrutture tecnologiche per l'erogazione di corsi in e-learning, Crick (2017) evidenzia che il Learning Analytics indirizza scelte tecniche, politiche, pedagogiche e teoriche su aspetti legati al *Learning Design* con lo scopo di migliorare o trasformare gli *outcomes* di questi corsi. Gli esiti di queste ricerche possono indirizzare la progettazione didattica verso scelte di due tipi:

1. Migliorare le potenzialità di apprendimento individuali attraverso l'implementazione di strategie volte ad agevolare l'auto-regolazione del flusso di informazioni;
2. Rispondere in modo più accurato ai bisogni di apprendimento dell'utenza (Crick, 2017).

3. Il contesto di analisi

L'ambiente di apprendimento on-line predisposto per l'erogazione della formazione a distanza è stato ospitato all'interno di una piattaforma di e-learning Moodle (*Modular Object-Oriented Dynamic Learning Environment*). Moodle è un ambiente informatico per la gestione di corsi basato sull'ideologia costruzionista secondo la quale ogni apprendimento sarebbe facilitato dalla produzione di oggetti intangibili.

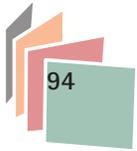
L'ambiente di apprendimento che ha ospitato le attività formative del progetto è ospitato sul sito insoelearning.it. La progettazione dell'ambiente è stata fondata sul concetto di presenza sociale sostenendo la comunità attraverso la coesione, necessaria affinché gli studenti si sentano coinvolti nel processo di apprendimento e per ridurre la percezione di isolamento e il conseguente rischio di inefficacia dell'intervento che può verificarsi nei contesti di formazione a distanza.

Tra le risorse messe a disposizione dalla piattaforma Moodle, il corso INSO ha usufruito di:

- video-lezioni, materiale di approfondimento in PDF e mappe concettuali per veicolare le conoscenze in un'ottica di auto-formazione;
- compiti e quiz per la valutazione degli apprendimenti e delle competenze acquisite;
- forum e chat per facilitare il confronto tra gli studenti e la costruzione di una comunità.

L'offerta formativa è stata elaborata in 6 moduli:

- *Building micro-enterprises*. Contentente 19 videolezioni con relativi materiali PDF di approfondimento;
- *Organizations' Sostenibility*. Contentente 14 videolezioni con relativi materiali PDF di approfondimento;
- *Marketing & Management*. Contentente 11 videolezioni con relativi materiali PDF di approfondimento;
- *Social networks and relationship dynamics*. Contentente 9 videolezioni con relativi materiali PDF di approfondimento;



- *Job search techniques*. Contentente 12 videolezioni con relativi materiali PDF di approfondimento;
- *Social media for micro-enterprises*. Contentente 8 videolezioni con relativi materiali PDF di approfondimenti.

Il progetto prevedeva il coinvolgimento di 140 studenti del *Comboni College of Science & Technology of Khartoum*; di costoro 47 non erano in possesso di un indirizzo e-mail personale. Per tale ragione gli studenti coinvolti sono stati 103 (il 68,67% degli studenti previsti). Di costoro, 16 studenti non hanno mai fatto accesso in piattaforma e non hanno fruito della formazione erogata. Pertanto, le attività formative sono state portate avanti dall'84,47% degli studenti idonei alla frequenza, ossia da 87 studenti dei 103 registrati in piattaforma. A richiamo dell'impianto pedagogico interculturale prima citato, il Comboni College (gestito da padri della congregazione comboniana e quindi cattolici) tenendo conto dell'altissimo numero di studenti di religione islamica presenti (oltre il 90%) e tenendo conto dei precetti religiosi che guidano la quotidianità di tali studenti, ha organizzato un calendario delle lezioni (seguite presso i laboratori di informatica) e un servizio mensa che ha accolto tali specificità, attivando tra studenti di religione cattolica e studenti di religione islamica un percorso di conoscenza e rispetto delle reciproche tradizioni e *weltanschauung*.

I dati per le analisi sono stati raccolti mediante il plugin di Moodle "Analytics Graphs".



4. Analisi dei dati

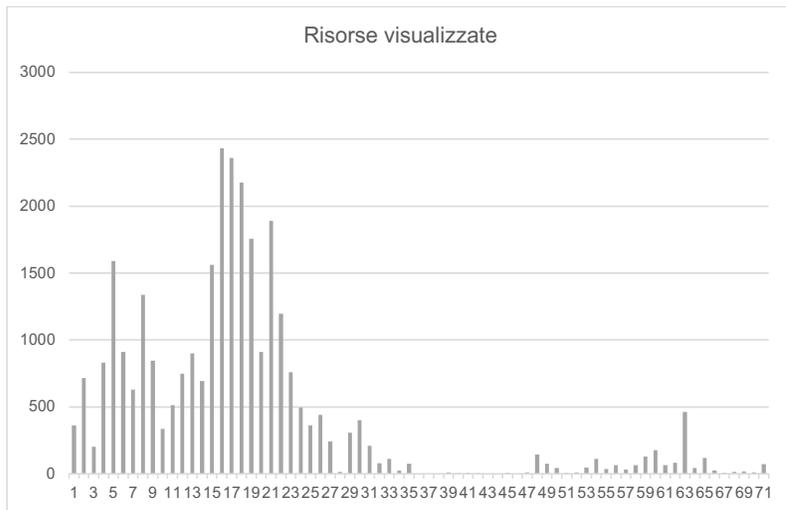
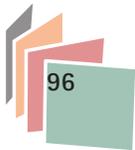
L'analisi della partecipazione degli studenti alle attività del corso prende in esame alcuni aspetti salienti quali:

- L'andamento, nel corso delle 71 settimane, delle risorse visualizzate in piattaforma;
- L'andamento dell'accesso alle diverse risorse didattiche disponibili: video-lezioni e materiale di approfondimento in PDF;
- L'andamento, nel corso dei 17 mesi, delle interazioni nei forum;
- L'andamento dello svolgimento delle prove di valutazione finale;
- L'analisi di questi dati mira a identificare quegli aspetti del corso che sono stati progettati efficacemente e a fornire indicazioni utili per progettazioni future.

4.1 *Andamento, nel corso delle 71 settimane, delle risorse visualizzate in piattaforma*

Analizzare l'andamento nel tempo delle risorse visualizzate in piattaforma consente di avere una prima panoramica generale del grado di coinvolgimento degli studenti nelle attività didattiche per tutta la durata del corso. La presenza online degli studenti costante per tutta la durata del corso sarebbe indice di un'efficace progettazione delle attività didattiche soprattutto nella pianificazione della durata del corso, nonché nella gestione dei materiali didattici resi disponibili e delle attività richieste online.

Nel grafico seguente viene mostrato il numero totale di accessi avvenuto settimanalmente dall'inizio del corso, il 20 ottobre 2017, alla sua conclusione, il 28 febbraio 2019, per un totale di 71 settimane.



Graf. 1: numero di accessi settimanali al corso

Dalla lettura dei dati si riscontra, in generale, che l'accesso alle risorse online è stato più intenso nelle prime 30 settimane del corso, ossia nei primi 6 mesi. In questo arco temporale si sono registrate 27904 visualizzazioni su un totale di 30297, ossia il 92,1% del totale.

Entrando nel dettaglio, si evidenzia che le maggiori attività si sono concentrate tra la 15^a e la 22^a settimana, ossia tra l'ultima settimana di gennaio 2018 e la terza settimana di marzo 2018, in un arco temporale pari a 2 mesi, in cui si sono registrate 14288 visualizzazioni, pari al 47,16% del totale.

Questo arco temporale corrisponde alla fase in cui tutti i materiali didattici del corso sono stati pubblicati e resi disponibili online. Conseguentemente i corsisti in questo periodo sono stati particolarmente attivi nello studio delle risorse. Superata la 30^a settimana, la presenza degli studenti è gradualmente scemata.

Tra dicembre 2018 e gennaio 2019 si è registrato un lieve incremento nelle visualizzazioni corrispondente alla fase di valutazione finale del corso.

4.2 *L'andamento dell'accesso alle diverse risorse didattiche disponibili: video-lezioni, materiale di approfondimento in PDF e mappe concettuali*

Ulteriore polarità d'interesse è stata quella dell'analisi approfondite degli accessi ai materiali didattici. Perché fonte di informazione su una ampia gamma di aspetti legati alla progettazione didattica: dalla qualità dei materiali stessi, alle preferenze di studio degli utenti in merito alla tipologia del materiale, nonché l'interesse nei confronti dei diversi moduli del corso.

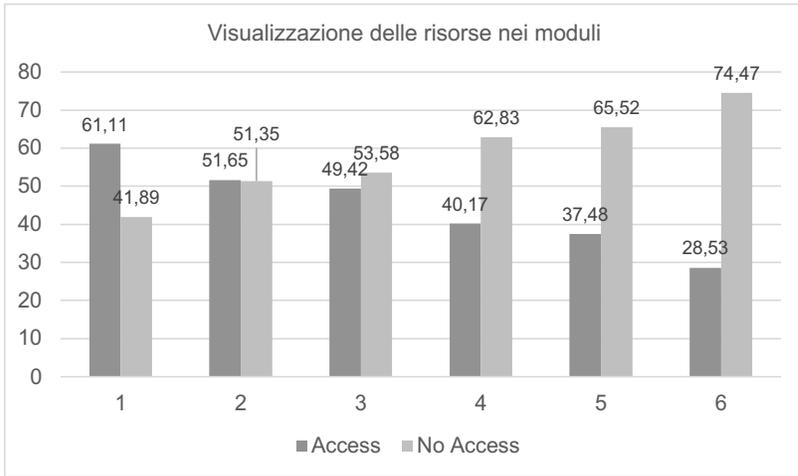
Per questi scopi l'analisi dell'accesso ai materiali didattici verrà di seguito discussa su 2 differenti piani di lettura: il confronto negli accessi tra i diversi moduli formativi e il confronto negli accessi tra le diverse tipologie di materiale.

L'accesso alle risorse didattiche: confronto tra moduli

L'accesso alle risorse didattiche è stato analizzato a partire dai dati relativi al numero di studenti che hanno o non hanno visualizzato le singole risorse caricate online.

Il grafico e la tabella seguenti mostra il numero e la percentuale di studenti che hanno visualizzato le risorse di ciascun modulo e quello degli studenti che non le hanno visualizzate. I dati sono riportati in media rispetto al numero complessivo di risorse caricate all'interno di ciascun modulo.





Graf. 2: Medie degli studenti che hanno (e non hanno) visualizzato le risorse dei 6 moduli del corso



Modulo	Accessi (%)	Non Accessi (%)
1	59,33%	40,67%
2	50,15%	49,85%
3	47,98%	52,02%
4	39%	61%
5	36,39%	63,61%
6	27,7%	72,3%

Tab. 1: Percentuale di studenti che hanno visualizzato le risorse dei moduli

Dalla lettura dei dati si rileva che l'accesso da parte degli studenti ai materiali di studio è stato più intenso nei primi moduli ed è diminuito gradualmente fino all'ultimo modulo in cui si registra che in media solo il 27,7% degli studenti ha visualizzato i materiali didattici. Il primo modulo formativo ha registrato, invece, una percentuale media di studenti che hanno visualizzato le risorse pari al 59,33%.

Poiché i materiali didattici dei 6 moduli erano tutti disponibili online nei periodi di maggiore presenza degli studenti, la ragione di questo decremento nell'accesso alle risorse dei moduli è da ricercarsi in altri fattori. Sono state avanzate due possibili ipotesi interpretative, una relativa ai contenuti dei moduli, l'altra di tipo prettamente progettistico:

- I. Ragioni contenutistiche: i primi tre moduli, che hanno registrato maggiore intensità d'accesso alle risorse, hanno incontrato maggior-

mente i fabbisogni formativi degli utenti, rispetto agli ultimi tre. L'offerta formativa del corso, infatti, potrebbe considerarsi come divisa in due macro-aree tematiche: marketing e management delle micro-imprese (primi 3 moduli); social media per le micro-imprese e dinamiche relazionali (ultimi 3 moduli). Si potrebbe dire, dunque, che l'interesse degli utenti fosse maggiormente orientato verso una formazione sui temi di marketing e management delle imprese. Considerando che gli utenti erano studenti dei corsi di laurea in computer science e information technology presso il Comboni College in Sudan, appare coerente ritenere che l'interesse degli studenti sia stato maggiormente catalizzato da quei contenuti maggiormente lontani dalle loro pre-conoscenze;

II. Ragioni progettistiche: da un punto di vista progettistico i moduli del corso sono stati caricati in piattaforma in modo sequenziale, uno sotto l'altro. I primi moduli, inoltre, contenevano un maggior numero di video-lezioni e dispense di approfondimento rispetto agli ultimi; le video-lezioni, inoltre, avevano anche una durata media maggiore. Si può ipotizzare, dunque, un calo d'interesse progressivo causato dalla percezione di irraggiungibilità dell'obiettivo di completare il corso.

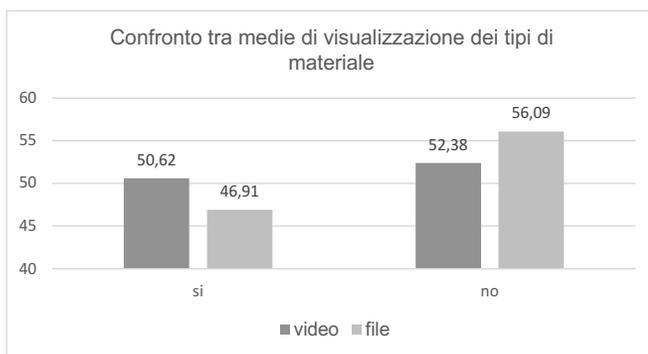


L'accesso alle risorse didattiche: confronto tra tipologie di materiali

L'analisi dei dati di accesso alle risorse didattiche con specifico focus sul confronto tra le diverse tipologie di risorse disponibili persegue l'obiettivo di individuare le modalità di studio privilegiate dagli studenti nei contesti di autoformazione al fine di migliorare la progettazione didattica dei corsi di formazione a distanza potenziando l'offerta formativa erogata mediante quei tipi di materiali didattici che l'utenza ritiene più efficace per lo studio.

All'interno del corso, il materiale didattico erogato è stato prodotto in due essenziali formati: video-lezioni esplicative dei contenuti specifici dei singoli moduli; dispense e slide in pdf di supporto allo studio formale e autonomo.

Per una prima lettura generale dei dati di accesso ai materiali didattici finalizzato a operare un confronto tra le diverse tipologie di risorse, nel grafico e nella tabella seguenti vengono mostrate le medie e le percentuali degli utenti che hanno visualizzato i materiali didattici considerati per tipologia, raffrontate a coloro che non vi hanno fatto accesso.



Graf. 3: Confronto tra medie di visualizzazione dei diversi tipi di materiale didattico

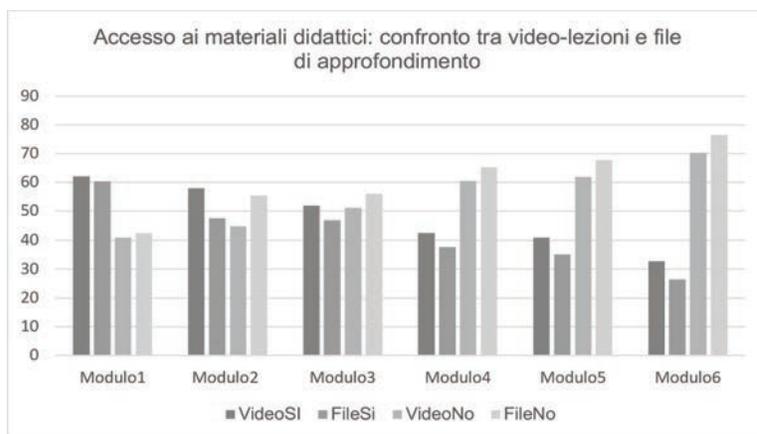


Visualizzazione	Video (%)	File (%)
si	49,15	45,54
no	50,85	54,46

Tab. 2: confronto tra le percentuali di visualizzazione dei diversi tipi di materiale didattico

Dalla lettura dei dati mostrati, si riscontra che, in linea generale, gli utenti hanno preferito studiare i contenuti del corso attraverso la visualizzazione delle video-lezioni più che attraverso le dispense testuali.

Considerate le difformità generali nell'accesso ai materiali didattici dei diversi moduli del corso, pare opportuno analizzare nel dettaglio le differenze di accesso ai diversi tipi di materiali all'interno dei moduli. Nel grafico e nella tabella seguenti, dunque, vengono mostrate le medie e le percentuali degli utenti che hanno visualizzato e quelle di coloro che non hanno visualizzato le video-lezioni e i file di approfondimento all'interno dei 6 moduli del corso.



Graf. 4: Medie degli utenti che hanno (e non hanno) visualizzato i materiali del corso

Modulo	Accessi Video (%)	Accessi File (%)	Non Accessi Video (%)	Non Accessi File (%)
1	60,24	58,65	39,76	41,35
2	56,31	46,19	43,69	53,81
3	50,40	45,63	49,60	54,37
4	41,32	36,65	58,68	63,35
5	39,81	34,15	60,19	65,85
6	31,80	25,73	68,20	74,27

Tab. 3: Percentuali degli utenti che hanno (e non hanno) visualizzato i materiali del corso all'interno dei moduli: confronto tra video-lezioni e file di approfondimento

Guardando nel dettaglio i dati di accesso ai materiali all'interno dei moduli del corso, si può, innanzi tutto, confermare la tendenza individuata nell'analisi generale dei dati: ossia che gli utenti hanno sempre preferito studiare i contenuti del corso mediante video-lezioni più che attraverso le dispense testuali. In alcuni moduli didattici la differenza tra chi ha preferito le video-lezioni e chi le dispense testuali è più accentuata rispetto ad altri. In particolare, si nota che nel modulo 1 la differenza nella visualizzazione di video-lezioni e dispense è minima, suggerendo che il primo modulo didattico si è configurato come un momento di socializzazione per gli utenti a questa modalità di formazione. Si presume che all'inizio del corso essi abbiano esplorato le possibilità offerte dalla piattaforma per poi scegliere in autonomia la propria modalità di studio più efficace.



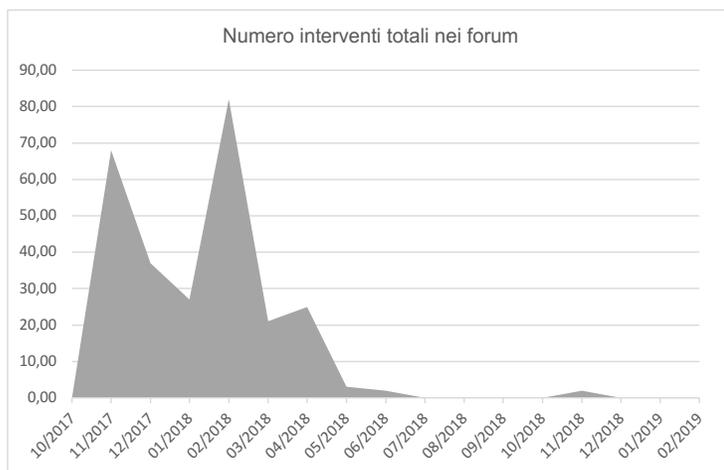
4.3 L'andamento, nel corso dei 17 mesi, delle interazioni nei forum

Per analizzare le interazioni nei forum si è scelto di leggere i dati nel loro andamento nel tempo lungo la durata del corso. Questa chiave di lettura consente di individuare la funzione che ha assunto il forum all'interno del corso, nonché avere un ulteriore elemento di dettaglio in merito all'efficacia della sua progettazione.

I dati verranno dapprima presentati in modo globale al fine di presentare l'andamento generale delle interazioni all'interno di tutti i forum per tutta la durata del corso al fine di cogliere una overview globale sull'andamento del corso. In seguito, si entrerà nel dettaglio dell'andamento delle interazioni nei forum dei singoli moduli al fine di comprendere più analiticamente la natura dei dati presentati.

Il grafico e la tabella seguenti mostrano il numero e la percentuale

di interventi totali che sono stati inseriti all'interno dei forum nell'arco dei 17 mesi del corso.



Graf. 5: numero totale degli interventi nei forum nell'arco dei 17 mesi del corso (punteggi grezzi)

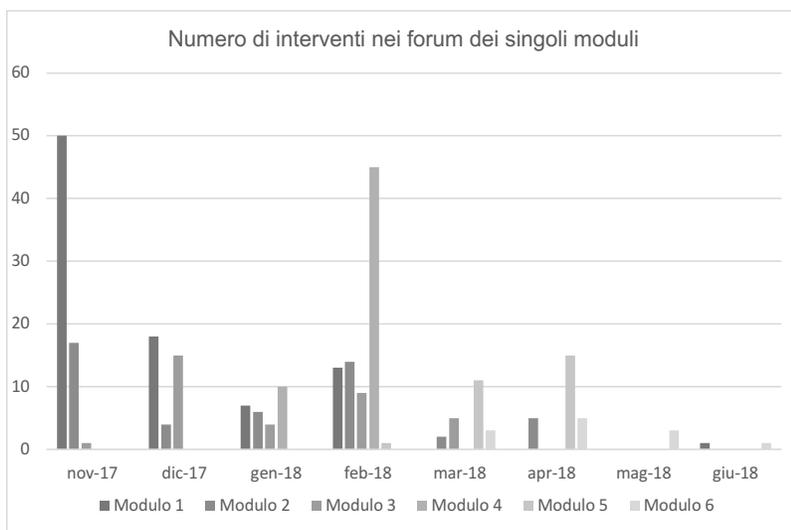
Mese	Percentuale degli interventi
Ottobre 2017	0%
Novembre 2017	25,47%
Dicembre 2017	13,86%
Gennaio 2018	10,11%
Febbraio 2018	30,71%
Marzo 2018	7,87%
Aprile 2018	9,36%
Maggio 2018	1,12%
Giugno 2018	0,75%
Luglio 2018	0%
Agosto 2018	0%
Settembre 2018	0%
Ottobre 2018	0%
Novembre 2018	0,75%
Dicembre 2018	0%
Gennaio 2019	0%
Febbraio 2019	0%

Tab. 4: Percentuali degli interventi nei forum nell'arco dei 17 mesi del corso

Dalla lettura dei dati esposti si riscontra che gli studenti hanno prevalentemente utilizzato i forum nei primi 6 mesi del corso, arco temporale in cui si è concentrato il 97,38% di tutti gli interventi inseriti nei forum per tutta la durata del corso. Questo dato suggerisce che i forum sono stati utilizzati dagli studenti prevalentemente nei primi tempi del corso, ossia nelle fasi di approccio ai contenuti formativi come strumenti di confronto e supporto allo studio iniziale.

Questi dati, tuttavia, non sono sufficienti a confermare questa ipotesi dal momento che questi tempi coincidono con il periodo di maggiore attività generale all'interno del corso, a seguito del quale la presenza generale degli utenti on-line è calata significativamente. Si rende, pertanto, necessario analizzare nel dettaglio l'andamento nel tempo dei forum nei singoli moduli.

Il grafico e la tabella seguenti mostrano il numero e le percentuali degli interventi inseriti nei forum dei 6 moduli nel corso dei mesi. Per ragioni di leggibilità del grafico, dal dataset sono stati eliminati i dati di ottobre 2017, in cui non ci sono stati interventi, e quelli successivi a Giugno 2018 dal momento che sono tutti pari a zero tranne che per 2 interventi inseriti nel forum del modulo 1 a novembre 2018.



Graf. 6: numero di interventi nei forum dei singoli moduli distribuiti nei mesi di attività del corso: novembre e dicembre 2017, gennaio, febbraio, marzo, aprile, maggio e giugno 2018 (punteggi grezzi)

Moduli								
6	0	0	0	0	25,00%	41,67%	25,00	8,33%
5	0	0	0	3,70%	40,74%	55,56%	0	0
4	0	0	18,18%	81,82%	0	0	0	0
3	2,94%	44,12%	11,76%	26,47%	14,71%	0	0	0
2	35,42%	8,33%	12,50%	29,17%	4,17%	10,42%	0	0
1	54,95%	19,78%	7,69%	14,29%	0	0	0	1,10%
Mesi	nov-17	dic-17	gen-18	feb-18	mar-18	apr-18	mag-18	giu-18

Tab. 5: percentuali degli interventi nei forum dei singoli moduli durante gli 8 mesi di principale attività del corso



Da un'attenta analisi dei dati presentati si riscontra che le interazioni all'interno dei forum hanno seguito un andamento graduale. I post nei forum dei primi due forum si registrano a partire dal mese di novembre 2017. Nel forum del modulo 3 le interazioni sono cominciate a partire da dicembre 2017. Così, in ordine progressivo, i moduli successivi sono stati gradualmente "popolati" gradualmente uno dopo l'altro. Questi dati suggeriscono che gli studenti hanno adottato una strategia studio e accesso alla conoscenza sostanzialmente lineare. I materiali didattici sono stati visualizzati in sequenza, sicché anche l'accesso ai forum è stato sequenzialmente graduale di modulo in modulo.

Si conferma e rafforza, inoltre, l'ipotesi avanzata precedentemente per la quale la possibilità di interagire nei forum è stata utilizzata dagli studenti allo scopo di confrontarsi e supportarsi vicendevolmente nella fase di approccio ai nuovi temi di studio. Durante le video-lezioni, in accordo con l'impianto pedagogico interculturale esplicitato nell'introduzione, si è lavorato attraverso "una didattica dei punti di vista" per favorire (rispetto agli argomenti oggetto del corso) interpretazioni plurime stimolando così un confronto critico tra gli studenti teso alla formazione di un pensiero plurale, aperto. Confronto che è emerso con vivacità all'interno dei forum.

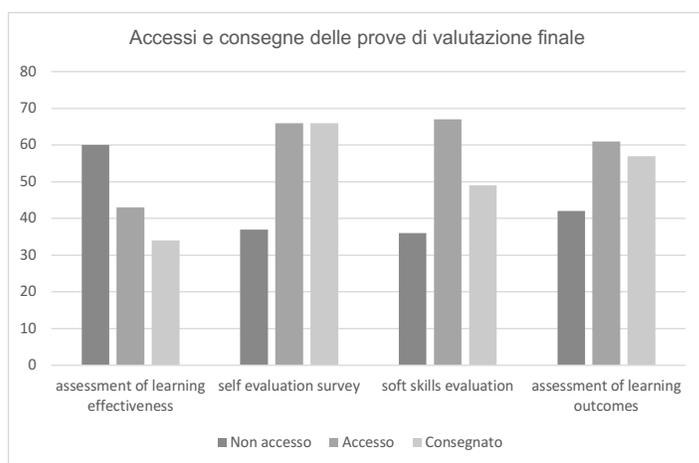
4.4 L'andamento dello svolgimento delle prove di valutazione finale

Per la valutazione finale sono stati somministrati 4 strumenti:

1. *Assessment of learning effectiveness*: somministrato mediante un link a un form di google moduli;
2. *Self evaluation survey* about career development competencies: somministrato mediante un link a un form di google moduli;

3. *Soft skills evaluation*: somministrato attraverso un assessment task in piattaforma composto da 2 domande semistrutturate;
4. *Assessment of learning outcomes*: somministrato attraverso un quiz in piattaforma.

Dai risultati ottenuti si riscontra un accesso differme da parte degli utenti: non tutti gli studenti hanno fatto tutte le prove. Nel grafico e nella tabella seguenti vengono mostrati i dati di accesso e completamento delle prove.



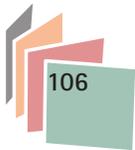
Graf. 7: Numero di studenti che hanno visualizzato, consegnato e non visualizzato le prove di valutazione finale (punteggi grezzi)

Prova	N. di studenti che hanno visualizzato la prova	N. di studenti che hanno completato la prova	% di studenti che hanno completato la prova rispetto a coloro che hanno visualizzato	% di studenti che hanno completato la prova rispetto al n. degli studenti totali
Assessment of learning effectiveness	43	34	79,07%	33,01%
Self evaluation survey	66	66	100%	64,08%
Soft skills evaluation	67	49	73,13%	47,57%
Assessment of learning outcomes	61	57	93,44%	55,34%

Tab. 6: Riepilogo del numero di studenti che hanno completato le prove di valutazione rapportato al numero di coloro che le hanno visualizzate e al numero totale di studenti del corso

Analizzando i dati presentati è possibile innanzi tutto osservare non solo una difformità nell'accesso alle prove ma anche nel completamento delle stesse. Infatti, non tutti gli utenti che hanno visualizzato le prove le hanno, poi, completate. Nonostante le prove siano state pubblicate online tutte nello stesso periodo, vi sono state delle significative differenze negli accessi. La prova che ha registrato il più basso tasso di partecipazione è *l'assessment of learning effectiveness* che non ha raggiunto nemmeno il 50% degli utenti che vi accedessero. Invece, la prova che ha registrato il più alto livello di partecipazione e il maggior numero di completamenti è la *self evaluation survey*.

5. Conclusioni



Le attività formative del progetto erogate in e-learning sono state elaborate con l'obiettivo di favorire negli studenti l'acquisizione di soft skills e conoscenze specifiche che potessero risultare utili all'inserimento nel mondo del lavoro all'interno del territorio di origine. Proprio in virtù di questo il *Comboni College of Science & Technology of Khartoum* ha dato vita ad una *start-up incubator* che ha seguito gli studenti partecipanti al corso nella realizzazione di diversi progetti imprenditoriali che sono stati poi finanziati da investitori locali. Appare doveroso rilevare come l'impianto interculturale del progetto INSO, più volte citato, lo contraddistingua dai tanti progetti di cooperazione internazionale, nati sì con l'intenzione di aiutare i Paesi in via di sviluppo spesso lacerati da guerre intestine, ma spesso connotati da prospettive culturali etnocentriche (spesso eurocentriche) che impongono valori e modelli molto distanti da quelli diffusi nelle culture locali disegnando sovente i tratti di un neo-colonialismo umanitario. I progetti finanziati sono stati infatti costruiti dal basso partendo dalle necessità espresse dal territorio.

Le analisi, proposte all'interno del contributo, svolte sull'andamento delle attività in e-learning, suggeriscono una partecipazione piuttosto autonoma e auto-organizzata da parte degli studenti alle attività formative proposte. Questo lo si evince dalle significative differenze emerse nell'accesso ai materiali didattici dei moduli del corso, nonché alle notevoli differenze emerse nel completamento delle prove di valutazione finale. Svincolati da un senso di obbligatorietà della formazione, hanno approcciato il corso piuttosto come la possibilità di colmare, integrare e aggiungere conoscenze differenti dalle proprie pregresse (Fornasari, 2016) cogliendo pienamente il principio dell'auto-formazione nei contesti di formazione superiore, in cui ciascun utente coglie, dalla

proposta formativa, non l'intera offerta didattica, ma ciò che gli è più utile rispetto ai propri obiettivi personali. È, questo principio, una delle basi della progettazione modulare dei corsi erogati in e-learning. Nel complesso per una efficace riprogettazione, si ritiene necessario ripensare la scansione temporale delle attività didattiche, nonché le modalità di presentazione dei moduli formativi, cercando di svincolarsi da una logica lineare proposta dall'alto indiscriminatamente per tutti gli utenti, ma in modo personalizzato per ciascun utente, abilitando l'avanzamento nel corso attraverso step di completamento. Questa modalità di erogazione dei contenuti, consentirebbe, inoltre, non solo di favorire agli utenti di seguire un proprio ritmo di studio, ma consentirebbe anche agli amministratori del corso di avere contezza, nel corso delle attività formative, di quali siano quegli studenti che rischiano di restare indietro nell'accesso alle attività didattiche, e, conseguentemente, di abbandono del corso. Questo permetterebbe di intervenire attraverso azioni mirate e personalizzate per questi studenti, rendendo la formazione erogata più flessibile e adattata attorno alle necessità individuali dell'utente.



Riferimenti bibliografici

- Ackoff R.L. (1989). From Data to Wisdom. *Journal of Applied Systems Analysis*, 16, pp. 3-9.
- Baldassarre M., Tamborra V. (2016). *La formazione transmediale del docente in servizio*. In L. Perla, M. Tempesta (Eds.), *Teacher Education in Puglia. Università e Scuola per lo sviluppo della professione docente* (pp. 93-126). Lecce: Pensa MultiMedia
- Cambridge D. (2008). Audience, integrity, and the living document: eFolio Minnesota and lifelong and lifewide learning with ePortfolios. *Computer & Education*, 5 (3), pp. 1227-1246.
- Colvin C., Rogers T., Wade A., Dawson S., Gasevic D., Buckingham Shum S., ..., Fisher J. (2015). *Student retention and learning analytics: A snapshot of Australian practices and a framework for advancement*. Canberra, ACT: Australian Government Office for Learning and Teaching.
- Commissione delle Comunità Europee (2000). *Memorandum sull'istruzione e la formazione permanente*, Bruxelles, 20 ottobre 2000, SEC (2000) 1832.
- Crick R. (2015). Learning Analytics: Layers, Loops and Processes in a Virtual Learning Infrastructure. In C. Lang, G. Siemens, A. Wise, & D. Gasevic (Eds.), *Handbook of Learning Analytics* (pp. 291-307). Solar.
- Daniel B. (2015). Big data and analytics in higher education: Opportunities and challenges. *British Journal of Educational Technology*, 46(5), pp. 904-920.
- Fornasari A., (2016), *Teacher Education Through Ict. Moocs: New Learning Environments*, *Tojet: the Turkish online Journal of Educational Technology*, Turkey

- Laurillard, D. (2014). *Insegnamento come scienza della progettazione. Costruire modelli pedagogici per apprendere con le tecnologie*. Milano: FrancoAngeli.
- Limone P. (2012). *Valutare l'apprendimento on-line. Esperienze di formazione continua dopo la laurea*. Bari: Progedit.
- Mouzakittis S., Tuncay N. (2011). E-Learning and Lifelong Learning. *Turkish Online Journal of Distanca Education*, 12 (1), pp- 166-173.
- Oliva A. (2010). Introduzione e guida alla lettura. *Il lifelong learning e l'educazione degli adulti in Italia e in Europa. Dati, confronti e proposte, Quaderno n. 9, dicembre 2010*, pp. 14-29.
- Pinto Minerva F. (2004). *Mediterraneo. Mare di incontri interculturali*. Bari: Progedit.
- Siemens, G., & Long, P. (2011). Penetrating the fog: Analytics in learning and education. *Educause Review*, 46(5), 30.
- Siemens G., Dawson S., & Lynch G. (2013). *Improving the quality and productivity of the higher education sector: Policy and strategy for systems-level deployment of learning analytics*. Sydney, Australia: Australian Government Office for Teaching and Learning.
- Sirna Terranova C. (2003). *Pedagogia interculturale. Concetti, problemi, proposte*. Milano: Guerini.



Dati demografici degli studenti e selezione dei MOOC su Eduopen. Uno studio esplorativo sui MOOC erogati da UniFg

Student demographics and MOOC selection on Eduopen. An exploratory study on UniFg MOOCs

Anna Dipace

Department of Education and Human Sciences, University of Modena and Reggio Emilia, anna.dipace@unimore.it

Lucia Maria Borrelli

Department of Human Study, University of Foggia, lucia.borrelli@unifg.it

Sara Perrella

Department of Human Study, University of Foggia, sara.perrella@unifg.it

Katia Caposeno

Department of Human Study, University of Foggia, katia.caposeno@unifg.it

Today's generation of Massive Open Online Courses (MOOC) based on Open Educational Resources (OER) is able to offer high quality education to all those who decide to use this new type of online lifelong learning. Students who decide to enroll in these courses represent an increasingly diverse audience in terms of age and place of origin. The literature is being enriched with research studies that study the demographic data of the students in relation to different variables, including the type of course MOOC chosen.

17 universities join the Eduopen project, including the University of Foggia, which in the first three years has contributed to the growth and development of the platform by designing and delivering various MOOCs, involving numerous teachers and stimulating an increasing number of students from non-geographical areas. necessarily surrounding. Specifically, in April 2019 UniFg counts:

- 34 Mooc Courses;
- 4 Pathway;
- 45 Lecturers and Tutors.
- more than 8000 students.

This contribution presents an exploratory study carried out on the demographic data of the students enrolled in the courses offered by the University of Foggia. Specifically, the demographic data of the students were analyzed based on age, educational qualifications and city of origin and studied the correlations between these data and the choice of available courses.

Keywords: Students demographic characteristics; MOOC; lifelong learning; eduopen

La generazione odierna di Massive Open Online Courses (MOOC) basati su Open Educational Resources (OER) è in grado di offrire un'istruzione di qualità a tutti coloro che, per varie ragioni, utilizzano questa nuova metodologia di formazione online. Gli studenti che decidono di iscriversi a questi corsi costituiscono un pubblico sempre più diversificato per quanto riguarda età e luogo di provenienza.

I ricercatori, educatori, e il pubblico in generale recentemente si è interessato molto su come differisce la provenienza dei corsisti e sulla relazione tra questa variabile e la scelta di un corso MOOC. Al progetto Eduopen aderiscono 17 Atenei, tra cui l'Università di

1 La struttura del presente contributo è stata pensata in maniera condivisa dalle autrici, mentre nella presentazione dei diversi aspetti e nella stesura dei paragrafi si è seguita la seguente suddivisione: § 1 e 4 Sara Perrella - § 2 e 2.1 Lucia Maria Borrelli - § 3 Anna Dipace - § 5 Katia Caposeno.

Foggia, che nel primo triennio ha contribuito alla crescita e sviluppo della piattaforma progettando ed erogando diversi MOOC, coinvolgendo numerosi docenti e stimolando un numero sempre più crescente di studenti provenienti da zone geografiche non necessariamente circostanti. Nello specifico, ad aprile 2019 Unifg conta:

- 34 Corsi MOOC;
- 4 Pathway;
- 45 Docenti e Tutor.
- più di 8000 studenti.

Questo contributo presenta uno studio esplorativo effettuato sui dati demografici degli studenti iscritti ai corsi offerti dall'ateneo foggiano. Nello specifico, sono stati analizzati i dati demografici degli studenti in base a età, titolo di studio e città di provenienza e studiate le correlazioni tra questi dati e la scelta dei corsi disponibili.

Parole chiave: Students demographic characteristics; MOOC; lifelong learning; eduopen



1. Introduzione

La didattica *online* è oggi una delle forme più diffuse per sviluppare percorsi di formazione; l'ultima frontiera della *Distance Education* risulta essere l'erogazione di corsi e-learning nella forma dei MOOC.

A circa 10 anni dalla loro comparsa, i MOOC (*Massive Open Online Courses*) risultano essere un fenomeno in continua espansione poiché in grado di offrire risposte a molteplici esigenze (Savelli, 2017): la loro capacità di diffusione massiva e la possibilità di fruire dei materiali indipendentemente dai vincoli spazio-temporali hanno richiamato una larga percentuale di corsisti (Calise, 2018).

La generazione attuale di MOOC basati su OER (*Open Educational Resources*) è in grado di rispondere alle esigenze formative moderne attraverso un'offerta didattica variegata e di qualità.

L'utenza dei corsi massivi è composta da un pubblico sempre più diversificato per età e luogo di provenienza. Su queste premesse risulta opportuno predisporre un percorso di apprendimento personalizzato in cui:

- gli obiettivi siano diversificati da studente a studente;
- gli studenti, nella costruzione del proprio percorso, possano sviluppare le diverse dimensioni (cognitiva, emotiva, relazionale, ecc ...) e rafforzare le conoscenze e abilità precedenti (Cordova & Lepper, 1996).

Considerando le differenze tra il paradigma dell'apprendimento tradizionale e i MOOC, sono in corso numerose ricerche volte a indagare e approfondire il fenomeno della previsione, dell'abbandono degli studenti e dei tassi di completamento nei MOOC.

Questo contributo presenta uno studio esplorativo effettuato sui dati demografici degli studenti iscritti ai corsi offerti dall'ateneo foggiano e si pone come obiettivo principale quello di leggere e analizzare le caratteristiche principali degli utenti e quindi provare a definire il profilo dello studente open per pensare ad una offerta formativa mirata e centrata sui bisogni del territorio (sebbene in una dimensione massiva).

La necessità di realizzare questo studio emerge proprio dalla letteratura di settore e più precisamente dall'ambito specifico dei Learning Analytics che hanno come focus proprio l'applicazione di modelli predittivi nei sistemi di istruzione attraverso la descrizione di dati e risultati utilizzando tecniche specifiche, quali: statistica, visualizzazione SNA, sentiment analysis, influence analytics, discourse analysis, concept analysis, and sense-making models. Attraverso i LA è possibile non solo raccogliere dati, ma anche elaborarli e analizzarli in profondità al fine di favorire la previsione di comportamenti degli studenti e quindi progettare interventi di progettazione didattica volti a migliorare i risultati di apprendimento e quindi il successo formativo (Siemens, Baker, 2012).



2. Lo studio esplorativo e lo stato dell'arte

L'introduzione delle ICT (*Information and Communications Technology*) nel campo dell'educazione ha generato un cambiamento profondo nei processi didattici all'interno dei contesti formativi. L'e-learning ha guadagnato terreno non solo per la sua efficienza, ma anche per la moltitudine di studi volti a migliorare il processo educativo (Llorente, 2006).

L'affermarsi del *lifelong learning*, ovvero la formazione intesa come processo che accompagna l'uomo durante l'intero corso della vita, ha coinvolto fasce sempre più ampie della popolazione ed ha dato spazio ad un nuovo mercato in grado di offrire contenuti specializzati che vanno ad affiancare i piani didattici di scuole e università. Questi cambiamenti profondi hanno generato un fenomeno di grandi dimensioni sia per quanto riguarda il numero di utenti coinvolti che nelle nuove figure professionali di cui questo mercato ha bisogno.

Il continuo progresso delle nuove tecnologie ha permesso la creazione di nuovi scenari basati su metodologie non più tradizionali ma caratterizzate da una forma di apprendimento collaborativo (piuttosto che individuale) e da una costruzione dei saperi anziché dalla trasmissione della conoscenza (Wenger, 1998).

L'apprendimento partecipativo si presenta come processo *learner-centered* e facilitatore delle interazioni fra gli studenti. Lo studente, che

prima rivestiva i panni di attore passivo diventa agente principale capace di determinare il proprio processo di apprendimento (Sansone, Ligorio, Dillenburg, 2012); la prospettiva “*learner centered*” si focalizza sulle sue esperienze, sui talenti, sulle sue capacità e sui suoi bisogni.

Le pratiche di insegnamento così ragionate risultano essere più efficaci poiché promuovono un livello di motivazione molto profondo ed un livello di successo molto alto (McCombs, Whisler, 1997).

Un apprendimento di questo tipo favorisce lo sviluppo di quelle competenze chiave per l'apprendimento permanente che il sistema formativo odierno è chiamato a supportare (Day, 2002).

Secondo i documenti ministeriali, i sistemi formativi odierni devono offrire una didattica finalizzata alla formazione di studenti critici e capaci di intervenire con atteggiamenti appropriati in situazioni complesse.

In questa prospettiva gli ambienti di apprendimento risultano interattivi ed efficienti. I nuovi ambienti di apprendimento non si basano esclusivamente sulle piattaforme virtuali ma includono anche risorse come *blog*, *wiki*, *social network*, ecc... Una moltitudine di scenari per costruire una realtà dinamica e accessibile da tutti (Cabero, Llorente, 2005).

A partire dalle indicazioni europee, l'Università di Foggia sta orientando sempre più la propria didattica nell'ottica della digitalizzazione dell'offerta formativa. Tale scelta è nata in risposta all'esigenza di realizzare esperienze di apprendimento sempre più personalizzate al fine di prevenire il costante *drop out* di studenti, causato spesso da un'emigrazione dei giovani locali verso altre zone del Paese e dalla difficoltà di frequentare le lezioni in presenza a causa di un territorio, come quello della Daunia, che circonda Foggia e che con essa è mal collegato.

Sulla base di tali premesse l'Università di Foggia ha scelto di ampliare la propria offerta formativa promuovendo una varietà di corsi MOOC all'interno della piattaforma Eduopen.

I corsi offerti dall'ateneo foggiano possono distinguersi in quattro tipologie:

- *Courseware* (non prevedono il rilascio di un certificato di partecipazione);
- *Courses* (corsi MOOC singoli);
- *Pathways* (moduli didattici composti da corsi MOOC);
- *Master* (Master universitari/Corsi di perfezionamento che ai fini del rilascio del titolo richiedono un'iscrizione presso l'Università che li eroga).

I *Courses* hanno una durata che varia dalle 3 alle 5 settimane (ad eccezione dei corsi “*self-placed*” che restano sempre disponibili sulla piattaforma) mentre i *Pathways* hanno una durata superiore alle 5 settimane.



Dalla tabella (1) si evince che in questo primo triennio di Eduopen l'ateneo foggiano ha prodotto 32 Corsi, 4 *Courseware*, 3 *Pathways* (corrispondenti a 3 corsi di perfezionamento o *Nano Master*) ed 1 Master (uno tra i primi master italiani ad essere erogato secondo la modalità MOOC da titolo "Innovazione didattica e CLIL, promosso nell'a.a. 2017/2018), con una produzione pari a 1278 ore di videolezioni registrate e 19.409 studenti iscritti.

2.1. *L'analisi del contesto*

La piattaforma Eduopen (learn.eduopen.org) è una piattaforma MOOC federata tra 21 Atenei pubblici italiani e istituti di ricerca e alta formazione che offre ambienti di apprendimento in cui si sviluppano diverse tipologie di interazione: docente-studente, studente-studente e studente-risorse. Questo spazio virtuale d'interazione ed acquisizione delle conoscenze è in grado di affiancare l'apprendimento tradizionale tramite l'utilizzo di sistemi informatici che aiutano gli utenti a capitalizzare le risorse.

All'interno della piattaforma gli utenti possono usufruire di supporti didattici classificabili in quattro categorie: videolezioni, materiali didattici di approfondimento, strumenti di valutazione (quiz, test di autovalutazione e compiti) e strumenti di comunicazione (wiki, forum, webinar e glossari).

I corsi proposti dall'ateneo foggiano sono in grado di offrire risposte a molteplici esigenze formative; all'interno del catalogo possiamo trovare corsi di preparazione alle PVI (Prove di Verifica Iniziali), corsi rivolti alla formazione iniziale e in servizio degli insegnanti e corsi di perfezionamento e aggiornamento professionale e un master post laurea di I livello (60 CFU).

Progettare un corso su piattaforma indica la definizione precisa degli obiettivi da perseguire, un'organizzazione dettagliata dei materiali didattici da somministrare e una pianificazione altamente strutturata delle prove di valutazione per verificare il raggiungimento degli obiettivi iniziali.

Lo staff del CEA UniFg (Centro E-learning di Ateneo) ha progettato e attivato diversi percorsi di formazione per tutti i docenti titolari dei corsi (Dipace et al., 2018); questi percorsi si sono sviluppati in quattro macro sessioni aventi i seguenti *focus*:

- innovazione delle pratiche didattiche.
- produzione di micro e macroprogettazione dei corsi.



- utilizzo degli strumenti mediali e delle e-tivities.
- metodologie di *online assessment* e *peer assessment*.

La fase successiva allo sviluppo di corsi MOOC ha riguardato la vera e propria produzione dei materiali audio-video con il supporto di video maker specializzati e alcune figure professionali specifiche, come *instructional designer* e informatici, presenti all'interno del team di ricerca e produzione CEA UniFg.

Dal 2016, sono stati complessivamente prodotti e implementati dall'Università di Foggia 32 MOOC.

Di seguito viene riportata la tabella (1) riepilogativa dell'offerta formativa dell'Università di Foggia sulla Piattaforma Eduopen.



Tipologia	Numero	Iscritti	Formato corso	Ore di formazione
Corsi	32	18153	984 videolezioni (durata dai 7 ai 15 minuti), 753 risorse didattiche aggiuntive, 64 prove di autovalutazione, 33 forum di discussione, 16 webinar	328
Courseware	4	161	65 videolezioni (durata dai 7 ai 15 minuti), 4 risorse didattiche aggiuntive	14
Pathway	3	1014	17 Corsi (rientrano all'interno del piano didattico del Master di I livello "Innovazione didattica e CLIL")	413
Master	1	81	60 CFU, 3 Pathway: "B2-English Grammar", "Cooperative Learning" e "CLIL"	623

Tab. 1: Mappatura dell'offerta formativa dell'Unifg erogata sulla piattaforma EduOpen

Al termine di ogni modulo sono state progettate e strutturate attività di valutazione quali quiz, test di autovalutazione e compiti, utili a guidare lo studente nel processo di autovalutazione (i primi) o di far comprendere al docente il livello dei discenti dell'aula virtuale (i secondi).

Il percorso di apprendimento in piattaforma, quindi, può modificarsi in base alle risposte che lo studente fornisce rispetto ad un dato stimolo.

Il docente che nella didattica tradizionale voglia seguire questa metodologia dovrà determinare le abilità/capacità necessarie al fine di ottenere il comportamento desiderato, per poi valutare gli studenti che vogliono impadronirsene.

Questo modello costruttivista mette in discussione la possibilità di una conoscenza oggettiva e tiene in considerazione il punto di vista di chi osserva e di chi esamina. L'acquisizione del sapere non avviene in forma passiva poichè ogni singolo utente inventa e costruisce in prima persona il suo mondo all'interno di una comunità virtuale attraverso la condivisione del linguaggio e delle esperienze.

In questo modo lo studente è posto al centro del processo formativo, contribuisce alla costruzione dei contenuti didattici mediante la negoziazione dei saperi (tramite l'utilizzo dei vari strumenti di comunicazione) ed è parte integrante di una rete d'interconnessioni che lo collegano con gli altri utenti.

Apprendimento e quotidianità in questo scenario si fondono in modo indistinguibile.

3. Metodologia

Lo scopo di questo studio è stato quello di esplorare la partecipazione degli studenti all'interno dei MOOC dell'Università di Foggia erogati sulla piattaforma EduOpen e di identificare i fattori caratterizzanti dell'utenza, per delineare quali aspetti contrassegnano chi decide di proseguire il proprio percorso all'interno dei 29 corsi MOOC Unifg (sono stati esclusi i *Courseware*).

L'approccio utilizzato è stato di tipo qualitativo (Case & Light, 2011), e lo strumento utilizzato per la raccolta dei dati è stato un questionario strutturato con il software *Survey Monkey*. Tramite il questionario sono state rilevate le 5 variabili oggetto di questa indagine:

- Genere
- Età
- Area geografica di residenza
- Titolo di studio:
- Precedenti esperienze MOOC

Il questionario è stato inviato agli 8041 studenti iscritti ai corsi Unifg (dati del 30 aprile 2019) ed è stato aperto da fine aprile a inizi maggio 2019.

Gli studenti che hanno partecipato all'indagine sono stati 1130, il 14,05% del totale.

Dei 1130 rispondenti, il 95,66% ha partecipato attivamente completando l'intero questionario, mentre il 4,34 % ne ha soltanto preso visione.



I rispondenti si sono iscritti a un totale di 2980 corsi per una media di 2,76 corsi per iscritto.

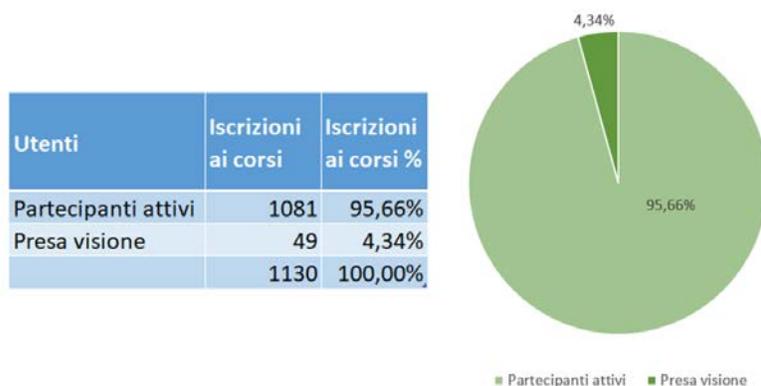


Fig. 1: Analisi partecipanti all'indagine



Ad ogni corso erogato dall'Università di Foggia (29 in totale escludendo dall'elenco i *Courseware*) è stato assegnato un ID univoco come indicato nella Tabella 2.

ID	Corso	ID	Corso
1	English Grammar for Food & Health	16	Principi di Patologia Generale: danno cellulare, infiammazione, riparazione, oncologia
2	English grammar for Person & Family	17	Approaches to economics, politics, law and criminal issues in Italian high schools according to the CLIL methodology
3	English Grammar for Study & Work	18	Percorsi di storia della letteratura italiana: l'epica cavalleresca
4	English Grammar for Travels & Sport	19	Insegnare con il CLIL materie scientifiche e tecniche
5	I 'come' e i 'perchè' della storia e della filosofia del diritto	20	Introduzione alla Fisica
6	Teorie e processi dell'apprendimento linguistico, language awareness	21	Processi di insegnamento e di apprendimento nelle classi CLIL
8	Promozione turistica attraverso i media digitali	22	Le 10 caratteristiche distintive dell'apprendimento cooperativo
9	Pre-corso di Matematica generale: Equazioni e Disequazioni	23	Le evidenze dell'apprendimento cooperativo

10	Progettazione e valutazione dell'apprendimento	24	Totò, la compravendita e il processo civile
11	Multifunzionalità e innovazione per l'agricoltura del futuro	25	Matematica per principianti: insiemi e operazioni elementari
12	Principi e obiettivi della valutazione, criteri per la valutazione, programmazione valutazione in CLIL. Rubriche e prodotti	26	Introduzione all'Economia
13	Principi delle buone pratiche di insegnamento nella classe CLIL, didattica per competenze	27	La forma di Governo. Attività amministrativa e garanzie procedurali del cittadino
14	Gli animali... chi sono? Chi siamo?	28	Banche e altri intermediari finanziari
15	Pedagogia ed Educazione: concetti base	29	La storia, le storie

Tab. 2: Elenco corsi UniFg con assegnazione ID



Successivamente alla raccolta dati sono state analizzate le diverse correlazioni tra i dati degli iscritti e le variabili indagate.

La prima correlazione studiata ha interessato la popolazione dei rispondenti (1081), il genere e l'iscrizione ai corsi (Figura 2). È emersa una prevalenza del genere Femminile (61,28%) sul genere Maschile (37,15%).

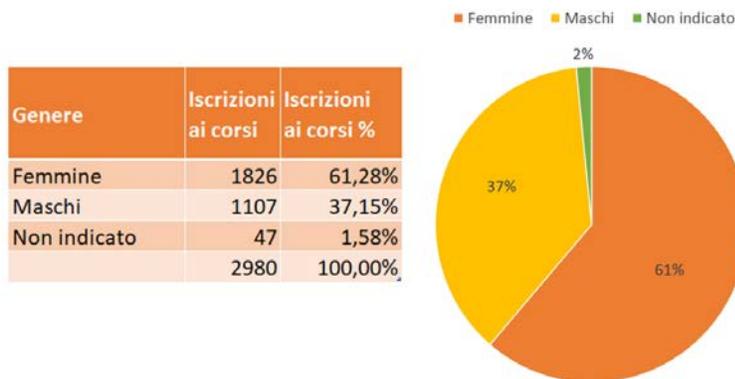


Fig. 2: Genere dei rispondenti all'indagine

La fase successiva ha riguardato l'analisi dettagliata del genere degli studenti iscritti in relazione ai corsi seguiti (Tabella 3).

ID	M	F	NI
1	70	126	4
2	82	154	4
3	79	166	4
4	66	129	3
5	51	27	1
6	15	48	1
8	47	76	2
9	48	33	3
10	29	90	3
11	12	10	2
12	12	42	1
13	13	53	1
14	18	23	0
15	53	115	1
16	38	33	0
17	63	115	3
18	37	85	2
19	15	41	1
20	44	27	2
21	17	48	1
22	25	71	2
23	24	74	2
24	13	8	0
25	84	71	2
26	17	17	0
27	30	26	0
28	29	15	1
29	76	103	1

Tab. 3: Correlazione dei corsi con la variabile genere



Emerge chiaramente una prevalenza del genere femminile in 20 corsi su 29, con un picco di iscritti di genere femminile in corrispondenza del corso con ID 3 (English Grammar for Study & Work).

La successiva analisi ha riguardato il fattore Età. I dati mostrano che circa la metà dei rispondenti è costituita da over 46 (50,94%) come da Figura 3.

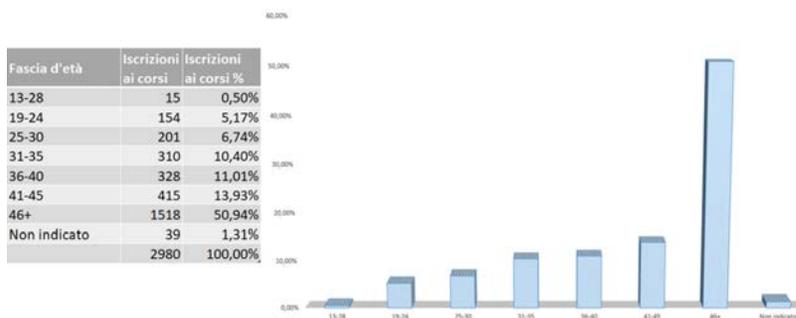


Fig. 3: Fasce d'età dei rispondenti all'indagine

Correlando la variabile Età all'iscrizione ai corsi UniFg emerge quanto segue in Tabella 4.

ID	13	19	25	31	36	41	46	NI
1	1	5	10	25	21	30	104	4
2	2	6	14	27	22	40	125	4
3	1	10	19	27	31	36	121	4
4	1	5	10	25	20	32	101	4
5	1	5	4	5	5	11	48	0
6	0	0	2	12	7	9	32	2
8	0	14	13	9	17	22	50	0
9	0	10	7	3	17	6	40	1
10	0	0	7	16	14	18	65	2
11	0	2	1	0	4	4	13	0
12	0	0	2	14	4	7	26	2
13	0	0	2	13	5	8	37	2
14	0	0	3	2	2	9	25	0
15	0	12	20	16	27	29	65	0
16	1	3	10	7	8	5	37	0
17	2	6	8	22	20	28	91	4
18	1	21	13	5	13	13	58	0
19	0	0	4	12	3	7	30	1
20	1	7	11	4	12	5	33	0
21	0	1	3	15	4	9	32	2
22	0	0	6	14	10	12	54	2
23	0	1	4	14	9	13	57	2
24	0	1	0	2	2	4	12	0
25	1	9	8	8	24	14	91	2
26	0	6	6	2	5	2	13	0
27	1	5	2	2	4	8	34	0
28	1	7	3	2	6	5	21	0
29	1	18	9	7	12	29	103	1

Tab. 4: Correlazione dei corsi con la variabile età

I dati mostrano chiaramente che su tutti i corsi la prevalenza di iscritti è di età superiore a 46 anni, con un picco in relazione al corso con ID 2 (English grammar for Person & Family).

Lo studio dei dati relativi alla variabile “Area geografica” mostra nettamente una maggioranza di iscritti provenienti dal Sud Italia, seguiti da studenti provenienti dal Nord-Est della penisola e dal Nord-Ovest.

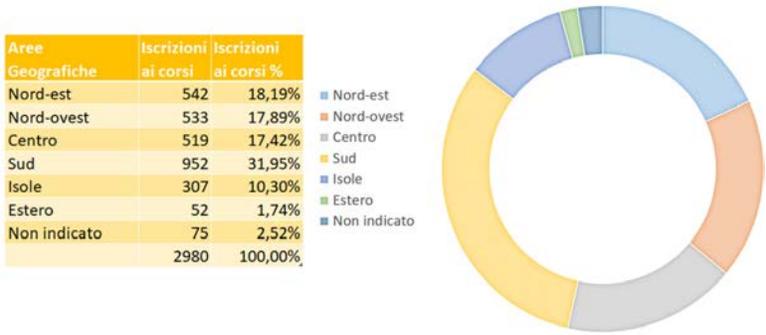


Fig. 4: Provenienza geografica dei rispondenti all'indagine

Correlando la variabile “Area geografica” ai corsi erogati UniFg emerge quanto segue in Tabella 5.

ID	NE	NO	CE	SU	IS	ES	NI
1	40	33	38	59	22	2	6
2	51	42	44	67	29	2	5
3	51	43	47	73	26	3	6
4	37	36	36	59	23	1	6
5	15	18	13	21	6	3	3
6	7	9	9	30	7	0	2
8	28	22	27	26	19	0	3
9	23	15	15	21	3	2	5
10	19	22	20	46	11	1	3
11	5	4	4	6	2	2	1
12	6	6	7	29	5	0	2
13	7	7	13	30	7	1	2
14	7	10	9	10	4	0	1
15	33	33	28	48	23	0	4
16	11	13	14	19	9	2	3
17	35	32	31	58	21	0	4
18	18	25	24	38	10	7	2
19	5	6	6	33	5	0	2
20	17	20	13	16	4	2	1
21	7	9	8	33	7	0	2
22	11	17	16	41	10	1	2
23	12	20	17	39	8	2	2
24	1	1	2	10	5	2	0
25	37	34	32	37	11	3	3
26	4	6	2	15	2	3	2
27	11	6	8	19	9	1	2
28	9	8	8	16	3	1	0
29	35	36	28	53	16	11	1

Tab. 5: Correlazione dei corsi con la variabile provenienza geografica

Il dato significativo riguarda nuovamente il corso con ID 3 (English Grammar for Study & Work), con 73 iscritti del Sud Italia.

Il corso successivo al corso ID 3 con un numero significativo di iscritti provenienti dal Sud Italia è il ID 2 (English Grammar for Person and Family).

L'analisi dei dati sulla variabile Titolo di studio mostra una percentuale maggiore di rispondenti in possesso di Laurea Magistrale o di Laurea a Ciclo Unico (30,20%) come da Figura 5.



Titolo di studio	Iscrizioni ai corsi	Iscrizioni ai corsi %
Scuola superiore di I grado	51	1,71%
Scuola superiore di II grado	821	27,55%
Laurea triennale	376	12,62%
LM o LMCU	900	30,20%
Post laurea	708	23,76%
Altro	116	3,89%
Non indicato	8	0,27%
	2980	100,00%

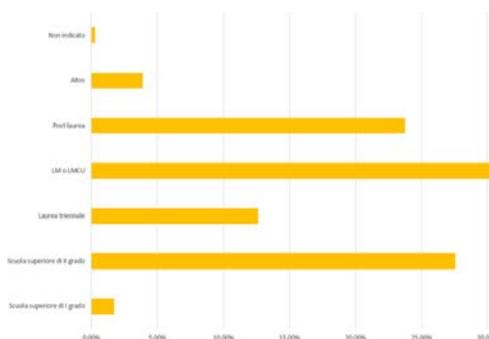


Fig. 5: Titolo di studio dei rispondenti all'indagine

Correlando la variabile Titolo di studio all'iscrizione ai corsi UniFg emerge quanto segue nella Tabella 6.



ID	1G	2G	LT	LM	PL	AL	NI
1	4	59	21	59	50	7	0
2	4	67	28	76	58	7	0
3	3	75	28	73	61	9	0
4	3	59	22	56	52	6	0
5	1	29	10	19	19	1	0
6	0	7	6	25	20	6	0
8	0	42	23	27	29	4	0
9	4	36	12	24	7	1	0
10	0	11	14	47	41	7	2
11	0	10	3	4	5	1	1
12	0	5	5	24	16	5	0
13	0	6	6	32	16	6	1
14	3	15	4	11	7	1	0
15	4	46	26	52	38	3	0
16	1	13	15	23	16	3	0
17	4	52	21	54	44	6	0
18	2	34	16	40	29	3	0
19	0	7	7	20	18	5	0
20	3	38	10	14	7	1	0
21	0	7	5	30	18	6	0
22	0	10	10	36	37	5	0
23	0	9	12	38	35	5	1
24	0	9	4	2	4	1	1
25	8	68	21	33	23	4	0
26	1	8	9	6	7	2	1
27	1	15	11	14	12	2	1
28	1	21	7	11	3	2	0
29	4	63	20	50	36	7	0

Tab. 6: Correlazione dei corsi con la variabile titolo di studio

Il dato significativo riguarda, ancora una volta, l'iscrizione al corso con ID 2 (English Grammar for Person & Family) che registra un totale di 76 iscritti in possesso di Laurea Magistrale.

Alla domanda "Prima di seguire un Mooc su EduOpen, avevi già seguito altri corsi online?" ha risposto in modo affermativo la maggioranza dei partecipanti all'indagine, per un totale di 1889 (63,39%).

Precedenti esperienze MOOC	Iscrizioni ai corsi	Iscrizioni ai corsi %
Sì	1889	63,39%
No	1002	33,62%
Non ricordo	88	2,95%
Non indicato	1	0,03%
	2980	100,00%

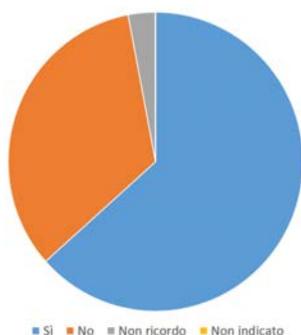


Fig. 6: Precedenti esperienze MOOC dei rispondenti all'indagine

Correlando la variabile Titolo di studio all'iscrizione ai corsi UniFg emerge quanto segue nella Tabella 7.

ID	Si	No	NR	NI	ID	Si	No	NR	NI
1	123	72	5	0	16	45	24	2	0
2	145	91	4	0	17	110	67	4	0
3	140	100	9	0	18	75	44	5	0
4	119	75	4	0	19	46	11	0	0
5	54	23	2	0	20	48	22	3	0
6	45	17	2	0	21	50	14	2	0
8	77	42	6	0	22	72	24	2	0
9	46	34	4	0	23	75	23	2	0
10	89	28	4	1	24	16	5	0	0
11	14	10	0	0	25	91	60	6	0
12	40	13	2	0	26	20	14	0	0
13	52	13	2	0	27	37	18	1	0
14	27	12	2	0	28	26	18	1	0
15	105	57	7	0	29	102	71	7	0

Tab. 7: Correlazione dei corsi con la variabile Precedenti esperienze MOOC

I dati mostrano un numero significativo in corrispondenza dell'ID 2 (English Grammar for Person & Family), seguito dal corso ID 3 (English Grammar for Study and Work).



4. Discussione dei risultati

Lo studio esplorativo ha permesso di tracciare un primo profilo dell'utente medio che si iscrive ai corsi erogati dall'Università di Foggia su piattaforma EduOpen. I dati rilevati e la loro analisi ed elaborazione hanno messo in evidenza la prevalenza di un'utenza femminile con età superiore ai 46 anni, con un titolo di studio magistrale, provenienti dal Sud Italia e che già hanno avuto esperienza in altri corsi MOOC online.

Gli utenti, inoltre, dimostrano una preferenza dei corsi con ID 2 e 3 in cui le singole variabili si concentrano maggiormente (come indicato in Tabella 8).

I due corsi, inclusi all'interno del primo Pathway del Master di I livello in "Innovazione didattica e CLIL", approfondiscono le materie di studio di un corso preparatorio alla certificazione di lingua inglese B2.

ID	Genere	Fascia d'età	Area Geografica	Titolo di studio	MOOC
	F	46	SU	LM	Si
1	126	104	59	59	123
2	154	125	67	76	145
3	166	121	73	73	140
4	129	101	59	56	119

Tab. 8: Profilo dell'utente tipo nei corsi UniFg

5. Conclusioni

Questo articolo ha esplorato le relazioni tra 5 variabili di studenti iscritti a corsi MOOC erogati dall'Università di Foggia sulla piattaforma Eduopen.

In particolare sono stati rilevati i dati circa il genere, la fascia di età, la provenienza geografica, il titolo di studio e la partecipazione in passato ad altri MOOC.

Dai risultati principali è emerso che la percentuale maggiore di studenti iscritti ai corsi è di genere femminile (20 corsi su 28) e di età superiore ai 46 anni (50,94%).

Accanto a questi risultati, è emerso che la percentuale maggiore di studenti proviene dal Sud Italia, è in possesso di una laurea magistrale ed ha seguito in passato altri corsi MOOC.

Tali risultati risultano interessanti poiché testimoniano come l'incertezza e la crescente complessità del mercato del lavoro porti alla riflessione dell'adulto che non riesce agevolmente a collocarsi nel mercato, riconosce il bisogno di apprendere nuove competenze e rafforzare i propri saperi e quindi investe sulla propria crescita in termini di formazione per tutta la vita, in particolare nel Sud Italia.

In tal senso, dunque, si può concludere che vi sono alcuni aspetti che probabilmente possono essere ulteriormente esplorati in studi futuri attraverso l'analisi dei learning analytics. Quest'ultimi presentano numerosi vantaggi poiché permettono di effettuare delle previsioni dei comportamenti degli studenti in un determinato ambiente di apprendimento online al fine di migliorare i processi didattici e, al tempo stesso, di monitorare i processi formativi. I Learning Analytics rappresentano un valido supporto a vari livelli:

- a livello dell'istituzione per il monitoraggio della qualità dei processi formativi
- livello dei docenti e degli studenti perché rendono possibile una riflessione e analisi al fine di prendere decisioni basate sui dati (data-driven), laddove processi decisionali data-driven comportano naturalmente un'intensa analisi dei dati (Long, Siemens, 2011; Jones, 2012). Inoltre, attraverso i LA, i docenti possono identificare tempestivamente gli studenti più a rischio e suggerire interventi personalizzati.

Pertanto, i dati raccolti, in linea con quanto permettono i Learning Analytics, potranno rappresentare la base per la progettazione di ambienti di apprendimento personalizzato, attraverso cui promuovere esperienze di apprendimento più efficaci (Greller, Drachslrer 2012; Siemens, 2010). I Learning Analytics forniscono informazioni per la pro-



duzione di interventi in tempo reale durante l'intero svolgimento del processo di apprendimento. I feedback personalizzati e immediati rappresentano un supporto importante per l'apprendimento autoregolato dello studente e agiscono sulla sua motivazione, sortendo così effetti significativi sul conseguente successo formativo (Dipace et al., 2019).

I dati raccolti in questo studio consentono ulteriori analisi che possono portare ad una ampia e complessa comprensione dell'andamento degli studenti all'interno dei corsi, nonché suggerire nuove opportunità di apprendimento o differenti possibilità di azione per migliorare l'andamento degli studenti in termini generali (Campbell, DeBlois, Oblinger, 2007).

I Learning Analytics possono, per esempio, essere utilizzati per predire le performance dello studente, suggerire risorse più rilevanti o più adatte per l'apprendimento, aumentare la riflessione e consapevolezza da parte dello studente, rilevare comportamenti indesiderati o stati affettivi che inibiscono l'apprendimento dello studente (ad es. noia, frustrazione, inappagamento, insoddisfazione, ecc.).



Ringraziamenti

Si ringraziano, in particolare, Bojan Fazlagic per la disponibilità nella raccolta dei dati in piattaforma e Dario Mazzoli per la collaborazione nella fase di analisi ed elaborazione dei risultati.

Riferimenti bibliografici

- Cabero J., Llorente M.C. (2005). Las plataformas virtuales en el ámbito de la teleformación. *Revista electrónica Alternativas de Educación y Comunicación*. Retrieved January 20, 2019, from https://idus.us.es/xmlui/bitstream/handle/11441/24672/file_1.pdf?sequence=1
- Calise M. (2018). MOOCs: sfide e opportunità. In *Conferenza dei Rettori delle Università Italiane – CRUI. I magnifici incontri CRUI 2018*. Piano Nazionale Università digitale. Retrieved January 19, 2019, from http://www2.cruil.it/cruil/magnifici_incontri_cruil_2018/Tav3A%20-%20MOOC%20Sfide%20e%20opportunit%C3%A0.pdf
- Campbell J.P., DeBlois P. B., Oblinger D. (2007). “Academic Analytics: A New Tool for a New Era.” *Educause Review* 42 (4): 40–57.
- Case, J. M., & Light, G. (2011). Emerging methodologies in engineering education research. *Journal of Engineering Education*, 100(1), 186–210.
- Crui, (2014), *MOOCs MASSIVE OPEN ON-LINE COURSES. Prospettive e Opportunità per l'Università italiana*. Retrieved June 18, 2019 from https://www.cruil.it/images/allegati/pubblcazioni/2015/MOOC_2015.pdf

- Day, C. (2002). *Developing teachers: The challenges of lifelong Learning*. Londra:Routledge.
- Dipace, A., Perrella, S., Bellini, C. & Limone, P. (2018). Innovazione didattica e formazione dei docenti universitari: azioni strategiche del progetto UniTutor. In *Proceedings della Conferenza EMEM ITALIA 2017*.
- Dipace, A., Loperfido, F., Scarinci, A. (2019, in press). From Big Data to Learning Analytics for a personalized learning experience. *REM—Research on Education and Media*
- Greller, W., Drachslrer H. (2012). “Translating Learning Into Numbers: A Generic Framework for Learning Analytics.” *Educational Technology & Society* 15 (3): 42–57
- Llorente, M.C. (2006). El tutor en E-learning: aspectos a tener en cuenta. *Revista electronica de Tecnologia Educativa*, 20. Retrieved February 5, 2019, from <http://www.edutec.es/revista/index.php/edutec-e/article/download/517/250>
- McCombs, B. L., & Whisler, J. S. (1997). *The learner-centered classroom and school: Strategies for increasing student motivation and achievement*. San Francisco: Jossey-Bass.
- Sansone, N., Ligorio, M. B. & Dillenbourg, P. (2012). Progettare il Role Taking a sostegno del Collaborative Knowledge Building. *Qwerty-Open and Interdisciplinary Journal of Technology, Culture and Education*, 6(2), 288-304. Retrieved February 14, 2019, from <http://www.ckbg.org/qwerty/index.php/qwerty/article/view/129>
- Savelli, S. (2017). Massive Online Open Courses: The New Frontiers of Experience. *DigitCult - Scientific Journal On Digital Cultures*, 2(1), 25-38. doi:10.4399/97888255045143
- Siemens, G. (2010). “What are learning analytics”. *ELEARNSPACE: Learning, networks, knowledge, technology, community*.
- Siemens, G., & d Baker, R. S. (2012, April). Learning analytics and educational data mining: towards communication and collaboration. In *Proceedings of the 2nd international conference on learning analytics and knowledge* (pp. 252-254). ACM.
- Wenger, E. (1998). Communities of practice: Learning, meaning and identity. *Journal of Mathematics Teacher Education*, 6(2), 185-194. Retrieved February 16, 2019, from https://www.researchgate.net/publication/225256730_Wenger_E_1998_Communities_of_practice_Learning_meaning_and_identity

Valorizzare i Learning Analytics per promuovere la valutazione per l'apprendimento.

Il caso del Master in Leadership e Management in Educazione dell'Università Roma Tre

The valorisation of the Learning Analytics to promote assessment for learning.

The case of the Master in Leadership and Management in Education of Roma Tre University

Giovanni Moretti

Department of Education, Roma Tre University, giovanni.moretti@uniroma3.it

Arianna Giuliani

Department of Education, Roma Tre University, arianna.giuliani@uniroma3.it

Arianna Lodovica Morini

Department of Education, Roma Tre University, arianna.morini@uniroma3.it

125

The article deepens the theme of the assessment for learning (Domenici, 2012; Lucisano, 2014; Notti, 2014; Trincherò, 2014) investigating how the functionalities connected to Learning Analytics (LA) (Fulantelli & Taibi, 2014; Gaševi et al. 2015; Blikstein & Worsley, 2016) could enhance their development in students.

The survey was carried out by analyzing diachronically data of the Master in Leadership and Management in Education (Roma Tre University) from a.y. 2013/14 to a.y. 2017/18; the Master is designed in blended mode on a Moodle platform.

The use of LA in the start-up and monitoring phases of the Master, integrated by the data collected by quali-quantitative tools, allowed to collect data useful for design individualized interventions. The training activities include the study of materials and different types of assessment tests. The trainees have the opportunity to verify the adequacy of their knowledge through self-assessment activities that provide feedback (Butler & Winne, 1995; Shute, 2008; Giuliani, Moretti & Morini, 2016). Furthermore, request students to elaborate critical analysis texts allow to give them also qualitative feedback. The ongoing use of the LA in the monitoring and assessment phases, in this regard, is strategic. In fact, by using the LA functions it is possible to observe the activities carried out by the students and to give them feedback.

From the analysis of the final questionnaires it emerges that students consider to be effective the articulation of the Master, both in didactic and assessment aspects. Also the analysis of the contributions in itinere and of the final test confirms how the planning of the formative activities has been effective considering the quality of the papers.

Keywords: Assessment; Feedback; Learning; Learning Analytics; University.

Il contributo approfondisce il tema della valutazione per l'apprendimento (Domenici, 2012; Lucisano, 2014; Notti, 2014; Trincherò, 2014) indagando in che modo le funzionalità connesse ai Learning Analytics (LA) (Fulantelli & Taibi, 2014; Gaševi et al. 2015; Blikstein & Worsley, 2016) possono valorizzarne lo sviluppo nei soggetti in formazione.

Il contributo rappresenta il risultato di un lavoro congiunto degli autori, in particolare G. Moretti ha scritto § 1 e § 6, A. Giuliani ha scritto § 3 e § 4, A.L. Morini ha scritto § 2 e § 5.

L'indagine è stata svolta analizzando diacronicamente i dati relativi al Master in Leadership e Management in Educazione (Università Roma Tre) dall'a.a. 2013/14 all'a.a. 2017/18, progettato in modalità blended su piattaforma Moodle. L'utilizzo dei LA nelle fasi di avvio e di monitoraggio del Master, integrati dai dati raccolti mediante strumenti di tipo qualitativo, ha permesso di rilevare dati utili a predisporre interventi individualizzati. Le attività formative prevedono lo studio di materiali di approfondimento e consegne valutative di diversa tipologia. I corsisti hanno l'opportunità di verificare l'adeguatezza delle proprie conoscenze mediante attività di autovalutazione con restituzione di feedback (Butler & Winne, 1995; Shute, 2008; Giuliani, Moretti & Morini, 2016). La richiesta di elaborare testi di analisi critica, inoltre, consente di restituire giudizi valutativi anche di tipo qualitativo. L'utilizzo dei LA nelle fasi di monitoraggio e valutazione, a tal proposito, è di rilevanza strategica. Avvalendosi delle funzioni dei LA è possibile infatti osservare le attività svolte dai corsisti e restituire loro un feedback.

Dall'analisi dei questionari finali emerge che i corsisti ritengono efficace l'articolazione del percorso formativo nelle fasi didattiche e valutative. Anche l'analisi dei contributi in itinere e della prova finale conferma come la progettazione delle attività formative sia stata efficace considerata la qualità degli elaborati.

Parole chiave: Apprendimento; Feedback; Learning Analytics; Università; Valutazione.



1. Contesto della ricerca

L'introduzione delle nuove tecnologie nella didattica e l'avvio di attività formative in modalità blended o a distanza, hanno notevolmente modificato il modo di intendere l'educazione e di progettare l'articolazione dei differenti percorsi formativi. L'utilizzo delle risorse caratterizzanti i sistemi informatici in educazione, in particolare, ha consentito negli anni di predisporre disegni di ricerca volti ad approfondire in che modo rendere le funzionalità connesse all'utilizzo delle tecnologie una risorsa potenziale per l'apprendimento e per la qualificazione dei percorsi formativi (Galliani, 2004; Domenici, 2014; Calvani, 2014; Vivanet, 2014).

Alcune ricerche di settore hanno approfondito in che modo negli ambienti online fosse possibile avvalersi anche dei sistemi di tracciamento digitali per disporre di informazioni utili sull'efficacia della proposta didattica, acquisendo dati sull'accessibilità e sulle modalità con cui gli studenti usufruiscono delle risorse digitali per migliorare il loro percorso di apprendimento (Trentin, 2004; Anderson, 2008; Albanese et al., 2010).

La riflessione sui Learning Analytics (LA) accompagna il complesso processo di trasformazione tecnologica degli ambienti di apprendimento a partire dal 2011, anno in cui, nell'ambito della prima Conferenza internazionale sui LA, viene formulata una definizione di tale area di indagine. I LA vennero definiti strumenti utili per misurare, raccogliere, analizzare e presentare dati sugli studenti e sui loro contesti allo sco-

po di comprendere l'apprendimento e ottimizzare l'organizzazione degli ambienti in cui esso ha luogo (Siemens, 2013).

A partire da questa precisazione è stata ampia la risonanza che i LA hanno riscontrato nella letteratura scientifica di ambito educativo (Buckingham Shum, 2012; Ferguson, 2014). Costituisce una testimonianza di tale interesse condiviso anche la sinergia attivata tra la Società Italiana di Ricerca Didattica (SIRD) e la Società Italiana di e-Learning (Sle-L). Fin dall'inizio l'obiettivo dei ricercatori è stato quello di approfondire le modalità con cui integrare i LA nella progettazione didattica della formazione a distanza (Fulantelli, Taibi, 2014; Gašević et al., 2015; Blikstein, Worsley, 2016). In tale contesto, tenendo conto degli aspetti più rilevanti da indagare, gli attori individuati come strategici sono i docenti e gli studenti e tra gli *stakeholders* gli esperti di *computer science*, capaci di elaborare i modelli di LA.

Un utilizzo efficace dei LA nella didattica dovrebbe allontanarsi da una visione dei *tracking methods* rigida e connessa alle sole funzioni di controllo e valutazione. Le evidenze che emergono dal monitoraggio sistematico dei dati relativi agli studenti dovrebbero, piuttosto, costituire una base di dati solida su cui progettare interventi specifici finalizzati a restituire feedback individualizzati utili a sviluppare il senso di autoefficacia e l'autoregolazione dei soggetti in apprendimento (Zimmerman, 2000; Pellerey, 2006), nonché le loro conoscenze e competenze.

I LA rappresentano una risorsa strategica per qualificare e innovare i processi formativi se il loro utilizzo è preceduto da una approfondita riflessione su come sia possibile impiegarli al fine di migliorare le pratiche didattiche, le strategie di valutazione e i processi di apprendimento. In particolare è importante riflettere su come i dati rilevati con i LA possano essere utili per rispondere in modo efficace alle esigenze formative degli studenti (Calvani, Rotta, 2000; Trentin, 2001; Bonaiuti, 2006) e su come le risorse offerte dalle tecnologie e dai LA consentano di predisporre online attività di autovalutazione finalizzate a promuovere l'acquisizione di competenze cognitive e metacognitive (Domenici, 2012; Lucisano, 2014; Notti, 2014; Trincherò, 2014). La possibilità di rilevare dati per restituire feedback formativi ai corsisti, a questo proposito, è una delle motivazioni che incoraggia a proseguire nello studio di come sia possibile applicare in modo efficace i LA nella didattica. Il feedback, infatti, quando è formulato in modo accurato e incoraggiante, è individuato da numerose ricerche come elemento di qualità che consente di sviluppare nei soggetti in formazione l'autoefficacia e la disposizione positiva ad apprendere (ad esempio: Butler, Winne, 1995; Shute, 2008; Giuliani, Moretti, Morini, 2016).



2. Metodologia della ricerca

La ricerca ha previsto uno studio longitudinale sui dati riguardanti i corsisti che dall'a.a. 2013/14 all'a.a. 2017/18 hanno preso parte alle attività formative del Master universitario di II livello in *Leadership e Management in Educazione* (LME), attivo presso il Dipartimento di Scienze della Formazione dell'Università Roma Tre.

Il principale obiettivo della ricerca è stato approfondire in che modo i LA sono stati utilizzati negli anni per rilevare dati utili alla riprogettazione del percorso formativo sia in itinere sia dell'anno accademico successivo, sulla base delle esigenze manifestate dai corsisti. L'obiettivo trasversale è stato approfondire in che modo le risorse messe a disposizione dai LA consentono di restituire feedback individualizzati ai soggetti in apprendimento e dunque favorire una valutazione per l'apprendimento (OECD/CERI, 2008; Weeden et al., 2009; Aquario, Grion, 2017).

Nell'ambito della ricerca sono stati approfonditi gli strumenti di monitoraggio e di raccolta dati di tipo quali-quantitativo utilizzati nel corso di Master nei diversi anni accademici. Particolare attenzione è stata prestata anche alla modalità con cui i LA hanno consentito di triangolare i dati raccolti rispetto ai corsisti e al percorso formativo, al fine di consentire interventi tempestivi in risposta alle esigenze formative emerse.

Nel complesso dei 5 anni accademici presi in considerazione, l'unità di analisi è composta da 260 soggetti, di cui il 70% donne e il 30% uomini. Quasi la totalità dei corsisti, durante il periodo in cui ha partecipato al Master, ha ricoperto anche la funzione di docente in istituzioni scolastiche di diverso ordine e grado. In particolare il 34,2% dei corsisti ha insegnato nella Scuola primaria, il 33,8% nella Scuola secondaria di II grado, il 20,8% nella Scuola secondaria di I grado, il 5% nella Scuola dell'infanzia e il 6,2% ha svolto altri lavori, comunque connessi all'ambito educativo (es. docente distaccato presso l'Ufficio Scolastico Regionale, Dirigente Scolastico).

3. L'articolazione del Master LME



Il Master LME è attivo presso il Dipartimento di Scienze della Formazione dell'Università Roma Tre dall'a.a. 2003/2004. Esso ha il principale obiettivo di far sviluppare a coloro che operano nei diversi settori della formazione conoscenze e competenze strategiche per una gestione efficace dei processi educativi di cui sono attori.

Il Master articola le sue attività formative in modalità *blended* e si avvale di una piattaforma Moodle (<http://lme.crisfad.uniroma3.it/>). Il percorso formativo prevede una durata complessiva di 12 mesi, finalizzati ad acquisire un totale di 60 CFU.

Le attività che i corsisti sono tenuti a svolgere consistono nello studio di unità di apprendimento, nella realizzazione di attività di laboratorio (individuali e di gruppo), nella partecipazione a due seminari in presenza e nell'attuazione di attività di tirocinio, che prevede due fasi, la prima dedicata all'analisi del proprio contesto scolastico, la seconda di osservazione di un Dirigente Scolastico che opera e gestisce situazioni problematiche in un ordine scolastico diverso dal proprio.

Nel percorso formativo vengono valorizzate attività di autovalutazione riguardanti i temi approfonditi in ogni unità di studio. La funzione dell'autovalutazione e delle modalità di restituzione del feedback sarà oggetto di approfondimento specifico nel paragrafo che segue.

A conclusione del Master, è prevista una prova finale, in presenza, volta ad accertare la capacità dei corsisti di risolvere un caso problematico simulando di essere un Dirigente Scolastico utilizzando le conoscenze e le competenze acquisite nell'ambito del Master.

Con la prospettiva di migliorare la qualità della progettazione del Master LME le attività sono state accompagnate nel corso degli anni da diverse ricerche (Domenici, Moretti, Ciraci, 2012; Barzanò, Moretti, 2015; Moretti, 2019); in particolare è stata oggetto di approfondimento l'efficacia delle strategie didattiche utilizzate con gli studenti che si avvalgono dei percorsi formativi a distanza.

4. Strumenti di raccolta dati

Negli anni accademici che sono stati oggetto di approfondimento nella ricerca, il Master LME si è avvalso di diversi strumenti di tipo qualitativo per raccogliere dati sui corsisti frequentanti il percorso formativo. Alcuni strumenti sono stati integrati con quanto emerso dal monitoraggio delle attività svolte dai corsisti in piattaforma, in particolare sono state utilizzate alcune funzionalità proprie dei LA.

Gli strumenti utilizzati sono rimasti gli stessi dall'a.a. 2013/14 all'a.a. 2017/18, tuttavia negli anni ne sono state modificate e/o inte-



grate alcune sezioni al fine di renderli sempre più rispondenti alla esigenza di raccogliere dati utili alla riprogettazione del percorso formativo o alla restituzione di feedback individualizzati.

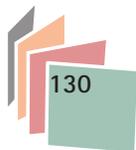
Il primo strumento somministrato ai corsisti nella fase di avvio del Master è il *Questionario di ingresso*.

Il questionario è implementato online e chiede di rispondere a domande di carattere socio-anagrafico e a domande che indagano ambiti specifici, quali ad esempio le motivazioni personali e professionali che hanno portato ad iscriversi al Master, le attese che si hanno del percorso formativo nel suo complesso e il livello di familiarità che si ha con gli ambienti di apprendimento online.

Le informazioni rilevate con il *Questionario di ingresso*, soprattutto nei primi anni considerati nella ricerca, hanno consentito di raccogliere dati utili ad assumere decisioni tempestive volte a rispondere agli interessi e alle esigenze manifestate dai corsisti cui il Master non aveva considerato in fase di progettazione. Le risposte date in merito alle capacità dei corsisti di avvalersi della piattaforma, ad esempio, hanno consentito di perfezionare le strategie utilizzate per rendere maggiormente accessibile l'ambiente di apprendimento online in tutte le sue sezioni, predisponendo dei tutorial utili nelle fasi di familiarizzazione con la piattaforma. Le risposte date in relazione agli interessi e alle motivazioni che avevano portato a scegliere di iscriversi al Master LME, d'altra parte, hanno costituito una risorsa per individuare materiali e strumenti ulteriori da condividere con i corsisti durante il percorso formativo.

Il principale elemento che ha consentito di avvalersi delle funzionalità dei LA nell'ambito del Master in ogni anno accademico è il *Monitoraggio sistematico* delle attività svolte in piattaforma.

In particolare, mediante l'analisi dei singoli interventi fatti nei forum da parte dei corsisti (forum sia di carattere generale sia di carattere specifico per ogni unità di studio) e delle interazioni sviluppate è stato possibile rilevare in che misura i corsisti prendevano parte attivamente alle attività del Master. L'analisi dei forum è stata di carattere sia qualitativo che quantitativo. Nel primo caso è stata approfondita la pertinenza e la consapevolezza con cui i corsisti scrivevano i propri interventi relativi ai contenuti di apprendimento del Master. Per fare questo a una prima fase di lettura da parte dei ricercatori è succeduta una fase di lavoro mediante il software di analisi testuale T-LAB, che ha consentito di individuare, nelle interazioni tra i diversi utenti, i concetti più ricorrenti sia in termini di singole parole chiave sia in termini di cluster (o nuclei tematici). Il software permette infatti di riconoscere sequenze di parole e di selezionare e classificare gli aspetti più significativi del testo. Quanto emerso è stato importante per verificare se le



discussioni avviate dai corsisti fossero pertinenti rispetto agli ambiti tematici del Master e non fossero solo di carattere informale.

La seconda modalità con cui è stata effettuata l'analisi dei forum è quella quantitativa; a tal fine sono stati calcolati gli indici di frequenza di alcune parole ritenute chiave per le diverse unità di studio prese in considerazione ed è stato conteggiato il numero delle interazioni tra utenti.

Nella fase di monitoraggio delle attività svolte dai corsisti, una delle risorse dei LA utilizzata sistematicamente negli anni è stata quella del *reporting*. Avere informazioni circa l'accesso, la lettura, la produzione di testi e lo scambio di interventi con i colleghi è stato strategico per avere informazioni circa la frequenza e l'intensità della partecipazione dei singoli corsisti alle attività del Master. Considerato che per ogni unità di studio si chiede ai corsisti di rispondere a più domande di approfondimento, inoltre, il monitoraggio delle risposte fornite consente ai soggetti in apprendimento di verificare il livello di comprensione che ognuno ha raggiunto, al fine di individuare eventuali strategie di recupero da suggerire ai singoli o al gruppo.

L'analisi dei forum e il sistema di *reporting* hanno costituito delle importanti risorse nell'ambito del Master LME per valutare l'andamento delle attività formative e per restituire feedback ai corsisti sia in caso di criticità nel proseguimento del percorso di studio – per rinnovare la disponibilità dei tutor al supporto didattico e tecnico – sia in caso di assolvimento positivo delle attività previste – per incoraggiare a proseguire in tal senso fino alla fine del percorso di studio. Le modalità con cui, negli anni accademici considerati, il Master LME ha utilizzato il monitoraggio come risorsa per progettare in itinere e negli anni successivi le attività formative sarà oggetto di approfondimento specifico nel paragrafo 5.

Uno degli strumenti utilizzati nell'ambito del Master, elemento distintivo di qualità fin dal suo avvio, è costituito dalle *Prove autovalutative* che integrano ogni unità di studio.

Le *Prove autovalutative* sono costituite da 10 domande con risposta a scelta multipla (di cui una sola corretta). Una volta completata la prova ogni corsista deve inviarla al sistema, che è progettato in modo che ognuno riceva un feedback immediato così da poter verificare in tempi rapidi il livello di comprensione e le eventuali lacune su cui tornare a lavorare.

Le prove, in particolare, prevedono la restituzione di un feedback di tipo individualizzato. In questo modo non solo è possibile leggere se la risposta fornita è corretta o sbagliata, ma è data a ogni corsista la possibilità di avere un messaggio approfondito in cui si spiega nel dettaglio il motivo per cui la risposta data o corretta o sbagliata (senza



eventualmente indicare la risposta corretta). Tale tipologia di feedback è stata individuata negli anni come la più efficace per migliorare la qualità degli apprendimenti dei corsisti e per far sviluppare loro capacità auto-regolative.

Le strategie di rilevazione dati e restituzione di feedback sono state utilizzate, nell'ambito dei diversi anni accademici approfonditi nella ricerca, in modo da favorire l'acquisizione da parte dei corsisti di competenze utili da trasferire in contesti anche diversi da quello dell'apprendimento online e della leadership educativa, ambito tematico fondamentale del Master LME. Uno degli strumenti con cui viene accertato il possesso di tali competenze è la *Prova finale*.

La *Prova finale* del Master LME consiste in una prova di simulazione semistrutturata. La consegna prevede di immaginare di assumere la funzione di Dirigente Scolastico, di definire una situazione problematica da risolvere e di formulare ipotesi utili a risolvere il caso individuato, argomentando in modo pertinente i diversi passaggi del processo decisionale. La strutturazione della prova valorizza un approccio critico alla conoscenza, così come fatto nel percorso formativo in itinere utilizzando gli strumenti descritti in precedenza. L'utilizzo sistematico dei LA per restituire feedback individualizzati ai corsisti costituisce in ogni annualità del Master una risorsa per consentire di lavorare in fase conclusiva sulla verifica di tali complesse competenze.

A conclusione del percorso viene somministrato anche il *Questionario finale*, che permette di rilevare dati circa l'efficacia del percorso formativo e di sviluppare una forma di follow-up con i corsisti.

Il questionario è articolato con domande chiuse e aperte e tra i diversi ambiti che approfondisce ci sono: la soddisfazione delle aspettative iniziali che si avevano rispetto al percorso formativo intrapreso; l'efficacia delle attività didattiche proposte e delle azioni tutoriali garantite; la pertinenza degli interventi seminari e di laboratorio proposti; la formulazione di un giudizio complessivo sulla progettazione educativa del Master.

Negli anni l'analisi dei dati rilevati con il *Questionario finale* si è dimostrata importante per individuare strategie utili per la riprogettazione delle attività di studio e di tirocinio negli anni accademici successivi.

5. L'utilizzo dei Learning Analytics nel Master LME



Nel Master LME i Learning Analytics sono stati utilizzati sistematicamente nel percorso, con finalità differenti. Durante l'avvio del Master, i LA permettono di raccogliere dati utili ad allestire l'ambiente e predisporre interventi individualizzati per rispondere alle esigenze emerse sulla base del profilo del singolo corsista. In questo senso, alcune delle strategie didattiche vengono progettate e messe a punto in maniera flessibile a seguito della rilevazione dei dati.

In itinere ci si avvale dei LA con funzione di monitoraggio, *reporting* e valutazione al fine di osservare le attività svolte dai corsisti e restituire loro feedback formativi individualizzati.

La ricerca di tipo longitudinale ha previsto quattro fasi che si sono ripetute nel corso degli anni accademici posti sotto osservazione (dal 2013 al 2018):

1. analisi dei fabbisogni formativi attraverso lo studio dei dati emersi dai LA e dagli strumenti di raccolta dati di natura quali-quantitativa;
2. progettazione di un intervento didattico che potesse rispondere alle esigenze emerse;
3. implementazione dell'intervento;
4. verifica dell'efficacia dell'intervento.



La Figura n. 1 rappresenta la ciclicità delle quattro fasi con cui è stata condotta la ricerca.



Fig.1: Fasi della ricerca di tipo longitudinale

Vengono presentati di seguito i principali esiti emersi dallo studio.

Nell'a.a. 2013/2014 è stata rilevata la necessità da parte dei corsisti di avere a disposizione maggiori occasioni di collaborazione e confronto tra pari. È importante precisare che l'interesse per il Master è sia na-

zionale che internazionale, infatti i corsisti provengono da diverse aree geografiche. Le attività del Master sono state quindi progettate per permettere a tutti di partecipare attivamente. Il Master viene avviato nella modalità a distanza, il primo seminario in presenza è generalmente previsto a distanza di cinque mesi dall'inizio del corso. In questo tempo i corsisti hanno la possibilità di familiarizzare con l'ambiente online, di comunicare sulla piattaforma con i tutor e gli altri corsisti attraverso le risorse di comunicazione sincrona e asincrona, e di avviare lo studio e approfondimento delle tematiche proposte.

A seguito del primo seminario, in cui i corsisti hanno occasione di conoscersi di persona e di entrare in contatto con i tutor, i docenti e gli esperti del Master, nell'a.a. 2013/2014 è stato rilevato un incremento nell'uso della chat e dei forum relazionali. In particolare sono stati gli stessi corsisti ad aprire nuovi argomenti di discussione in cui sollecitavano una riflessione condivisa.

Inoltre è aumentato il numero di corsisti che ha visualizzato gli elaborati di studio prodotti dai colleghi. Questi nuovi scambi online sono stati interpretati come una possibile richiesta di maggiore interazione tra pari. L'ipotesi è stata confermata nel questionario finale in cui, tra i suggerimenti indicati per migliorare il corso di Master, è stata proposta la messa a punto di attività specifiche in cui favorire il lavoro e la collaborazione a distanza tra studenti.

L'attività didattica che è stata progettata e implementata nell'a.a. 2014/2015 è stata la richiesta di approfondire una Unità di Studio con la modalità di lavoro di gruppo. La consegna chiedeva di studiare individualmente un determinato argomento e di rispondere collegialmente ad alcune domande stimolo. I gruppi sono stati istituiti tenendo conto dei dati emersi utilizzando i LA dal monitoraggio. In particolare, sono state integrate le informazioni relative l'accesso, la lettura, la produzione di testi e lo scambio di interventi in piattaforma per rilevare la frequenza, la motivazione e l'intensità della partecipazione dei singoli. In questo modo è stato possibile individuare diversi profili per costituire gruppi eterogenei in cui, ad esempio, corsisti più motivati hanno potuto contribuire a incoraggiare una maggiore interazione da parte dei corsisti meno motivati.

Per facilitare il lavoro di gruppo sono stati messi a disposizione degli appositi forum. Gli elaborati di ogni gruppo sono stati caricati online per permettere a tutti di visualizzare il lavoro svolto anche dagli altri. L'analisi effettuata sui testi inseriti dai corsisti nei forum, sia per caricare il proprio lavoro sia per interagire con i colleghi, ha evidenziato come, con il passare del tempo, i corsisti utilizzassero progressivamente un linguaggio sempre più specifico e tendessero a dialogare piuttosto che



a fare interventi monodirezionali.

A conclusione delle attività sono stati restituiti dei feedback formativi, formulati sulla base di alcuni parametri di riferimento tra cui: la completezza del lavoro, la qualità del contenuto, l'originalità dell'elaborato, il rispetto dei tempi, le modalità con cui il gruppo ha organizzato e portato avanti la consegna, con particolare riferimento ai processi negoziali attivati tra i membri.

I dati emersi grazie all'uso dei LA hanno permesso di analizzare le interazioni interne e il livello di partecipazione all'attività, garantendo la formulazione di un feedback individualizzato. Sono stati considerati gli scambi presenti nei forum e lo storico della risorsa "Wiki", messa a disposizione per la costruzione condivisa e collaborativa dell'elaborato finale. Attraverso il "Wiki" è infatti possibile monitorare le azioni del singolo in termini di inserimento e modifica dei contenuti al fine di rilevare le modalità e l'intensità della partecipazione.

Nell'a.a. 2014/2015 è stato manifestato l'interesse da parte dei corsisti di poter avere un confronto con gli esperti e i docenti che hanno predisposto le Unità di Studio.

L'esigenza è stata documentata attraverso lo studio dei forum in cui i partecipanti hanno formulato domande di chiarimento e approfondimento relative agli argomenti affrontati. Le domande sono state raccolte e sistematizzate al fine di costruire un'intervista strutturata da somministrare agli esperti. Le interviste sono state videoregistrate e caricate nell'ambiente online.

I video sono stati largamente apprezzati dai corsisti. La totalità ha infatti preso visione del nuovo materiale e, nel questionario finale, è stato confermato l'interesse verso la nuova risorsa messa a disposizione nell'ambiente online che è stata ritenuta strategica per migliorare la qualità degli apprendimenti.

Nell'anno accademico successivo (2015/2016), questa dimensione è stata ulteriormente arricchita. È stato rilevato che, nonostante i video avessero risposto alla necessità di chiarimento su alcune questioni relative alle Unità di Studio, i corsisti sentissero la necessità di un confronto diretto con gli esperti per poter accrescere le loro conoscenze e le loro competenze sui temi connessi alla dirigenza e al management nei contesti educativi.

Nel Master, prevedendo attività in modalità blended, è stato possibile progettare un intervento didattico nell'ambito dei seminari di studio in presenza. In questa occasione sono state allestite delle sessioni tematiche parallele. I corsisti hanno così avuto la possibilità di confrontarsi in piccoli gruppi di lavoro, facilitando lo scambio di punti di vista sia tra pari sia con esperti. L'attività ha contribuito a valorizzare la



riflessione individuale, sviluppando consapevolezza e spirito critico.

A partire dall'a.a. 2015/2016, è stato riscontrato un incremento nell'utilizzo delle risorse dedicate all'autovalutazione. Come descritto nel paragrafo precedente, ogni Unità di Studio (UdS) prevede, oltre alle domande stimolo di approfondimento, una prova strutturata di tipo autovalutativo per permettere ai corsisti di rilevare il livello di apprendimento raggiunto. La prova è costituita da domande a scelta multipla che contribuiscono a riflettere sui contenuti della UdS. Inizialmente la prova prevedeva una restituzione esclusiva del punteggio ottenuto, sulla base delle risposte corrette. L'interesse dimostrato dai corsisti è stato di ricevere un feedback maggiormente approfondito e dettagliato. Per qualificare il processo di apprendimento e migliorare la funzione autovalutativa dello strumento sono stati costruiti dei feedback formativi individualizzati per ciascuna alternativa di risposta. Il feedback, oltre a specificare se l'alternativa individuata è corretta o meno, fornisce una spiegazione che argomenta il motivo. Avvalersi di una restituzione specifica e immediata, permette di contribuire a sviluppare le capacità auto-regolative e di orientare nuovamente il processo di apprendimento. Confermata la validità del dispositivo, tra il 2015 e il 2017 tutte le prove autovalutative delle UdS sono state arricchite con i feedback individualizzati.

Nella Figura n. 2 viene riportato lo sviluppo temporale in cui è possibile rilevare l'analisi e gli interventi didattici progettati e implementati nei diversi anni accademici.



Fig.2: Analisi dei dati e strategie di intervento nella didattica del Master LME dall'a.a. 2013/2014 all'a.a. 2017/2018

Dall'analisi dei dati raccolti in maniera diacronica mediante il questionario finale e i forum presenti in piattaforma, sono emersi alcuni

punti di forza in merito alle attività didattiche e valutative implementate nel corso degli anni.

I corsisti rilevano come l'introduzione di diverse modalità di elaborazione e riflessione, tra cui le attività in gruppi di lavoro, siano state ritenute efficaci. Come scrive A.P. (a.a. 2015/2016) *“la tipologia delle attività di approfondimento e delle modalità di condivisione online sono un punto di forza”*.

L'integrazione dei feedback nelle prove autovalutative, sono state valutate positivamente negli anni in cui sono state progettate e implementate. Nello specifico, sono state considerate utili per promuovere la capacità riflessiva e metariflessiva (T.S, a.a. 2016/17 e R.B. a.a. 2017/2018). Ad esempio F. D. (a.a 2017/2018) segnala che *“le attività proposte rendono maggiormente consapevoli i corsisti delle proprie competenze”*.

Conclusioni

Nel Master in *Leadership e Management in Educazione* i LA sono stati utilizzati in modo attivo e diacronico per rilevare dati utili alla riprogettazione flessibile del percorso formativo sia in itinere sia in vista della successiva edizione del Corso.

I dati raccolti attraverso i questionari finali nei diversi anni accademici hanno evidenziato una valutazione positiva dei corsisti in merito all'articolazione del percorso formativo. Gli interventi didattici progettati e implementati a seguito del monitoraggio effettuato valorizzando i LA e utilizzando differenti strumenti di rilevazione dati, sembrano essere stati accolti positivamente dai corsisti e considerati efficaci per migliorare il loro percorso formativo.

L'analisi integrata delle diverse fonti di dati ha consentito di qualificare l'offerta formativa del Master LME nel suo complesso e di promuovere una valutazione per l'apprendimento, soprattutto mediante l'attenzione specifica che negli anni è stata riservata in modo sistematico all'importanza del ruolo dei feedback formativi individualizzati.

L'introduzione di modalità con cui autovalutare individualmente e verificare tra pari il proprio percorso formativo ha innalzato la qualità della proposta didattica a distanza, favorendo la costruzione di una comunità di apprendimento.

I corsisti hanno avuto l'opportunità di monitorare l'adeguatezza delle proprie conoscenze, sia mediante attività di confronto, sia avvalendosi di feedback formativi. L'analisi dei risultati emersi dalla prova finale conferma che la progettazione e la riprogettazione delle attività didattiche,



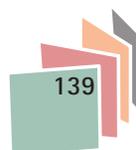
sia online che in presenza, sono state efficaci e hanno contribuito negli anni a migliorare la qualità degli apprendimenti e a rispondere in modo flessibile alle esigenze formative manifestate dai corsisti.

Riferimenti bibliografici

- Albanese O., Businaro N., Cacciamani S., De Marco B., Farina E., Ferrini T., Vanin L. (2010). Riflessione metacognitiva in ambienti online e autoregolazione nell'attività di studio nei corsi universitari. *Italian Journal of Educational Technology*, 18(1), 50-50.
- Anderson T. (Eds.). (2008). *The theory and practice of online learning*. Canada: Athabasca University Press.
- Aquario D., Grion V. (2017). "Valutazione per l'apprendimento": autovalutazione e valutazione fra pari in alcuni corsi dell'Università di Padova. In E. Felisatti, A. Serbati (Ed.), *Sviluppare la professionalità docente e innovare la didattica universitaria* (pp. 232-249). Milano: Franco Angeli.
- Barzanò G., Moretti G. (2015). Towards Headship: Developing Leadership and Management Skills in an Italian University Master Course. *Journal of Educational, Cultural and Psychological Studies (ECPS Journal)*, (11), 225-252.
- Blikstein P., Worsley M. (2016). Multimodal Learning Analytics and Education Data Mining: using computational technologies to measure complex learning tasks. *Journal of Learning Analytics*, 3(2), 220-238.
- Bonaiuti G. (2006). *E-Learning 2.0: Il futuro dell'apprendimento in rete, tra formale e informale*. Trento: Erickson.
- Buckingham Shum S. (2012). UNESCO Policy Brief: Learning Analytics. *UNESCO Institute for Information Technologies in Education*. Retrieved (in www.ite.unesco.org/publications/3214711).
- Butler D.L., Winne P. H. (1995). Feedback and self-regulated learning: A theoretical synthesis. *Review of educational research*, 65(3), 245-281.
- Calvani A., Rotta M. (2000). *Fare formazione in Internet: manuale di didattica online* (Vol. 2). Trento: Erickson.
- Calvani A. (2014). L'innovazione tecnologica nella scuola: come perseguire un'innovazione tecnologica sostenibile ed efficace. *LEA-Lingue e Letterature d'Oriente e d'Occidente*, 2, 567-584.
- Domenici G. (2012). Valutazione e autovalutazione come risorse aggiuntive nei processi di istruzione. *Education Sciences & Society*, 2(2), 69-82.
- Domenici G. (2014). Tecnologie digitali, successo formativo e qualificazione della didattica). *Journal of Educational, Cultural and Psychological Studies (ECPS Journal)*, (10), 13-22.
- Domenici G., Moretti G., Ciraci A.M. (2012). The master's course in leadership and management in education: results of the monitoring and reasearch activities. *8th International Conference on Education, Research and Training Institute of the East Aegean (INEAG)*, Greece EDITOR, 239-247.
- Gašević D., Dawson S., Siemens G. (2015). Let's not forget: Learning analytics are about learning. *TechTrends*, 59(1), 64-71.



- Ferguson R. (2014), Learning Analytics: fattori trainanti, sviluppi e sfide. *Tecnologie didattiche*, 22(3), 138-147.
- Fulantelli G., Taibi D. (2014). Learning Analytics: opportunities for schools. *Italian Journal of Educational Technology*, 22(3), 157-164.
- Galliani L. (2004). *La scuola in rete*. Roma-Bari: Laterza.
- Giuliani A., Moretti G., Morini A. (2016). Didactic tutoring services and obligations for additional learning, an empirical exploratory research: the case of the Department of Educational Science at Roma Tre University. *Italian Journal of Educational Research*, (15), 63-78.
- Lucisano P. (2014). Responsabilità sociale, valutazione e ricerca educativa. *Italian Journal of Educational Research*, 13-20.
- Moretti G. (2019). Sviluppo professionale degli insegnanti e ruoli di leadership: lo sviluppo di competenze degli insegnanti in un corso di Master. In L. Perla, M. Berta (a cura di), *Professione insegnante. Idee e modelli di formazione* (pp. 158-170). Milano: FrancoAngeli.
- Notti A. M. (2014). La valutazione come attribuzione di valore. In A. M. Notti (Ed.), *A scuola di valutazione* (pp. 11-22). Lecce: Pensa MultiMedia.
- OECD/CERI (2008). *Assessment for learning: Formative assessment. International Conference, Learning in the 21st Century: Research, Innovation and Policy*. Parigi: CERI/OECD.
- Pellerey M. (2006). *Dirigere il proprio apprendimento*. Brescia: La Scuola.
- Shute V. J. (2008). Focus on Formative Feedback. *Review of Educational Research*, 78(1), 153- 189.
- Siemens G. (2013), Learning Analytics: The Emergence of a Discipline, *American Behavioural Scientist*, 57(10), 1380-1400.
- Trentin G. (2004). *Apprendimento in rete e condivisione delle conoscenze: ruolo, dinamiche e tecnologie delle comunità professionali online* (Vol. 6). Milano: FrancoAngeli.
- Trincherò R. (2014). *Valutare l'apprendimento nell'e-learning*. Trento: Erickson.
- Vivanet G. (2014). Sull'efficacia delle tecnologie nella scuola: analisi critica delle evidenze empiriche. *Italian Journal of Educational Technology*, 22(2), 95-100.
- Weeden P., Winter J., Broadfoot P. (2009). *Valutazione per l'apprendimento nella scuola. Strategie per incrementare la qualità dell'offerta formativa*. Trento: Erickson.
- Zimmerman B.J. (2000). Self-efficacy: An essential motive to learn. *Contemporary educational psychology*, 25(1), 82-91.



Learning Analytics per il monitoraggio e l'analisi delle carriere degli studenti universitari

Learning Analytics for monitoring and analyzing student's careers at University

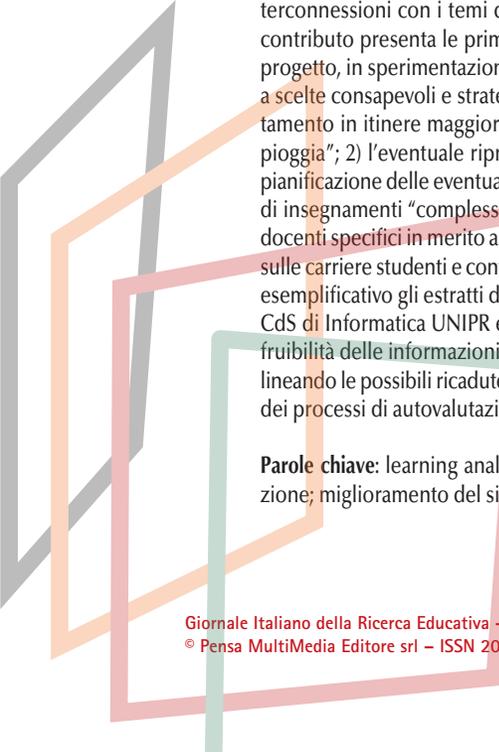
Daniela Robasto

University of Parma



After a preliminary analysis on the concept of Learning Analytics and its necessary interconnections with Big Data and Business Intelligence in the academic field, the paper presents the first steps of a pilot project at the University of Parma. The pilot project, being tested, is aimed at using Learning Analytics in order to lead to informed and strategic choices regarding: 1) an enhancement of guidance actions in itinere more calibrated on the type of problematic and not distributed "in the rain"; 2) the possible redesign of the study plans of the CdS subject to analysis and the planning of possible preparatory courses between courses; 3) the early identification of "complex / problematic" courses) the design of specific teacher training seminars on teaching and docimological skills, starting from real data on student careers and contextualised by disciplinary sectors. The contribution is given as an example the "screens" of a DEMO that is under construction and experimentation at the CdS of Informatica UNIPR and which is considered as an example for the impact in terms of usability of the information for different types of stakeholders. The work ends by emphasizing the possible effects of pedagogical research on LA, in the view of consolidation of the self-assessment and improvement processes of Italian universities.

Keywords: learning analytics; career monitoring for university students; self evaluation; improvement of the university system; university teaching.



Dopo un'analisi preliminare sul concetto di Learning Analytics e delle sue necessarie interconnessioni con i temi dei Big Data e Business Intelligence in ambito accademico, il contributo presenta le prime tappe di un progetto pilota presso l'Università di Parma. Il progetto, in sperimentazione, è volto ad utilizzare i Learning Analytics al fine di condurre a scelte consapevoli e strategiche in merito a: 1) un potenziamento delle azioni di orientamento in itinere maggiormente calibrate sul tipo di problematica e non distribuite "a pioggia"; 2) l'eventuale riprogettazione dei piani di studi dei CdS oggetto di analisi e la pianificazione delle eventuali propedeuticità tra insegnamenti; 3) l'individuazione precoce di insegnamenti "complessi/problematici" 4) la progettazione di seminari di formazione docenti specifici in merito a competenze didattiche e docimologiche, a partire da dati reali sulle carriere studenti e contestualizzati per settori disciplinari. Il contributo riporta a titolo esemplificativo gli estratti di una DEMO che è in costruzione e sperimentazione presso il CdS di Informatica UNIPR e che si considera esemplificativa per l'impatto nei termini di fruibilità delle informazioni per diverse tipologie di stakeholders. Il lavoro termina sottolineando le possibili ricadute della ricerca pedagogica sui LA nei termini di consolidamento dei processi di autovalutazione e miglioramento delle università italiane.

Parole chiave: learning analytics; monitoraggio carriere studenti universitari; autovalutazione; miglioramento del sistema universitario; didattica universitaria.

Con *Learning Analytics* (LA) ci si riferisce alla misurazione, alla raccolta, all'analisi e alla presentazione dei dati sugli studenti e sui loro contesti, ai fini della comprensione e dell'ottimizzazione dell'apprendimento e degli ambienti in cui ha luogo¹ (Siemens, 2013; Ferguson, 2012). I LA, rispetto ad altre indagini educative con rilevazioni empiriche sugli studenti, si contraddistinguono tuttavia per un uso di dati che solitamente è raccolto per *fini altri*, preesistente, e che solo successivamente può essere utilizzato, monitorato, trattato per trarre informazioni sui processi e percorsi di apprendimento. Vi è pertanto differenza sostanziale tra un dato raccolto intenzionalmente ai fini di una ricerca educativa e all'interno di un preciso impianto di ricerca, e un dato raccolto per fini prevalentemente anagrafici-amministrativi-organizzativi. Probabilmente questa iniziale assenza di intenzionalità "educativa" della rilevazione ha portato la ricerca pedagogica a non volgere prontamente lo sguardo verso le opportunità offerte dalla digitalizzazione dei dati relativi agli attori dei sistemi educativi e formativi salvo avvertire, più recentemente, che gli studenti e i docenti lasciano dietro di sé *molte tracce*, la cui analisi può convertirle in benefici concreti per il miglioramento del settore istruttivo (Duval, 2011).

Il tema dei LA è quindi intrinsecamente connesso con il tema dei 'big data' (Manyika et al., 2011, p.1) e necessariamente intrecciato con i continui sviluppi della *business intelligence*. Ora, che l'Università si avvii ad utilizzare taluni strumenti della business intelligence *anche* in riferimento ai processi di apprendimento dovrebbe riportare sul piatto della bilancia alcune considerazioni relativamente al *core business* dell'università (i processi di apprendimento e insegnamento) e al destinatario principale del suo "servizio" (lo studente). Se infatti il mondo impresa, per suo necessario sostentamento, ha imparato ad impiegare tecniche di analisi per estrarre *valore aggiunto* da insiemi di dati, usando per identificare modelli di comportamento, sviluppare campagne pubblicitarie e guidare i sistemi di raccomandazione (Ferguson, 2014), processi sempre orientati e finalizzati al suo *core* principale, il mondo universitario potrebbe invece rischiare di avere a disposizione milioni di dati e non utilizzarli, e/o utilizzarli in modo non strumentale rispetto al *core business* o rispetto al *target* principale.

Sappiamo che nel frangente in cui uno studente si iscrive all'università e diventa "matricola" (e oggi frequentemente anche prima, grazie

1 Tale è la definizione di Learning Analytics fu quella riportata nella call for papers della prima conferenza internazionale sul Learning Analytics and Knowledge (LAK 2011) e adottata dalla Society for Learning Analytics Research (SoLAR).

al monitoraggio dei sistemi di orientamento in ingresso), l'ateneo inizia ad accumulare dati ancorati al numero di matricola; la raccolta terminerà solo dopo il conseguimento della laurea o, nei casi meno favorevoli, dopo l'abbandono universitario. I database degli atenei ogni giorno accumulano una quantità crescente di dati sugli studenti, dati personali, informazioni di sistema e informazioni sul percorso accademico (Mazza & Milani, 2004; Romero, Ventura, & García, 2008). Ma cosa accade dopo la raccolta? Anche se le capacità di tracciare gli studenti sono in genere incluse tra le funzionalità generiche di molti software in uso negli atenei, le funzionalità di estrazione e aggregazione, di reporting e di visualizzazione di questi strumenti di analisi sono spesso di base o inesistenti (Dawson, 2009).

Può dunque accadere, non così raramente, che le decisioni strategiche anche sul tipo di report o sulle *query* da avviare sulle innumerevoli matrici dati, non sia sempre calibrato sul core business (l'apprendimento) o sul target (lo studente). Il conseguente rischio è che l'analisi e la reportistica nasca da esigenze più o meno segmentate e contingenti (monitoraggio degli indicatori per FFO², monitoraggio per risorse PRO³; monitoraggio dei CFU conseguiti; monitoraggio degli iscritti etc.) che potrebbero determinare provvedimenti altrettanto contingenti e non sistemici, nei termini del miglioramento dei processi di insegnamento e apprendimento. La differenza tra un processo sistemico e un processo contingente potrebbe dunque risiedere proprio nelle competenze di *Analytics*⁴, cioè in quelle competenze di analisi indirizzate a trasformare i dati grezzi in *informazioni di valore* per i decision maker accademici. Se è pur vero che il monitoraggio degli indicatori ai fini



- 2 L'acronimo FFO sta per *Fondo di Finanziamento Ordinario* destinato alla copertura delle spese istituzionali, tra cui i costi di personale, e di funzionamento. L'evoluzione normativa degli ultimi anni ha modificato radicalmente le modalità di attribuzione delle risorse statali al sistema universitario introducendo criteri che, gradualmente, riducono il peso dei finanziamenti su base storica a favore di parametri quali: il costo standard per studente; la quota premiale in relazione ai risultati della didattica e della ricerca; gli interventi perequativi a salvaguardia di situazioni di particolare criticità (Fonte MIUR <https://www.miur.gov.it/finanziamenti>).
- 3 Il Riferimento è alla Programmazione Triennale di Sviluppo secondo quanto previsto dal Decreto Ministeriale 12 maggio 2017 n. 264 - Linee generali di indirizzo 2016 – 2018 – Ammissione a finanziamento dei progetti presentati dalle Università.
- 4 Le aree di competenza dell'Analytics si dividono in quattro strategie di Analytics: Descriptive, Predictive, Prescriptive e Automated (si veda Osservatorio del Politecnico di Milano, Big Data Analytics & Business Intelligence, Report 2018, *Casi di adozione di progettualità di Analytics*). Si veda anche Gartner, 2013.

dell'FFO è essa stessa un'informazione di *valore*, quanto mai necessaria al sostentamento e miglioramento dell'università, va detto che i database delle università sono densi di dati che potrebbero restituire altrettante informazioni di *valore*, dove per valore si intenda *la capacità del bene di soddisfare un bisogno* e il bisogno sia quello dell'apprendimento e dell'alta formazione degli studenti.

In tale sfumato confine tra un *valore* denso di imprescindibili interpretazioni economiche ed un valore, reinquadrato in una cornice squisitamente pedagogica, va ad inserirsi la riflessione sulle competenze di Data Analytics necessarie alle università per proseguire in un percorso di miglioramento continuo: all'*analytics* va ad aggiungersi il *learning* e il compito diviene tanto audace quanto affascinante.

2. Il progetto pilota presso l'area informatica dell'Università di Parma

Si presenta di seguito un progetto di monitoraggio delle carriere nato all'interno del Corso di Studi triennale in Informatica (CdS L3027) presso l'Università degli studi di Parma⁵. Il progetto ha avuto avvio nel 2018 da una prima collaborazione tra l'area amministrativa della UO Controllo di Gestione UNIPR e il CdS di Informatica al fine monitorare l'andamento degli indicatori MIUR AA1 e D21⁶. In tale occasione furono analizzati 43 corsi di studio triennali e preso in considerazione gli insegnamenti del primo anno di ciascun CdS che permettevano il superamento dei 40 CFU. In questo primo processo il totale di studenti analizzati fu di 4348 casi e di questi 412 studenti risultavano sotto la soglia dei 40 CFU per un solo esame da sostenere o verbalizzare. Con l'introduzione di una sessione di appello straordinaria si riuscì dunque a raggiungere l'obiettivo previsto. La collaborazione sugli indicatori MIUR ha successivamente indotto il CdS di Informatica ad avviare alcune riflessioni in merito all'utilizzo dei dati per analizzare i livelli di performance dei CdS, giungendo ad avviare, nel 2019, un progetto pi-



5 Referente scientifico del processo di Data Mining Prof. Alessandro Dal Palù, Presidente del CdS di Informatica, Dipartimento di Scienze Matematiche, Fisiche e Informatiche presso UNIPR.

6 Il riferimento è agli indicatori contenuti nella Programmazione Triennale 2016-2018, di cui al D.M. 635/2016. Nello specifico l'indicatore AA1 si riferisce a "Proporzione di studenti che si iscrivono al II anno della stessa classe di laurea o laurea magistrale a ciclo unico, avendo acquisito almeno 40 CFU in rapporto alla coorte di immatricolati nell'a.a. precedente." mentre il D21 si riferisce a "Proporzione di studenti iscritti entro la durata normale del corso di studi che abbiano acquisito almeno 40 CFU nell'anno solare".

lota sui LA. Attualmente il progetto pilota prevede la sinergia tra il Dipartimento di Scienze Matematiche, Fisiche e Informatiche (tramite il Presidente del Cds in Informatica), il Presidio di Qualità di Ateneo, la U.O Controllo di gestione e l'area Pedagogica tramite la collaborazione di un docente di Pedagogia Sperimentale.

Il progetto pilota nasce con l'obiettivo di rilevare, nei tempi più rapidi possibili, le sorgenti di ritardo nelle carriere studenti al fine di poter attivare diverse linee di intervento, tra cui scelte strategiche in merito a: 1) un potenziamento delle azioni di orientamento in itinere maggiormente calibrate sul tipo di problematica e non distribuite "a pioggia"; 2) l'eventuale riprogettazione del piano di studi dei CdS oggetto di analisi e la pianificazione delle eventuali propedeuticità tra insegnamenti; 3) l'individuazione precoce di insegnamenti "complessi/problematici" che oggi in molti atenei rischia di non essere oggetto di analisi, o in altri casi, di essere presa in carico basandosi sui dati raccolti tramite lo strumento OPIS-ROS⁷ (strumento per sua stessa natura percettivo e pertanto non utilizzabile per inferire dati sugli esiti dell'apprendimento); 4) la progettazione di seminari di formazione docenti specifici in merito a competenze didattiche e docimologiche, a partire da dati reali sulle carriere studenti e contestualizzati per settori disciplinari.

Il progetto prevede inoltre un'analisi delle dipendenze dei tempi di sostenimento degli esami, nella carriera del singolo studente e il controllo di correlazione tra voto / giorni per il conseguimento / n. di tentativi, anche al fine di caratterizzare l'andamento di una coorte.

Il Presidio della Qualità ha stabilito inoltre che nel corso del prossimo anno accademico il progetto pilota raccolga dati di tre diversi CdS appartenenti a tre aree formative tra loro molto diverse: un CdS di area umanistica; un CdS di area tecnico-scientifica; un CdS di area Medica. Il fine è quello di monitorare e tarare lo strumento grazie ai learning analytics su corsi di studi differenti, sia per tipo di insegnamenti, sia per architettura interna del CdS, sia per modalità di selezione studenti in ingresso.

Nelle sezioni che seguono si presentano alcuni estratti delle analisi pilota svolte sulla "demo"⁸; le videate riportate nelle figure sottostanti possono essere esemplificative delle estrazioni che il sistema permette, utilizzando tre soli fattori facilmente monitorabili ed inseriti su una ti-

7 Si vedano anche i recenti lavori sulla revisione dello strumento di rilevazione dell'opinione studenti <https://www.anvur.it/attivita/ava/opinioni-studenti/> che tuttavia non ne modificano l'impianto di natura percettiva.

8 La fase "Demo" attualmente ha nel data base i dati di due CdS di Informatica e di Matematica.



me line: la data di fine frequenza dell'insegnamento; la data di prima iscrizione ad un appello; la data di verbalizzazione dell'insegnamento (uguale o successiva alla prima iscrizione ad un appello).



Fig. 1: Estratto del processo temporale tra la fine dell'erogazione di un insegnamento e la verbalizzazione dell'esito, a seguito di superamento della prova

2.1 Storia di uno studente

L'estrazione "storia di uno studente" (Fig.2) permette in una videata di ricostruire la carriera di uno studente ad esempio in un triennio di iscrizione al corso di studio. Sulla ascissa vi sono gli insegnamenti previsti nel piano di studi, sulle ordinate vi sono le sessioni di appello previste nel triennio. La linea continua nera orizzontale indica il tempo intercorrente tra la fine dell'erogazione dell'insegnamento e la prima iscrizione ad un appello (Fig.1). La linea in verticale di colore rosso indica il momento di iscrizione all'appello e verbalizzazione coincidente. La barra in orizzontale di colore rosso indica invece il tempo intercorrente tra una prima iscrizione ad un appello e un periodo successivo di "preparazione" a seguito di fallimento (ritirato, bocciato o assente).

La videata riporta un caso non particolarmente problematico dove risultano evidenti due "fallimenti" su ventuno insegnamenti. Il primo "fallimento", sull'insegnamento di fisica, si è risolto molto velocemente, probabilmente in un altro appello della medesima sessione; la risoluzione del secondo fallimento non è invece ancora avvenuta nonostante siano passati circa due anni dal primo tentativo. Nel frattempo, tuttavia lo studente risulta attivo, si iscrive ad altri 10 appelli d'esame e li supera. Cosa stia accadendo nel processo di apprendimento dello studente "x" con l'insegnamento "Algoritmi e Strutture e dati II" non si può ovviamente comprendere da quanto offerto dai LA. L'estratto mette in luce, a colpo d'occhio, un elemento di possibile rallentamento della carriera studente che un buon percorso di tutorato potrebbe approfondire, partendo proprio da tale elemento di difficoltà, ancora irrisolto ma ben circoscritto.



In fig. 3 è riportata la storia dello stesso studente ma con l'indicazione non solo del caso di difficoltà di cui sopra, ma anche l'evidenza di un insegnamento che ha portato nello studente un più lento percorso di pianificazione sull'insegnamento "Fondamenti dell'Informatica". Il tempo intercorrente tra la fine dell'insegnamento e la verbalizzazione è di quasi due anni. Nel frattempo, lo studente sostiene altri 6 esami al primo tentativo. Se la situazione fosse analoga nelle carriere di diversi altri studenti iscritti al CdS, si potrebbe ipotizzare una necessaria riprogettazione del corso di studio con lo spostamento dell'insegnamento ad un anno successivo o con la verifica di eventuali propeuticità implicite nella preparazione dell'insegnamento (ad esempio "Intelligenza Artificiale" o "Reti di calcolatori" che precedono la verbalizzazione dell'insegnamento a lungo pianificato).

La fig. 4 riporta ancora una volta la storia dello studente "x" ma a cui si aggiunge la tracciatura dei periodi di massima intensità di impegno dello studente con la barra di intensità di colore viola; quando la tonalità tende a scurirsi sta ad indicare un periodo "di carico", particolarmente intenso per lo studente. La barra è realizzata sovrapponendo i periodi di studio per tutti gli esami durante i relativi periodi tra frequenza e sostenimento.

L'analisi delle sfumature della barra di intensità permette di comprendere quando un singolo caso è stato accademicamente più attivo e quando meno e permette inoltre di comprendere un percorso di apprendimento a carico continuo e costante o a "picchi d'impegno". Si può notare come i picchi siano concentrati durante le sessioni di esame, mentre tra le sessioni l'intensità rimane più scura nei casi in cui vi siano esami in sospenso.

In questo caso, ad esempio, si vede un carico particolarmente "leggero" tra il luglio 2016 e il febbraio 2017. Il periodo agosto 2017 e marzo 2018 invece è particolarmente intenso. Ancora una volta, il dato letto sul singolo studente potrebbe non dire molto; lo stesso dato di periodi "leggeri/sovraccarichi" rilevato su un numero consistente di studenti potrebbe invece condurre ad una miglior pianificazione degli insegnamenti o eventualmente la messa a controllo di una relazione tra una certa preparazione iniziale dello studente (tipo di diploma) e il tipo di impegno richiesto nel proseguo degli studi.



Storia di uno studente

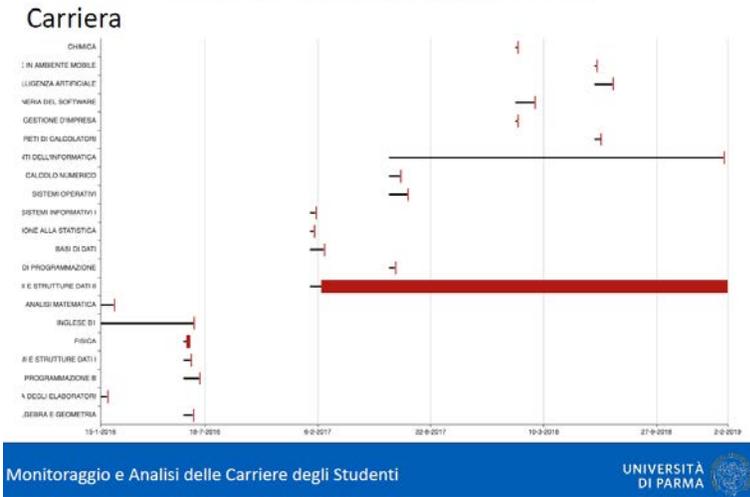


Fig. 2: La videata “storia di uno studente” permette di ricostruire la carriera di uno studente in un triennio di iscrizione al corso di studio



Storia di uno studente

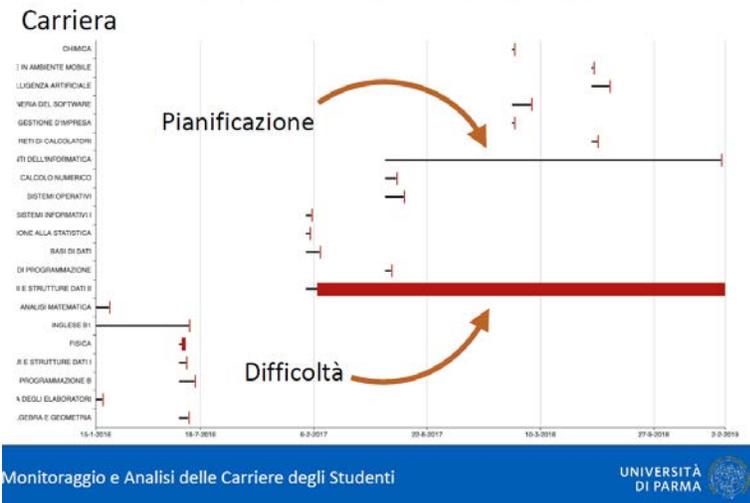
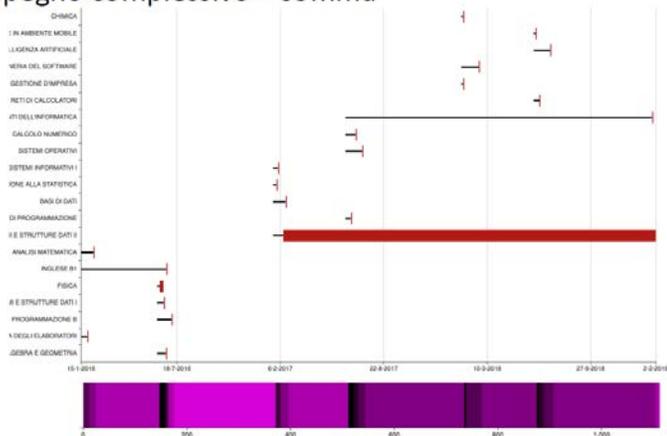


Fig. 3: La videata storia di uno studente permette di ricostruire la carriera di uno studente sottolineandone i periodi di difficoltà e pianificazione

Storia di uno studente

Impegno complessivo = somma



Monitoraggio e Analisi delle Carriere degli Studenti UNIVERSITÀ DI PARMA

Fig. 4: Monitoraggio della carriera di uno studente con la barra di intensità dell'impegno richiesto sul periodo. Le sezioni più scure della barra indicano i periodi maggiormente intensi per l'attività accademica dello studente

2.2 Storia di un insegnamento

Utilizzando sempre gli stessi tre fattori (frequenza, prima iscrizione, verbalizzazione), è possibile invece analizzare le performance degli studenti relativamente ad uno stesso insegnamento (Fig. 5). La fig. 6 mostra un estratto delle performance di diverse matricole verso un unico insegnamento, evidenziando molti casi di insuccesso dopo il primo tentativo. È possibile controllare l'andamento di più matricole (riportate sull'ascisse) e verificare come alcuni studenti abbiano in "cantier" tale insegnamento da più di 4 anni. In taluni casi, ove non si evidenziasse nessuna iscrizione a questo o ad altri appelli di altri insegnamenti, si arriverebbe ad ipotizzare perfino un abbandono universitario.

Uno spaccato di questo tipo, per molti versi allarmante, dovrebbe condurre all'attivazione di alcuni interrogativi interni al CdS. Il presidente del CdS è al corrente di tale situazione? I tutor hanno attivato un percorso sulle matricole a rischio di abbandono? Il docente dell'insegnamento ha avuto modo di analizzare questi dati? Sono state messe in atto delle azioni di monitoraggio degli interventi didattici o di verifica della coerenza tra quanto erogato e quanto oggetto di valutazione sommativa? L'insegnamento è "disciplinarmente" coerente/affine con gli altri insegnamenti del CdS? Gli studenti hanno i prerequisiti necessari per poterne affrontare la preparazione?

È proprio in questi casi che i LA, messi a disposizione nel più breve tempo possibile, divengono essenziali per un'analisi complessiva di un corso di studio, delle sue difficoltà e dei suoi punti di forza. Anche in casi allarmanti di questo tipo, certamente già noti alla popolazione interna al CdS, se non suffragati da dati di un certo peso e di una certa costanza, potrebbero rischiare di rimanere "alert" inevasi o affrontati con interventi di tutorato non calibrati su un intervento specificatamente indirizzato all'insegnamento oggetto di difficoltà.

Storia di un esame

- Quante volte lo studente si è iscritto (=tentativi)
- Causa (spesso non disponibile): assente, ritiro, rifiuta voto

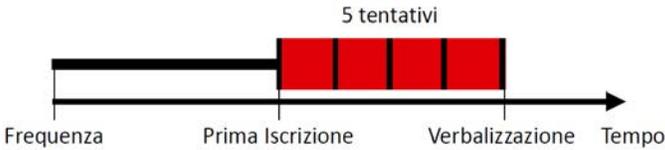


Fig. 5: Estratto della storia di un esame, inserito su una linea del tempo

Storia di un insegnamento

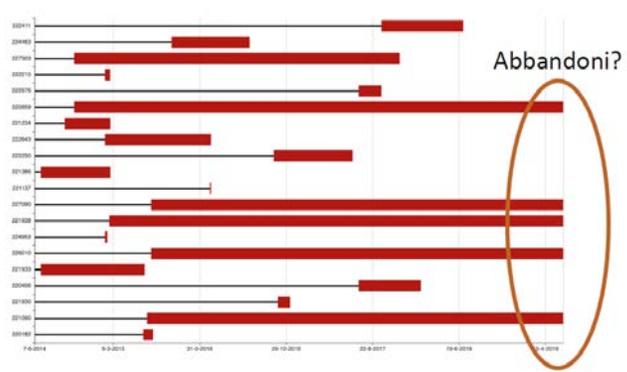
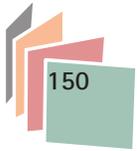


Fig. 6: La videata storia di un insegnamento, con l'analisi delle performance delle matricole, portata ad individuare un alto numero di casi che non riesce a superare tale insegnamento anche a distanza di anni

2.3 *Lo studente lungo un percorso di apprendimento. Analisi delle coorti.*

Prendendo in considerazione altri fattori è possibile analizzare con prospettive differenti le caratteristiche della popolazione studentesca ed eventualmente calibrare o progettare diverse attività di orientamento in itinere. I LA a disposizione degli atenei potrebbero prevedere infatti la possibilità di analisi di correlazione e di dipendenze statisticamente rilevanti tra coppie di esami. In particolare, si vogliono analizzare i casi in cui il sostenimento di un esame risulta ritardato dal sostenimento di un altro e i casi in cui una coppia di esami vengono conseguiti in ordine inverso rispetto al rispettivo ottenimento della frequenza. Inoltre, tali indagini si possono applicare alla carriera del singolo studente tramite il controllo di correlazione tra voto/giorni per il conseguimento/ n. di tentativi, anche al fine di caratterizzare l'andamento di una coorte. Il confronto tra i dati di diverse coorti, al variare ad esempio delle modalità di orientamento e/o tutorato (situazione abbastanza frequente nella vita di un CdS), permetterebbe inoltre al CdS di verificare se le modifiche introdotte si siano dimostrate efficaci o meno e, in caso affermativo, in quale misura (Voti più alti? Minor tempo impiegato per la preparazione degli esami? Minori fallimenti al primo tentativo?). La fig. 7 mostra il confronto tra due casi della stessa coorte ad uno stesso insegnamento; entrambi sostengono con successo l'esame di un insegnamento al primo tentativo, con talune differenze. Il caso a sinistra esemplifica la performance di uno studente che si iscrive ad un appello subito dopo averne terminato la frequenza e supera l'appello con un voto sufficiente. Il caso a destra invece illustra la performance di un altro studente iscritto allo stesso appello, ravvicinato alla fine del corso, ma che supera in modo eccellente. Quali sono le differenze tra i due studenti? Il loro profilo in ingresso è simile? L'ateneo ha al proprio interno numerosi dati anagrafici e curricolari per poter controllare diverse ipotesi sulla dipendenza tra fattori. Un'analisi di tal tipo svolta ad esempio su una coorte potrebbe mettere in evidenza, fin dalle prime sessioni di esame, le caratteristiche della popolazione studentesca che proseguirà nel CdS.



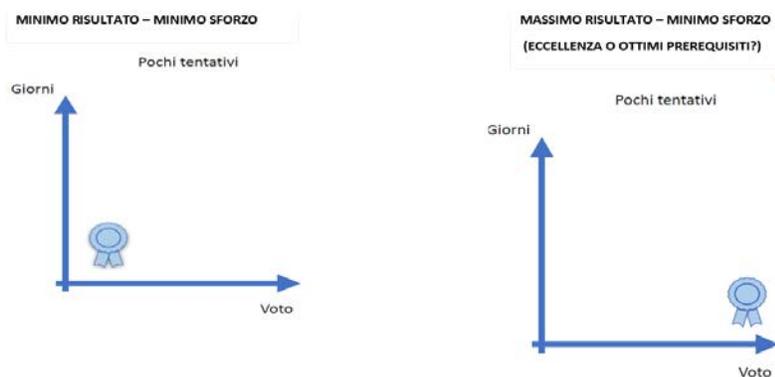


Fig. 7: Confronto tra due studenti ad uno stesso appello posto subito dopo il termine dell'erogazione dell'insegnamento

La fig. 8 invece riporta un estratto di analisi su uno stesso cluster di studenti all'interno di una coorte, alle prese con due insegnamenti differenti. Nell'insegnamento a sinistra i casi si distribuiscono in modo abbastanza eterogeneo per giorni impiegati e voto conseguito. Vi è un numero consistente di studenti che hanno raggiunto un buon voto, spostato verso i valori alti della scala in trentesimi. Nell'insegnamento a sinistra invece tutti gli studenti non si presentano ai primi appelli ma attendono quelli successivi, ci si sposta infatti sul numero maggiore di giorni impiegati per la preparazione e i voti conseguiti hanno una bassa distribuzione di frequenza ed un'alta concentrazione sui valori centrali della scala, senza che nessuno raggiunga i valori alti. Una situazione di questo tipo potrebbe significare un tempo di preparazione all'esame superiore, potrebbe indicare un insegnamento che preveda più prove prima della verbalizzazione finale, potrebbe ancora derivare da una propedeuticità tra insegnamenti non formalmente riconosciuta. Un'analisi di secondo livello, anche tramite i controlli incrociati con la "storia dello studente" potrebbe fornire numerose risposte per *spiegare* le differenze tra i due insegnamenti. Gli stessi percorsi di formazione docenti che vengono erogati da quasi tutti gli atenei italiani, potrebbero approfondire anche un diverso utilizzo di processi docimologici, talvolta ancorati a tradizioni di settore o dettati da esigenze specifiche dei settori disciplinari. L'estratto in fig. 8 mette in luce senz'altro un diverso utilizzo della scala ad intervalli 0-30 tra i due insegnamenti. Se il dato venisse confermato nel raffronto tra diverse coorti, allora si potrebbe lecitamente escludere che tale distribuzione dipenda solo dalle caratteristiche degli studenti valutati.

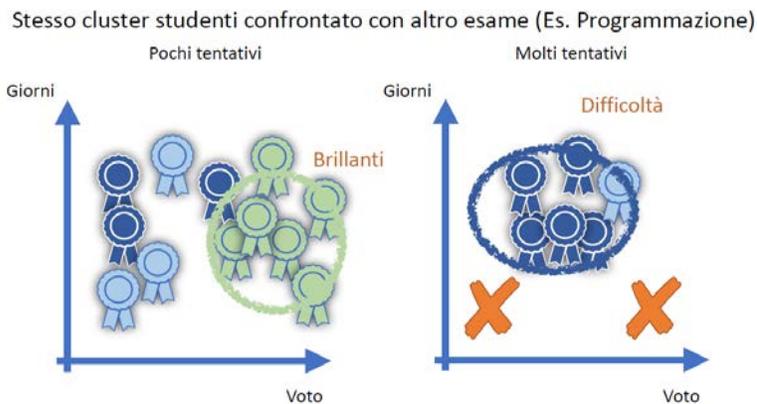


Fig. 8: Estratto di LA su uno stesso cluster di studenti all'interno di una coorte, alle prese con due insegnamenti differenti

3. Conclusioni e sviluppi futuri nei processi autovalutativi

Il progetto pilota nato in seno all'Università di Parma rappresenta una tra le tante e infinite modalità di utilizzare i LA per rispondere a problemi conoscitivi prevalentemente pedagogici. Le potenzialità sono evidenti: l'analisi sistemica di alcuni fattori ed una restituzione "user friendly", non solo per gli addetti ai lavori, possono restituire alla popolazione accademica (nelle sue tre diverse anime: docenti; studenti e PTA) una base empirica notevole ma al contempo ordinata per prendere delle decisioni informate e consapevoli relativamente alla progettazione dei percorsi di apprendimento, di orientamento e di tutorato.

Diversi fattori guidano l'espansione emergente dei Learning Analytics: a) l'ampliamento e la proliferazione dell'utilizzo di Internet e della tecnologia in tutti gli ordini educativi-formativi; b) l'abbondanza di dati disponibili all'interno degli ambienti di apprendimento virtuali c) la disponibilità di strumenti che possono essere utilizzati per gestire e analizzare i dati; d) la crescente domanda di comprendere gli studenti e migliorare l'ambiente di apprendimento e il suo contesto (Khalil & Ebner, 2015); non è pertanto immaginabile che proprio gli atenei, con le loro decine di migliaia di studenti, non riflettano nel brevissimo periodo su come ottimizzare al meglio i dati in loro possesso e come riorientarne l'utilizzo verso il core business.

Di contro le università, come anche gli altri ordini di istruzione e formazione, stanno già da tempo prendendo le misure anche con la dimensione etica e legale del trattamento del dato che potrebbe porre un freno non irrilevante ad utilizzo massivo dei dati sensibili (Prinsloo & Slade, 2013; Fung, Wang, Chen, 2010) messi a disposizione dei diversi portatori di interesse, chiamati in causa nel complesso processo formativo.

Alcuni autori (Arnold, Pistilli, 2012; Bakharia, Dawson, 2011; Jackson, Read, 2012; Khalil, Ebner, 2015) hanno infatti messo in evidenza come un utilizzo sistemico dei LA possa essere di interesse per diversi stakeholders: studenti, docenti, mondo della ricerca e decision makers (Tab.1). Gli obiettivi delle diverse categorie di stakeholders sono necessariamente differenti tra loro, ma accomunati dal desiderio di riflettere criticamente sul proprio operato, al fine di migliorarsi anche grazie ai dati raccolti nel processo autovalutativo.

Stakeholder	Obiettivi dell'utilizzo dei LA	Esempi e riferimenti a studi condotti sulla popolazione di stakeholders
Studenti	Migliorare le loro prestazioni. Personalizzare l'apprendimento (online). Consigliare o ricevere consigli sui corsi da frequentare.	Gli studenti sono informati sul loro processo di apprendimento e confrontano le loro prestazioni con quelle degli altri. Nello studio viene utilizzata l'applicazione Signals (Arnold, Pistilli, 2012; Vozniuk, 2013).
Docenti	Migliorare i loro processi di insegnamento. Potere fornire feedback agli studenti in tempi molto rapidi.	I docenti procedono ad un monitoraggio dei progressi nell'apprendimento dei propri studenti (Miranda, Marzano, Lytras, Miltiades, 2017). In uno degli studi viene utilizzata l'applicazione SNAPP (Bakharia, Dawson, 2011; Vozniuk, 2013)
Ricercatori	Valutare i corsi. Migliorare i modelli didattici proposti dai corsi. Individuare nuove modalità per veicolare informazioni circa i processi educativi/formativi.	Attraverso diverse interfacce e processi di analisi, i ricercatori possono confrontare le tecniche di Analytics relative ai processi di apprendimento. (Castro, Vellido, Nebot, Mugica, 2007; Ferguson, 2012; Chatti, Dyckhoff, Schroeder, Thüs, 2012; Khalil, Ebner, 2015).
Decision Maker accademici	Supportare la presa di decisioni per raggiungere obiettivi formativi sempre più alti e intervenire prontamente sugli studenti a rischio di abbandono universitario	Il monitoraggio su alcune università australiane rivela che un terzo degli studenti che inizia un percorso di studi non giunge alla laurea. Nello studio si fa riferimento ad un utilizzo combinato di indicatori di performance (Program Connect for Success C4S) basato sui LA e volto a contrastare il fenomeno dell'abbandono universitario. (Connect for Success C4S) (Jackson, Read, 2012) (Gaeta, Marzano, Miranda, Sandkuhl, 2016).

Tab. 1: Learning Analytics Stakeholders
(a partire da Khalil, M. & Ebner, M. 2015; rivista e aggiornata dall'autrice)

I LA, anche nel sistema accademico italiano, possono così divenire un valido strumento per avviare processi autovalutativi e piani di miglioramento effettivamente orientati ed ancorati a dati sui processi di apprendimento e insegnamento. Nel frangente in cui le università italiane iniziano a riflettere e collaborare sinergicamente⁹ per l'individuazione dei core contents disciplinari e trasversali, comuni a medesime classi di laurea, una riflessione collettiva sulla base di selezionati LA potrebbe meglio supportare il processo di individuazione di core contents e di riflessione critica sull'effettivo raggiungimento degli obiettivi for-

mativi prefissati sulle diverse fasce della popolazione studentesca.

L'ampia letteratura sul tema ha già messo in luce le diverse potenzialità dei LA: da quelle predittive (Romero, Ventura, 2010); d'intervento e d'orientamento (Duval, 2011) fino a quelle di benchmarking (Vorhies, Morgan, 2005) e di individuazione di best practices trasferibili a diversi contesti. Sta alla flessibilità del mondo pedagogico di trovare spazi di ricerca comuni, certamente transdisciplinari e al servizio del miglioramento continuo dei propri atenei.

Riferimenti bibliografici

- 
- Arnold K.E., Pistilli M.D. (2012). Course Signals at Purdue: Using learning analytics to increase student success. In *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge (LAK'12)* (pp. 267-270). New York, USA: ACM.
- Bakharia A., Dawson S. (2011). SNAPP: a bird's-eye view of temporal participant interaction. In *Proceedings of the 1st international conference on learning analytics and knowledge (LAK'11)* (pp. 168-173). New York, USA: ACM.
- Castro F., Vellido A., Nebot A., Mugica F. (2007). Applying data mining techniques to e-learning problems. In L. C. Jain, T. Raymond, D. Tedman (Eds.), *Evolution of teaching and learning paradigms in intelligent environment* (Vol. 62, pp. 183-221). Berlin: Springer-Verlag.
- Chatti M.A., Dyckhoff A.L., Schroeder U., Thüs H. (2012). A reference model for learning analytics. *International Journal of Technology Enhanced Learning*, 4(5/6), 318-331. doi:10.1504/IJTEL.2012.051815
- Dawson S., Gašević D., Siemens G., Joksimovic S. (2014). Current state and future trends: a citation network analysis of the learning analytics field. In *Proceedings of the Fourth International Conference on Learning Analytics & Knowledge (LAK'14)* (pp. 231-240). New York, USA: ACM.
- Duval E. (2011). Attention please! Learning analytics for visualization and recommendation. In *Proceedings of the 1st International Conference on Learning Analytics and Knowledge (LAK'11)* (pp. 9-17). New York, USA: ACM.
- Ebner M., Schön M. (2013). Why Learning Analytics in Primary Education Matters. *Bulletin of the Technical Committee on Learning Technology*, Karagiannidis, C. & Graf, S (Ed.), 15(2), 14-17.
- Gaeta M., Marzano A., Miranda S., Sandkuhl K. (2016). The competence management to improve the learning engagement. *Journal of Ambient Intelligence and Humanized Computing*, 1-13, Springer.
- Ferguson R. (2012). Learning analytics: drivers, developments and challenges. *International Journal of Technology Enhanced Learning*, 4(5), 304-317.
- Ferguson R., Buckingham Shum S. (2012). *Social learning analytics: five approaches*. Proceedings of the 2nd International Conference on Learning Analytics and

9 Si vedano i gruppi di lavoro multidisciplinari per l'individuazione dei core contents delle aree sanitarie, filosofiche, pedagogiche e la costruzione delle prove pilota TECO-D e TECO-T. <https://www.anvur.it/gruppi-di-lavoro/?taxonomy=attivita&term=teco>

- Knowledge, (pp. 23-33). Vancouver, British Columbia, Canada.
- Ferguson R. (2014), Learning Analytics: fattori trainanti, sviluppi e sfide. *TD Tecnologie didattiche*, 22 (3), 138-147
- Fung B., Wang K., Chen R., Yu P. S. (2010). Privacy-preserving data publishing: A survey of recent developments. *ACM Computing Surveys (CSUR)*.
- Jackson G., Read M. (2012). *Connect 4 Success: A Proactive Student Identifications and Support Program*. Retrieved 12 November 2014 from http://fyhe.com.au/past_papers/papers12/Papers/9B.pdf
- Manyika J., Chui M., Brown B., Bughin J., Dobbs R., Roxburgh C., Byers A. H. (2011). *Big data: The next frontier for innovation, competition, and productivity*.
- Miranda S., Marzano A., Lytras Miltiades D. (2017). A research initiative on the construction of innovative environments for teaching and learning. Montessori and Munari based psycho-pedagogical insights in computers and human behavior for the “new school”. *Computers in Human Behavior*, 66, 282-290.
- Mostow J., Beck J. (2006). Some useful tactics to modify, map and mine data from intelligent tutors. *Natural Language Engineering*, 12(02), 195-208.
- Prinsloo P., Slade S. (2013). An evaluation of policy frameworks for addressing ethical considerations in learning analytics. In *Proceedings of the Third International Conference on Learning Analytics and Knowledge (LAK 2013)* (pp. 240-244). New York, USA: ACM.
- Romero C., Ventura S. (2010). Educational data mining: a review of the state of the art. *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*, 40(6), 601-618.
- Siemen, G. (2010). What are Learning Analytics? Retrieved 7 November 2014. Estratto da <http://www.elearnspace.org/blog/2010/08/25/what-are-learning-analytics/>.
- Siemens G. (2012). Learning Analytics: Envisioning a Research Discipline and a Domain of Practice. In *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge (LAK 2012)* (pp. 04-08). New York, USA: ACM.
- Slade S., Sharon, & Prinsloo, P. (2013). Learning analytics: ethical issues and dilemmas. *American Behavioral Scientist*, 57(10) 1509-1528.
- Society for Learning Analytics Research. (2011). *Open Learning Analytics: an integrated & modularized platform*. Retrieved 01 November 2014. Estratto da: <http://solaresearch.org/OpenLearningAnalytics.pdf>.
- Vorhies D.W., Morgan N.A. (2005). Benchmarking marketing capabilities for sustainable competitive advantage. *Journal of Marketing*, 69(1), 80-94. Originally published in: Khalil, M. & Ebner, M. (2015). Learning Analytics: Principles and Constraints. In *Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications* (pp. 1326-1336). Chesapeake, VA: AACE.
- Vozniuk A., Govaerts S., Gillet D. (2013). Towards portable learning analytics dashboards. In *Advanced Learning Technologies (ICALT), 2013 IEEE 13th International Conference on* (pp. 412-416). IEEE.

View publication

- Siemens G. (2013). Learning Analytics: The Emergence of a Discipline. *American Behavioural Scientist*, 57 (10), pp. 1380-1400, Sage Publication.
- Norris D. B. (2008). Action analytics: measuring and improving performance that matters in higher education. *EDUCAUSE Review*, 43(1).
- van Barneveld A., Arnold K. E., Campbell J.P. (2012). Analytics in Higher Education: Establishing a Common Language. *EDUCAUSE*.



Applicazione del machine learning ai learning analytics della piattaforma Moodle per creare gruppi eterogenei nei corsi on-line

Application of machine learning to the learning analytics of the Moodle platform to create heterogeneous groups in on-line courses

Giacomo Nalli

School of Science and Technology, Computer Science section, University of Camerino
giacomo.nalli@unicam.it

Leonardo Mostarda

School of Science and Technology, Computer Science section, University of Camerino
leonardo.mostarda@unicam.it

Andrea Perali

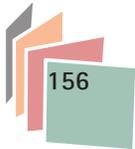
School of Pharmaceutical Sciences and Health Production, University of Camerino
andrea.perali@unicam.it

Sebastiano Pilati

School of Science and Technology, Physics Section, University of Camerino
sebastiano.pilati@unicam.it

Daniela Amendola

School of Biosciences and Veterinary Medicine, University of Camerino
daniela.amendola@unicam.it



In university courses to promote collaborative activities among students, on-line learning environments such as e-learning platforms are used. Effective collaborative activities involve the creation of heterogeneous groups of 4 or 5 students. In the university context the formation of groups is difficult due to the high number of students. Groups are often unbalanced and not very functional if chosen randomly. Some e-learning platforms, such as Moodle, lack an intelligent mechanism that allows the automatic creation of heterogeneous groups of students. We applied clustering algorithms on Moodle learning analytics (LA) that allowed to build groupings that identify the different characteristics of students based on their behaviors kept on the platform. Therefore we have developed an intelligent numerical tool which, using clusters obtained from Machine Learning on the LA, generates heterogeneous groups. These groups are made available on the platform for the teacher. The project will conclude with the development of a Moodle plugin to automate the exchange of data and information between the Machine Learning algorithm and the Moodle platform.

Keywords: Learning Analytics; Machine learning; Moodle; Clustering; Gruppi

Nei percorsi universitari, per favorire le attività collaborative tra gli studenti, vengono utilizzati ambienti di apprendimento on-line come le piattaforme e-learning. Attività collaborative efficaci prevedono la creazione di gruppi eterogenei di 4 o 5 studenti. Nel contesto universitario la formazione dei gruppi è difficile per l'elevato numero di studenti. Se scelti in maniera casuale, spesso i gruppi risultano sbilanciati e poco funzionali. Alcune piattaforme e-learning, ad esempio Moodle, mancano di un meccanismo "intelligente" che permetta di creare in automatico gruppi eterogenei di studenti. Il nostro lavoro consiste nel realizzare un software in Python in grado di creare gruppi eterogenei di studenti, utilizzando tecniche di Machine Learning con i dati estratti da Moodle. Abbiamo applicato

algoritmi di clustering sui learning analytics (LA) di Moodle che hanno permesso di costruire dei raggruppamenti che identificano le caratteristiche degli studenti in base ai loro comportamenti in piattaforma. Abbiamo quindi sviluppato uno strumento numerico “intelligente” che, utilizzando i cluster ottenuti dal Machine Learning sui LA, genera gruppi eterogenei. Questi gruppi vengono messi a disposizione in piattaforma per il docente. Il progetto si concluderà con lo sviluppo di un plugin di Moodle per automatizzare lo scambio di dati e informazioni tra l’algoritmo di Machine Learning e la piattaforma Moodle.

Parole chiave: Learning Analytic; Machine learning; Moodle; Clustering; groups

1. Introduzione

Le attività collaborative sono una metodologia didattica estremamente significativa per gli studenti poiché, promuovendo un apprendimento attivo (Foote, 2009), migliorano i risultati dell’apprendimento oltre che sviluppare le loro abilità sociali come il processo decisionale, la comunicazione, le capacità collaborative ed il pensiero critico (Smith and MacGregor, 2009; Stevens, Levy, 2005; Amendola, Miceli, 2018). A causa del gran numero di studenti iscritti nei percorsi universitari, talvolta si possono riscontrare difficoltà a livello organizzativo nella progettazione di attività collaborative. Oggi grazie allo sviluppo delle tecnologie digitali è possibile organizzare esperienze di apprendimento collaborativo on-line in maniera più flessibile, sia per gli studenti che per i docenti (Abedin et al., 2012). Nei contesti formativi universitari, per favorire le attività collaborative tra gli studenti, vengono spesso utilizzati ambienti di apprendimento on-line come le piattaforme e-learning che, come ambienti web 2.0, enfatizzano la partecipazione, la connessione, la condivisione di conoscenze ed idee tra gli studenti grazie a strumenti appositamente progettati (McLough, 2007).

La piattaforma e-learning Moodle (Modular Object-Oriented Dynamic Learning Environment, ambiente per l’apprendimento modulare, dinamico, orientato ad oggetti), molto utilizzata dalle università italiane e straniere, offre diversi strumenti per stimolare la partecipazione, l’interazione, la negoziazione e la collaborazione tra studenti, come ad esempio il forum, il wiki ed il workshop. Tali strumenti permettono al docente di poter effettuare e gestire le attività in maniera semplice e automatizzata.

Sia in ambiente on-line che in presenza, uno dei fattori fondamentali che può influenzare il successo dell’apprendimento collaborativo è la composizione di gruppi in base al numero ed all’eterogeneità dei componenti. In riferimento al numero dei componenti, gruppi accettabili vengono considerati quelli formati da almeno 3 o più persone, anche se in generale gruppi da 4 o 5 membri risultano essere più efficaci (Burke, 2011).



I gruppi possono essere formati anche da due persone, ma questo numero non incoraggia il lavoro di gruppo perché non è sufficiente per generare creatività e varietà di idee (Csernica et al., 2002). Un'altra caratteristica fondamentale per il successo nei lavori di gruppo è sicuramente l'eterogeneità degli studenti in termini di risorse cognitive, caratteristiche e comportamenti (Nijstad, De Dreu, 2002).

Tuttavia formare gruppi ottimali di studenti per attività collaborative non è semplice. Normalmente si utilizzano diversi approcci che però non garantiscono sempre la formazione di gruppi eterogenei: selezione casuale, selezione automatica e selezione del docente (Sadeghi, Kardan, 2016). Nel primo caso il docente assegna gli studenti casualmente ai gruppi, manualmente o tramite un sistema informatico. Questo è il metodo più semplice e veloce che consente di mescolare tutti gli studenti con la speranza di raggiungere l'eterogeneità all'interno dei gruppi (Bacon et al., 2001), fondamentale per il miglioramento delle performance sociali e cognitive, soprattutto per studenti meno motivati.

La formazione del gruppo può essere condotta anche attraverso l'approccio di auto-selezione da parte degli studenti che consente di avere gruppi con un'elevata empatia, ma non sempre eterogenei sotto il profilo pedagogico.

Il terzo approccio, che potrebbe garantire la formazione di gruppi eterogenei, è la selezione da parte del docente sulla base di caratteristiche prestabilite, come per esempio conoscenza, abilità, interessi e stile di apprendimento (Jackson et al., 1995).

Quest'ultimo metodo, che risulta essere il più funzionale per la realizzazione di gruppi eterogenei, in ambiente universitario è di difficile attuazione. Per il docente universitario, infatti, l'identificazione degli studenti che frequentano l'aula, in base a determinate caratteristiche e comportamenti, risulta essere complicata non solo per l'elevato numero di partecipanti ma anche in quanto i corsi hanno una durata relativamente breve e non sempre hanno frequenza obbligatoria.

Nei percorsi universitari, dove per le attività didattiche vengono utilizzate anche le piattaforme e-learning, l'elaborazione dei Learning Analytics (LA) prodotti può essere utile per creare gruppi eterogenei di studenti sulla base delle loro caratteristiche e comportamenti in piattaforma.

Il punto di debolezza è che in questi ambienti on-line manca ancora un meccanismo intelligente che permetta di creare questi gruppi in modo automatico, facilitando in tal modo il lavoro del docente.

Negli ultimi anni alcuni gruppi di ricerca hanno iniziato a lavorare su progetti di Machine Learning applicati ai LA estratti dalle piattaforme e-learning per creare gruppi eterogenei in modo automatico: al-



cuni autori si sono basati sui dati estratti da un questionario on-line (Razmerita, Brun, 2011), altri sui dati estratti dall'interazione degli studenti all'interno dello strumento forum di Moodle (Maina et al., 2017). I lavori presenti in letteratura per la formazione di gruppi eterogenei on-line si basano principalmente sull'elaborazione di LA estratti da un'unica attività svolta dagli studenti nella piattaforma e-learning.

Dutt et al. (2017) nel loro lavoro passano in rassegna la letteratura esistente sul clustering dei dati dai processi di apprendimento. In questa rassegna la possibilità di realizzare una classificazione di studenti in base ai loro stili di apprendimento viene presentata come applicazione di successo delle tecniche di clustering dei dati provenienti dai processi di apprendimento. Un'importante conclusione ribadita da Dutt et al. è che questi dati sono spesso non-indipendenti e con una struttura gerarchica multi-livello e quindi la scelta delle variabili per la separazione dei cluster assume un ruolo chiave e deve essere effettuata dal ricercatore in modo molto scrupoloso per ottenere risultati coerenti. Una sezione della rassegna di Dutt et al. è dedicata alle tecniche di clustering applicate ai dati provenienti da corsi erogati in e-learning, com'è il caso del nostro lavoro. Si evidenzia che numerosi articoli sono presenti in letteratura su questo argomento in quanto i corsi e-learning forniscono molti dati da poter utilizzare grazie agli strumenti di tracciamento delle piattaforme e-learning. Interessante sottolineare applicazioni delle suddette tecniche allo studio del problem-solving per individuare in modo automatico gli stili di risoluzione e di apprendimento degli studenti. La sezione di questa rassegna dedicata al collaborative learning evidenzia l'efficacia dell'algoritmo K-means per la formazione di gruppi e come gli ambienti e-learning siano i più adatti per gli studenti che collaborano attivamente. Conclusioni queste che sono in linea con i metodi adottati e i risultati ottenuti nel nostro lavoro.

Infine, un importante confronto tra i LA ed il cosiddetto Educational Data Mining (EDM) viene discusso da Papamitsiou & Economides (2014). Per quanto riguarda l'approccio per acquisire informazioni sui processi di apprendimento, LA adotta un quadro olistico, cercando di comprendere i sistemi nella loro piena complessità. D'altra parte, EDM adotta un punto di vista riduzionista analizzando i singoli componenti del processo di apprendimento, cercando nuovi modelli nei dati e modifica dei rispettivi algoritmi. Queste due ricerche sono complementari ed al fine di catturare l'intero quadro dei processi di apprendimento degli studenti, entrambi gli approcci andrebbero perseguiti.

Il nostro progetto di ricerca consiste nell'elaborare insieme i LA estratti dalle molteplici attività che gli studenti svolgono durante un percorso on-line, sfruttando al meglio le correlazioni presenti nei dati.



Lo scopo è quello di ottenere una visione complessiva dei comportamenti degli studenti per meglio caratterizzarli, permettendoci così di creare dei gruppi quanto più possibili eterogenei. Il nostro approccio consiste nella creazione di un'applicazione informatica che permetta la realizzazione di gruppi eterogenei in modo automatico, usando tecniche di Machine Learning non supervisionato (Kotsiantis, 2007) applicate ai LA prodotti dagli studenti durante la fruizione di un percorso on-line in Moodle ed elaborati tramite il linguaggio di programmazione Python. La scelta iniziale dei LA da utilizzare viene fatta dal Docente, tramite caselle di spunta visualizzate dalla piattaforma e-learning Moodle, nel primo step del processo. Lo scopo è di poter selezionare da subito i dati che corrispondono ai criteri generali in base ai quali il Docente intende comporre i gruppi. La scelta dei LA fatta dal Docente includerà in tal modo la sua idea didattica e costituirà il set iniziale sul quale andrà ad operare il software di clusterizzazione automatica realizzato e descritto in questo articolo. Al Docente rimarrà pertanto un controllo a monte del processo di formazione automatica dei gruppi, con possibilità di variare le sue scelte iniziali dalle quali far partire, o ripartire, il software e di analizzarne le relazioni di causa-effetto così ottenute, in primis la stabilità dei risultati, ovvero la composizione dei gruppi, al variare delle condizioni iniziali.

In questo articolo, come corso pilota per il nostro progetto di ricerca è stato scelto un percorso laboratoriale on-line di fisica per il primo anno del corso di laurea in lingua inglese in Biosciences and Biotechnology, organizzato in una parte individuale ed una parte collaborativa che richiede la formazione di gruppi di lavoro.

Il nostro progetto di ricerca si sviluppa in due fasi:

1. applicazione di algoritmi di clustering ai Learning Analytics di Moodle estratti alla fine del percorso didattico individuale per la creazione di raggruppamenti omogenei al loro interno basati sulle caratteristiche e comportamenti in piattaforma simili degli studenti. Per verificare l'efficienza delle tecniche di clustering per la formazione dei raggruppamenti omogenei di studenti, sono stati messi a confronto le caratteristiche dei diversi raggruppamenti con gli esiti di un elaborato svolto alla fine del percorso individuale;
2. sviluppo di un software intelligente che, utilizzando i raggruppamenti ottenuti, distribuisce automaticamente gli studenti degli stessi raggruppamenti in diversi gruppi che risultano in questo modo eterogenei al loro interno. Dopo la formazione dei gruppi eterogenei il software comunica al docente automaticamente i gruppi creati.

2. Aspetti Metodologici

2.1. *Descrizione dell'attività formativa*

Il laboratorio online di fisica è stato erogato tramite la piattaforma e-learning Moodle di Ateneo e vi hanno partecipato 55 studenti internazionali (19 maschi e 36 femmine).

Il percorso didattico è strutturato in due parti. Una prima parte individuale, della durata di un mese, caratterizzata da 5 video esperimenti, realizzati nei nostri laboratori, relativi alla forza elastica e all'oscillatore armonico. Ogni video esperimento dura in media 10 minuti per un totale di 50 minuti.

I video esperimenti on-line consentono agli studenti di esaminare i fenomeni elastici, raccogliendo ed elaborando dati sperimentali utilizzando il software gratuito Gnuplot, al fine di ottenere un fit dei dati e trovare la formulazione matematica della legge fisica corrispondente. Per l'utilizzo del software sono disponibili in piattaforma due video tutorial: i) come installare Gnuplot; ii) come utilizzare Gnuplot per creare un grafico, preparare figure ed elaborare i dati. I video tutorial durano in totale 15 minuti. Sono presenti, inoltre, altri files (pdf e pagine web) con testo ed immagini, che contengono ulteriori spiegazioni relative a Gnuplot ed alla parte teorica dei video esperimenti.

La parte individuale si conclude con la realizzazione di un elaborato finale, da caricare in piattaforma attraverso il modulo consegna di Moodle, dove lo studente deve rispondere ad un set di domande e risolvere esercizi guidati dai video tutorial e dai video esperimenti. L'elaborato viene poi valutato dal docente.

La seconda parte, caratterizzata dall'attività collaborativa, consiste nella realizzazione di un report finale sugli argomenti proposti nella prima parte del percorso, da realizzare in gruppo. Per rendere l'attività collaborativa più efficace in questa fase del percorso vengono creati gruppi eterogenei formati da 4 o 5 studenti ognuno, come illustrato di seguito.

2.2. *Machine Learning*

Per la realizzazione dei gruppi abbiamo applicato tecniche di Machine Learning non supervisionato ai Learning Analytics prodotti dagli studenti in piattaforma durante lo svolgimento del percorso individuale. Il Machine Learning è un insieme di tecniche sviluppate nel campo dell'intelligenza artificiale che include al suo interno vari modelli sta-



tistici complessi e i corrispondenti metodi di ottimizzazione. L'obiettivo è quello di costruire algoritmi che possano estrarre informazione utile da grosse moli di dati a disposizione ed individuare delle correlazioni tra di essi, fornendo all'utilizzatore un modello in grado di effettuare predizioni accurate su contesti nuovi. Spesso, ma non sempre, questi modelli sono costruiti tramite reti neurali artificiali.

Nell'ambito della didattica on-line, tecniche di Machine Learning possono essere usate per diversi scopi, come per esempio informare i docenti sull'andamento del corso (Feng, Heffernan, 2005) o nella predizione del voto finale dello studente (Lopèz et al., 2012).

Tra le tecniche di Machine Learning, definite di apprendimento non supervisionato, troviamo in particolare gli algoritmi di clustering.

Questa tecnica permette di individuare un gruppo di oggetti che hanno caratteristiche simili, secondo criteri definiti a priori, e di assegnarli ad un cluster o raggruppamento.

Negli ambienti di apprendimento on-line, il clustering può essere usato per trovare raggruppamenti di studenti con simili caratteristiche, relative, ad esempio, ai livelli di apprendimento (Bovo et al., 2013).

Sono disponibili diversi algoritmi di clustering che consentono di formare gruppi di studenti sulla base del loro comportamento durante la fruizione di un corso online in piattaforma e-learning. Questi algoritmi risultano essere utilizzati per diverse finalità: per esempio l'algoritmo "Self Organizing Maps" permette di raggruppare gli studenti in base al loro background, oppure l'algoritmo "Fuzzy c-means" che permette di raggruppare studenti in base alla loro personalità e strategia di apprendimento (Vellido et al., 2010). La nostra scelta è ricaduta sull'algoritmo di clustering K-Means perché oltre ad essere un algoritmo molto utilizzato e di facile uso grazie alla libreria scikit learn che ne fornisce l'implementazione, è un algoritmo basato sull'utilizzo della distanza euclidea per la computazione che assegna un determinato elemento (in questo caso lo studente) a uno specifico cluster. Queste tipologie di algoritmo risultano essere ottimali per il raggruppamento degli studenti in base al comportamento relativo alla navigazione online effettuata tra le varie attività all'interno di un corso e-learning (Vellido et al.) e per organizzare in cluster i differenti comportamenti degli studenti tenuti nella piattaforma e-learning (Dutt et al., 2017).

Nel nostro lavoro per la realizzazione dei gruppi eterogenei vengono applicate prima tecniche di clustering, utilizzando l'algoritmo K-means, per organizzare dei raggruppamenti di studenti con caratteristiche simili (comportamento durante la fruizione del percorso on-line) ricavate dall'elaborazione dei LA forniti dalla piattaforma Moodle.



Successivamente, un software intelligente, distribuisce automaticamente gli studenti degli stessi raggruppamenti in diversi gruppi che risultano in questo modo eterogenei al loro interno.

3. Formazione di gruppi eterogenei

In questa sezione descriviamo la metodologia utilizzata per la raccolta e l'elaborazione dei dati sul comportamento degli studenti estratti dalla piattaforma e-learning al fine di costituire gruppi di studenti al loro interno eterogenei per l'attività collaborativa prevista nella seconda parte del percorso laboratoriale di Fisica.

Numerosi studi in letteratura ci forniscono un quadro teorico di riferimento per il nostro lavoro. Possiamo riassumere lo stato dell'arte osservando che i learning analytics inerenti il comportamento degli studenti in piattaforma hanno una grande utilità in ambito didattico. Ad esempio la frequenza dei login ad una piattaforma e-learning può essere utilizzata per la predizione del voto finale di uno studente (Froissard et al., 2015). In Jo ed altri (2014), viene dimostrato un modo efficace per utilizzare i dati relativi ai log degli utenti, come il tempo totale online e la frequenza dei login, come indicatori predittivi delle prestazioni di apprendimento. In Macfadyen (2010), viene dimostrato inoltre come l'uso di dati di tracciamento dello studente come, ad esempio, tempo speso online, login effettuati, file visti, web link e caricamento di esercizi in piattaforma (consegne sottomesse), sia correlato con il voto finale dello studente.

In Fig. 1 è rappresentato il flusso delle attività on-line, della ricerca, dello sviluppo del software e della creazione dei gruppi eterogenei.

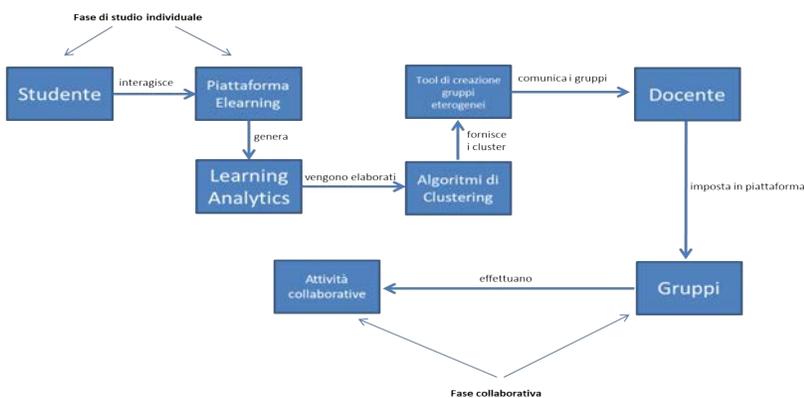


Fig. 1: Diagramma a blocchi del flusso delle attività, ricerca e sviluppo del software



Il processo per la formazione dei gruppi eterogenei ha richiesto due fasi di lavoro distinte:

- a) implementazione di tecniche di clustering per la formazione di raggruppamenti omogenei in termini di caratteristiche e comportamenti simili degli studenti durante la fruizione del percorso individuale on-line;
- b) sviluppo di un software automatico in grado di selezionare gli studenti dai diversi raggruppamenti omogenei per la formazione di gruppi eterogenei.

3.1. *Tecniche di clustering per la formazione di raggruppamenti omogenei*

Per la formazione di raggruppamenti omogenei abbiamo:

1. estratto e selezionato i learning analytics dalla piattaforma Moodle da utilizzare per la preparazione di un dataset da sottoporre alle tecniche di clustering;
2. implementato le tecniche di clustering ai LA per la realizzazione di raggruppamenti omogenei di studenti con caratteristiche e comportamenti simili tra loro.

Durante il primo step del progetto di ricerca abbiamo estratto, selezionandoli tra quelli disponibili, tutti i Learning Analytics relativi al comportamento degli studenti prodotti in piattaforma dopo la fruizione del percorso individuale online: frequenza login, ultimo login effettuato, tempo totale speso on-line, numero di video tutorial visti, frequenza video tutorial visti, numero video esperimenti visti, frequenza video esperimenti visti, numero pagine web viste, numero file pdf scaricati, numero di esercizi svolti. Per semplificare il file dei dati li abbiamo raggruppati in features, ovvero dati appartenenti allo stesso tipo di attività (per es. visualizzazione dei 5 video esperimenti) sono stati aggregati in una stessa feature. La feature, in italiano caratteristica, è una proprietà individuale e misurabile di un fenomeno osservato (Bishop, 2006).

Per la caratterizzazione degli studenti e la creazione dei raggruppamenti omogenei abbiamo utilizzato solo dati quantitativi, non prendendo in considerazione la valutazione degli elaborati da parte del docente per evitare che il peso del voto potesse influenzare il processo di suddivisione degli studenti. Le valutazioni degli elaborati individuali sono invece state utilizzate successivamente per verificare l'efficacia delle



tecniche di clustering. Abbiamo infatti messo a confronto le valutazioni con i cluster (raggruppamenti) ottenuti per vedere se esistesse una correlazione tra di essi, in modo da verificare se il comportamento degli studenti influisca o meno sulle loro performance finali.

Prima di applicare l'algoritmo per il clustering, i dati raccolti sono stati pre-processati. Ogni studente è stato rappresentato da un "vettore di input" con features costituite dai valori degli attributi associati allo studente.

Tutti i dati, organizzati in vettori di features, uno per ogni studente, sono stati inseriti all'interno di un unico file Excel, chiamato dataset, che rappresenta il file di input del nostro software per la generazione dei cluster. Per avere la stessa scala di valori i dati sono stati normalizzati in un intervallo che va da 0 a 1.

Successivamente siamo passati alla realizzazione del software per la creazione dei raggruppamenti omogenei utilizzando il linguaggio di programmazione Python poiché presenta una grande quantità di tool (librerie) per l'elaborazione di dati usando algoritmi di Machine Learning. Inizialmente il software prende in input il dataset e lo processa per determinare il numero di raggruppamenti da creare. Questo numero deve essere fornito all'algoritmo di clustering. Per generare i raggruppamenti è stato utilizzato l'algoritmo di clustering K-means, implementato attraverso una funzione della libreria scikit learn di Python (Hackeling, 2014). L'algoritmo utilizza un numero K di cluster (da impostare prima dell'esecuzione), per il quale si vuole suddividere il dataset e ad ogni cluster assegna un centroide, che deve rappresentare il punto centrale di ogni cluster.

Per determinare il numero adeguato di cluster è stata sviluppata una funzione in Python, utilizzando l'"Elbow Method" (Thorndike, 1953; Hackeling, 2014), un metodo di interpretazione all'interno dell'analisi del cluster finalizzato a trovare il numero appropriato di cluster in un set di dati. Il metodo permette di generare un grafico, avente nell'asse delle ascisse un intervallo di valori di K (ovvero il numero di cluster in cui viene suddiviso il dataset) e sull'asse delle ordinate la somma delle distanze dei dati osservati dai centroidi dei cluster, chiamato Within-Cluster-Sum-of-Squares (WCSS). Il numero di K (valore presente nell'asse delle ascisse) dove, nel grafico, la diminuzione del valore di WCSS all'aumentare di K comporta un calo significativo della velocità di incremento, viene chiamato "elbow". Tale valore rappresenta il numero ottimale di cluster da realizzare in base al dataset fornito.

Il passo successivo è l'esecuzione dell'algoritmo di clustering. L'algoritmo prende il numero K di cluster ottenuto dall'elbow method per suddividere il dataset e assegna ad ogni cluster un centroide. La scelta



del centroide, nella prima iterazione, avviene in maniera casuale. La vicinanza di un vettore (set di dati relative a un certo studente) al centroide di un cluster stabilisce l'appartenenza di tale vettore a quel determinato cluster. L'algoritmo calcola la distanza euclidea tra ogni vettore x e ogni centroide, assegnando il vettore al centroide c per il quale la distanza risulti minima:

$$c = \operatorname{argmin}_{c_i \in C} \operatorname{dist}(c_i, x)^2$$

dove “ c_i ” rappresenta un centroide dell'insieme C (insieme di centroidi), x rappresentano i vettori di input, mentre “ dist ” è la distanza euclidea standard. In seguito, vengono ricalcolati i valori dei centroidi. Il nuovo valore di un centroide sarà la media di tutti i vettori che sono stati assegnati al cluster del centroide. L'esecuzione continua per iterazione (un numero massimo di iterazioni vengono definite a priori per evitare un ciclo infinito) e termina quando:

- nessun vettore del dataset è soggetto a cambiamento di cluster;
- la somma delle distanze viene ridotta al minimo;
- viene raggiunto un numero massimo di iterazioni.

Alla fine dell'esecuzione vengono generati i raggruppamenti di studenti che risultano avere dei comportamenti e caratteristiche simili in piattaforma. In output viene generato il grafico con la rappresentazione dei vettori associati a diversi cluster in un piano bidimensionale di due tra le 10 features utilizzate ed il testo contenente il nome del raggruppamento, seguito da una lista di numeri, in cui ogni numero identifica ciascun studente.

Infine sono stati utilizzati i voti dell'elaborato individuale per poter vedere se risultano esserci delle somiglianze tra studenti, appartenenti ad uno stesso raggruppamento, non solo a livello di comportamento ma anche a livello di performance, per garantire una maggiore eterogeneità nella formazione dei gruppi eterogenei.

L'esecuzione del software continua con la formazione dei gruppi eterogenei, attraverso un algoritmo che preleva studenti da cluster diversi e li riorganizza in gruppi eterogenei.

3.2. *Formazione gruppi eterogenei*

Dopo la realizzazione dei cluster, l'elaborazione continua con l'esecuzione di una funzione Python basata su un nuovo algoritmo, sviluppato



appositamente, per suddividere in maniera automatica e uniforme gli studenti appartenenti a diversi cluster in diversi gruppi eterogenei.

Ogni raggruppamento di studenti ottenuto dal cluster rappresenta i diversi livelli di partecipazione alle attività in piattaforma e identifica tipologie di comportamento diverse per ogni cluster.

L'algoritmo prevede la selezione e l'inserimento all'interno di ogni gruppo di almeno uno studente appartenente a una tipologia diversa di cluster, in modo da garantire che in ogni gruppo siano presenti studenti con diverse caratteristiche e comportamento, in modo da migliorare così le performance degli studenti nelle attività collaborative.

Inizialmente la funzione calcola il numero di gruppi da creare in base al numero di studenti da inserire nel gruppo, che deve essere impostata a priori. Poi ordina le liste dei cluster in ordine crescente, dalla lista meno numerosa alla più numerosa e calcola, in base alle lunghezze, quanti studenti per ogni cluster inserire in un gruppo.

Successivamente l'algoritmo seleziona gli studenti in base al proprio ID, prelevando uno o più valori alla volta da ogni cluster (in base ai calcoli effettuati considerando le lunghezze dei cluster) in maniera sequenziale e li salva in un'unica nuova lista nell'ordine in cui essi sono stati prelevati dai diversi cluster. La lista ottenuta viene poi suddivisa in sotto liste attraverso la definizione di un intervallo che rappresenta il numero di partecipanti che si vuole ottenere all'interno di ogni gruppo.

In questo modo avremo la certezza che almeno un componente di ogni cluster (a patto che il numero di componenti dei gruppi non sia inferiore al numero di cluster) farà parte di ogni sotto lista (che rappresenta ogni singolo gruppo), ottenendo così al suo interno eterogeneità tra gli studenti.

Conclusa questa fase, una e-mail viene inviata direttamente al docente con la lista dei gruppi eterogenei contenenti al loro interno gli ID associati agli studenti.

4. Risultati e discussioni

Per la fase di test, è stato necessario creare il dataset, un file Excel da fornire in input al software, che presenta al suo interno diversi dati relativi al comportamento degli studenti. Sulla base del dataset fornito, il software, applicando l'algoritmo dell'"elbow method", ha restituito in output il relativo grafico dal quale è stato ottenuto il numero ideale di Cluster da generare, pari a 3 (Fig. 2).



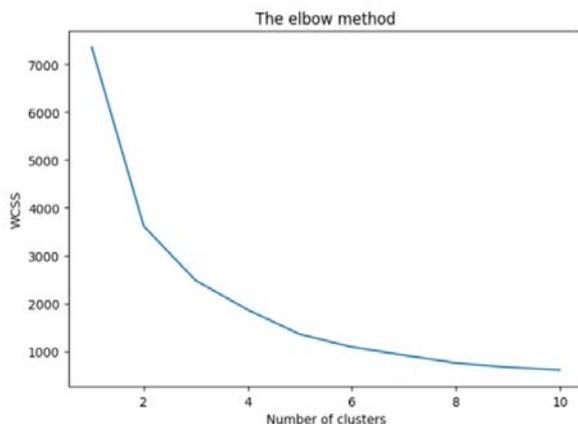


Fig. 2: Grafico elbow method



Successivamente, è stato eseguito l'algoritmo di clustering K-means che ha generato 3 raggruppamenti omogenei, caratterizzati rispettivamente da 10 (cluster 0 in rosso), 29 (cluster 1 in blu) e 16 (cluster 2 in verde) studenti (Fig. 3)

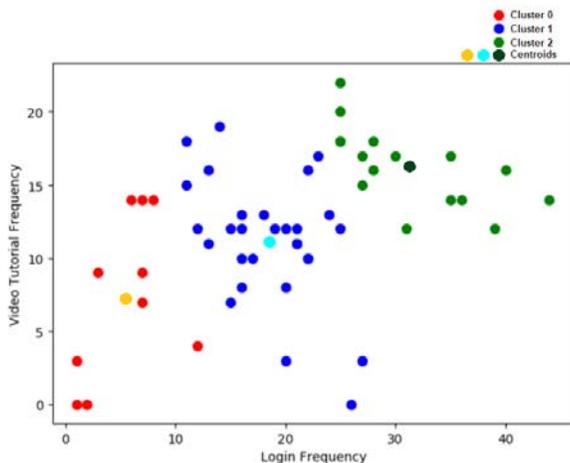


Fig. 3: Rappresentazione di una proiezione bidimensionale (frequenze di visualizzazione dei video tutoriali, frequenza di login) dei raggruppamenti emersi dalla fase di clustering in base alle diverse features selezionate

I cluster ottenuti sono stati poi analizzati per determinare i diversi comportamenti in piattaforma, evidenziando quali features risultano essere più influenti all'interno dei raggruppamenti e in tal modo profilare gli studenti che appartengono a un determinato raggruppamento. I risultati sono i seguenti. Come si evince dalla Tab. 1, il cluster 0 rap-

presenta gli studenti poco attivi in piattaforma con un basso numero di accessi (media = 5,4), bassa frequenza di visualizzazioni relativi ai video tutorial (media = 7,4), e uno scarso numero e frequenza di visualizzazioni relativa agli esperimenti (media = 2,6). La scarsa partecipazione si riflette anche sulla consegna dell'elaborato individuale, dove solo il 40% degli studenti lo ha consegnato in piattaforma. Il cluster 1 rappresenta gli studenti mediamente attivi in piattaforma, in cui gli utenti hanno ottenuto dei valori medi in quasi tutte le attività: media del numero di accessi pari a 18,4, media del numero di video tutorial visti uguale a 2,5, media della frequenza di visualizzazioni relativi ai video tutorial uguale a 11,2, media del numero di esperimenti visti uguale a 4,1, media della frequenza di visualizzazione dei video esperimenti uguale a 6,7. In questo caso la partecipazione ha consentito al 90% degli studenti di realizzare e consegnare l'esercizio finale. Il cluster 2 invece rappresenta gli studenti molto attivi in piattaforma con una media del numero di accessi di 31,2, media del numero di video tutorial visti uguale a 3, con una frequenza media di visualizzazioni pari a 16,2. La media del numero di esperimenti visti è di 4,3 (5 video esperimenti in totale) mentre la frequenza media di visualizzazioni è di 7. Questo livello di partecipazione ha consentito a tutti gli studenti di consegnare l'elaborato finale (100% degli studenti).



	cluster 0	cluster 1	cluster 2
	Studenti poco attivi in piattaforma	Studenti mediamente attivi in piattaforma	Studenti molto attivi in piattaforma
n° studenti	10	29	16
n° accessi in piattaforma (media)	5,4	18,4	31,2
n° video tutorial visti (media)	1,8	2,5	3
frequenza visualizzazione totale dei video tutorial (media)	7,4	11,2	16,2
n° video esperimenti visti (media)	2	4,1	4,3
frequenza visualizzazione totale dei video esperimenti (media)	2,6	6,7	7
% studenti che hanno svolto l'elaborato	40%	90%	100%

Tab. 1: Differenze tra cluster e dettagli delle features analizzate

Per verificare se il comportamento svolto dagli studenti nella piattaforma e-learning potesse avere un riscontro in termini di performance, abbiamo confrontato ogni cluster con la media dei voti relativi all'elaborato individuale svolto dagli studenti appartenenti al medesimo cluster. I risultati ottenuti sono molto interessanti. Come riportato in Fig. 4, il livello di partecipazione in piattaforma degli studenti rispecchia pienamente il voto finale (voto in centesimi). Gli studenti più attivi (cluster 2), infatti, hanno ricevuto un punteggio più alto rispetto agli altri con una media dei voti pari ad 82,5. Gli studenti mediamente attivi (cluster 1) risultano invece avere una votazione inferiore con una media del 72,04, mentre quelli del cluster 0 una media dei voti addirittura al di sotto della sufficienza, pari a 53. Questa forte correlazione evidenziata dall'indice di Pearson pari a 0,986 ci ha permesso di confermare ulteriormente le differenze esistenti tra gli studenti appartenenti ai diversi cluster, non solo più in termini di comportamento ma anche di performance.

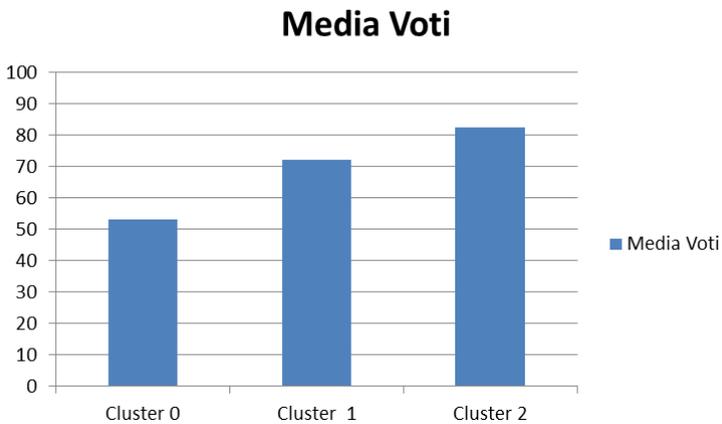


Fig. 4: Medie dei voti (intervallo 0-100) degli studenti nei diversi cluster

Una volta ottenuti i cluster, l'esecuzione del software ha restituito la creazione dei gruppi eterogenei. A priori abbiamo impostato il numero di studenti per gruppo pari a 5. Dalla Fig. 5 possiamo osservare che una volta impostato l'algoritmo, il software ha creato 11 gruppi e distribuito in automatico gli studenti in modo che in ogni gruppo fosse presente almeno uno studente appartenente a cluster differenti, creando quindi gruppi eterogenei sulla base dei comportamenti degli studenti.

```

cluster 0:
[4, 8, 12, 22, 25, 29, 30, 34, 43, 46]
cluster 1:
[0, 1, 2, 3, 5, 6, 7, 10, 11, 14, 15, 21, 23, 24, 26, 27, 28, 31, 33, 35, 36, 38, 39, 40, 41, 45, 47, 51, 52]
cluster 2:
[9, 13, 16, 17, 18, 19, 20, 32, 37, 42, 44, 48, 49, 50, 53, 54]

Gruppo 0[4, 9, 0, 1, 2]
Gruppo 1[0, 13, 3, 5, 6]
Gruppo 2[12, 16, 7, 10, 11]
Gruppo 3[22, 17, 14, 15, 21]
Gruppo 4[25, 18, 23, 24, 26]
Gruppo 5[29, 19, 27, 28, 31]
Gruppo 6[30, 20, 33, 35, 36]
Gruppo 7[34, 32, 38, 39, 40]
Gruppo 8[43, 37, 41, 45, 47]
Gruppo 9[46, 42, 44, 51, 52]
Gruppo 10[44, 48, 49, 50, 53]

```

Fig. 5. Cluster e gruppi eterogenei ottenuti dall'esecuzione del software

Infine il software ha inviato automaticamente una e-mail al docente, con la lista dei gruppi eterogenei, indicando il nome del gruppo (es. gruppo 0, gruppo 1, etc.) e accanto ad ogni gruppo la lista di 5 numeri, che sono gli ID che identificano gli studenti. In tal modo il docente ha potuto selezionare gli studenti in base al proprio ID, inserirli nei relativi gruppi all'interno di Moodle e iniziare la parte collaborativa del corso.

4. Conclusioni e sviluppi futuri

In questo lavoro abbiamo ideato e realizzato un nuovo software finalizzato ad aiutare i docenti nella definizione di gruppi di studenti destinati ad attività collaborative. La possibilità di utilizzare tecniche di Machine Learning ci ha permesso di dimostrare come queste contribuiscano al miglioramento della formazione di gruppi eterogenei, ma anche di testare l'efficienza dell'algoritmo K-means, analizzando per ogni cluster le features e delineando diversi profili di studente. Il confronto tra i cluster ottenuti e la media del voto finale degli studenti di ogni cluster, ha confermato le differenze tra i cluster anche in termini di performance. Grazie all'esecuzione di queste tecniche sui dati raccolti e le verifiche sull'efficienza delle tecniche, è stato possibile garantire e massimizzare l'eterogeneità degli studenti all'interno di ogni gruppo, accrescendo così le possibilità di successo per tutti gli studenti nello svolgimento di lavori di gruppo all'interno del corso. La scelta delle features è ricaduta sui Learning Analytics della piattaforma e-learning Moodle, per permettere al software di essere riutilizzato da altri docenti e tutor che svolgono attività didattiche on-line su Moodle. La scelta dell'ambiente Moodle non risulta comunque essere vincolante per il funzionamento del software, in quanto è facilmente adattabile a qualsiasi altro LMS (Learning Management System), attraverso la modifica



del dataset. Il software realizzato per il presente lavoro richiede attualmente, per l'esecuzione, un supporto tecnico ai docenti, soprattutto nella realizzazione del dataset e per la scelta del numero di cluster da creare. Il progetto di ricerca proseguirà con la realizzazione di un Plugin avanzato per Moodle che permetterà al docente in modo semplice, intuitivo e automatizzato di creare gruppi di studenti eterogenei, attraverso la sola selezione, tramite delle caselle di spunta, dei Learning Analytics di Moodle che il docente stesso riterrà maggiormente rilevanti per la profilazione delle caratteristiche degli studenti.

Riferimenti bibliografici

- 
- 172
- Abedin B., Daneshgar F., D'Ambra J. (2012). Pattern of non-task interactions in asynchronous computer-supported collaborative learning courses. *Interactive Learning Environments*, 22 (1), 1-17.
- Amendola D., Miceli C. (2018). Online peer assessment to improve students' learning outcomes and soft skills. *Italian Journal of Educational Technology*, 26(3), 71-84.
- Bacon R.D., Stewart K.A., Anderson E.S. (2002). Methods of Assigning Players to Teams: A Review and Novel Approach. *Simulation & Gaming*, 32(1), 6-17.
- Bishop C. M. (2006). *Pattern recognition and machine learning*. Berlin: Springer.
- Bovo A., Sanchez S., Héguy O., Duthen Y. (2013). "Clustering Moodle data as a tool for profiling students". The 2nd International Conference on e-Learning and e-Technologies in Education - ICEEE (pp. 121-126).
- Burke A. (2011). Group Work. *Journal of Effective Teaching*, 11(2), 87-95.
- Csernica J., Hanyka M., Hyde D., Shooter S., Toole M., Vigeant M. (2002). *Practical guide to teamwork, version 1.1*. Lewisburg: College of Engineering, Bucknell University.
- Dutt A., Ismail M. A., Herawan T. (2017). A systematic review on educational data mining. *IEEE Access*, 5, 15991-16005.
- Feng M., Heffernan N.T. (2005). *Informing teachers live about student learning: Reporting in the assistent system*. The 12th Annual Conference on Artificial Intelligence in Education Workshop on Usage Analysis in Learning Systems. Amsterdam: IOS Press.
- Foote E. (2009). *Collaborative Learning in Community College*. ERIC. Estratto da <http://www.ericdigests.org/1998-1/colleges.htm>.
- Froissard C., Richards D., Atif A., Liu D.Y. (2015). *An enhanced learning analytics plugin for Moodle: student engagement and personalised intervention*. Proceedings of the ASCILITE Conference, (pp. 180-189).
- Hackeling G. (2014). *Mastering Machine Learning with scikit learn*. Birmingham: Packt Publishing
- Jackson S.E., May K.E., Whitney K. (1995). Understanding the dynamics of diversity in decision making teams. In Guzzo & Salas (Eds), pp. 204-261.

- Jo I., Kim D., Yoon M. (2014). *Analyzing the log patterns of adult learners in LMS using learning analytics*. Proceedings of the Fourth International Conference on Learning Analytics And Knowledge (pp.183-187).
- Kotsiantis S.B. (2007). *Supervised Machine Learning: A Review of Classification Techniques*. Proceedings of the Conference on Emerging Artificial Intelligence Applications in Computer Engineering: Real Word AI Systems with Applications in eHealth, HCI, Information Retrieval and Pervasive Technologies, (pp. 3-24).
- Lopèz M., Luna J., Romero C., Ventura S. (2012). *Classification via clustering for predicting final mark based on student participation in forums*. Proceedings in the International Conference on Information Technology Systems and Innovation (ICITSI), (pp.148-151).
- Macfadyen L.P., Dawson S. (2010). Mining LMS data to develop an "early warning system" for educators: A proof of concept. *Computer and Education*, 54, 588-599.
- Maina E.M., Oboko R.O., Waiganjo P.W. (2017). Using Machine Learning Techniques to Support Group Formation in an Online Collaborative Learning Environment. *International Journal of Intelligent Systems and Applications*, 3, 26-33.
- McLoughlin C., J.W. Lee M. (2007). *Social software and participatory learning: Pedagogical choices with technology affordances in the Web 2.0 era*. Proceedings in Ascilite, Australian Society for Computers in Learning in Tertiary Education Annual Conference (pp. 664-675).
- Nijstad B.A., De Dreu C. K.W. (2002). Creativity and Group Innovation. *Applied Psychology: An International Review*, 51 (3), 400-406.
- Papamitsiou Z., Economides A. (2014). Learning Analytics and Educational Data Mining in Practice: A Systematic Literature Review of Empirical Evidence. *Educational Technology & Society*, 17, 49-64.
- Razmerita L., Brun A. (2011). *Collaborative learning in heterogeneous classes. Towards a Group Formation Methodology*. Proceedings of 3rd International Conference on Computer Supported Education (CSEDU 2011), 2, (pp. 189-194).
- Sadeghi H., Kardan A.A. (2016). Toward effective group formation in computer-supported collaborative learning. *Interactive Learning Environments*, 24(3), 382-395
- Smith B.L., MacGregor J.T. (2009). *What is collaborative learning? National Center on Postsecondary Teaching*. Pennsylvania State University Press.
- Stevens D.D., Levi A.J. (2005). *Introduction to rubrics: An assessment tool to save grading time, convey effective feedback and promote student learning*. Stylus Publishing.
- Thorndike R.L. (1953). Who Belongs in the Family? *Psychometrika*, 18, 267-276.
- Vellido A., Castro F., Nebot A. (2010). Clustering educational data. *Handbook of educational data mining*, 75-92.



Social Network Learning Analytics: identification of students at risk of early school leaving

Social Network Learning Analytics: identificazione degli studenti a rischio di abbandono scolastico

Caterina Bembich

Dipartimento di Studi Umanistici, Università degli studi di Trieste, cbembich@units.it

Social Network Analysis (SNA) is gaining increasing attention in educational research as branch of Learning Analytics. This contribution offers an overview of Social Network Analysis, describing the origins of this methodology, the implications of its use for educational research, and its application in the study of early school leaving.

This paper presents a case study of the application of SNA in a group of students attending a vocational school. The study analyzes the external and internal relationships of the students in the school. The results highlight that students at risk of dropping out of school, tend to have less dense and less cohesive social networks, and exhibit a higher probability of establishing relationships with classmates with similar tendencies to leave school early. SNA draws attention to the relational structures that are established during school activities, and can help teachers and schools improve learning processes and learning environments more broadly. This study shows how a relational approach can be used to explore the phenomenon of early school leaving and highlights dysfunctional relational structures that accentuate the risk situation.

Keywords: Social Network Analysis; Learning Analytics; Early school leaving; Interdependence; Identification of risk

La Social Network Analysis (SNA) sta attualmente emergendo nella ricerca educativa, come un settore dei Learning Analytics. La prima parte del contributo descrive le origini della Social Network Analysis, le implicazioni nella ricerca educativa e la sua applicazione nello studio del fenomeno dell'abbandono scolastico.

Il lavoro presenta successivamente un caso studio di applicazione della SNA. Lo studio analizza le reti di relazioni, sia esterne che interne alla scuola, costruite da un gruppo di studenti frequentanti una scuola professionale. Dai risultati ottenuti si evidenzia come studenti a rischio di abbandono scolastico tendono ad avere reti sociali meno dense e meno coese, e a mostrare una probabilità più alta di stringere rapporti con compagni di classe con la loro stessa tendenza all'abbandono. Lo studio evidenzia come un approccio relazionale può essere utilizzato per esplorare il fenomeno dell'abbandono scolastico e mettere in luce strutture relazionali disfunzionali che accentuano la situazione di rischio. L'analisi delle reti sociali può aiutare gli insegnanti e la scuola a migliorare i processi di apprendimento attraverso la strutturazione di percorsi di apprendimento che possano favorire strutture di rete funzionali al raggiungimento del successo scolastico, attraverso il coinvolgimento attivo degli studenti.

Parole chiave: Social Network Analysis; Learning Analytics; Abbandono scolastico; Interdipendenza; Identificazione del rischio

Introduction

“Learning Analytics” is understood as the sampling, measurement and examination of learning data, derived from environments and from learners, with the aim to improve learning settings and processes, using an evidence base approach (Buckingham Shum, 2012; Siemens, 2013). According to this definition, Learning Analytics (LA) can include different levels of analysis (macro, meso or micro levels). When the primary interest of a study is addressed to the learners and teachers, the analysis occurs at the micro level.

Micro-level analysis can aid in the identification of students at risk of failure, and subsequently, given the highly detailed nature of the analysis, can provide a strategy for improvement (Buckingham Shum, 2012). At this level of analysis, Social Network Analysis (SNA) can be adopted as a technique to collect and analyze interpersonal data. In educational research, SNA is currently emerging as branch of Learning Analytics, aiming to better understand and improve learning contexts (Siemens, 2005; Ferguson, Shum, 2012). SNA has a solid base in the learning sciences; however, from 2008 it begun to be increasingly applied within the literature related to Learning Analytics (Ferguson, 2014). Furthermore, pedagogical theories were increasingly referenced in the LA literature, due to the research of experts employing Social Network Analysis (Dawson, 2008; Dawson, McWilliam, 2008). These authors argue that learning processes are facilitated by the participation of individuals, adopting a socio-constructivist view, that considers learning as a process of social construction of knowledge and skills (Bruner, 1986).

SNA is a methodology that studies the relations and interactions between individuals in a social network (Domínguez, Hollstein, 2014). It originates in the field of sociology, and in network theory in particular, and was developed to study relationships and social structures among individuals in a group (Siemens, 2005). Moreno (1934) developed sociometric theory, from which sociomatrix and sociogram are derive as tools to undertake social analysis. Currently, new approaches to SNA have been developed, through the use of computerized statistical models, that have helped evolve the analysis of students’ interactions in learning environments (Siemens, Long, 2011; Robins et al., 2007).



1. SNA fundamental principles

The SNA methodology takes a relational perspective: it represents social networks through a graph structure, viewing actors as nodes in a network and ties representing links between pairs of actors (Wasserman, Faust, 1994). As a part of social learning analytics (e.g., Ferguson, Shum, 2012), network visualization is considered as a potentially helpful feedback for learners, that can stimulate them to reflect on their social interactions in a group, with the aim of building collective knowledge (e.g., Dawson, 2010).

The nature of the social network is defined by the type of pairwise connection, represented by different network structures. The data analysis of social network quantifies the importance of actors among the network and allows recognition of group(s) of actors connected more densely than others. Through SNA, we can understand if ties and relations between individuals are weak or strong, depending on the frequency, quality or importance of bonds (Domínguez, Hollstein, 2014).

The analysis of networks is founded on a mathematical approach based on graph theory (Van Steen, 2010), which represents the social networks as a binary matrix, called an adjacency matrix. The socio matrix is a method commonly used to conduct social network analysis, to analyze and understand in a quantitative way the interactions between individuals. It represents the presence or strength of ties between the members of a group. Moreover, there are statistical techniques and models to analyze and describe an observed network, which measure its properties (Skyler, Bruce, 2010).

The perspective of analysis can be directed towards the individual (egocentric approach) or to the entire network (complement network design). The first approach can be useful to identify contextual factors that influence learning and that support an individual's learning. In the second approach, a global view identifies elements that hold the network together and provides information's regarding a set of people (Haythornthwaite, de Laat, 2010); it is useful to identify communities and individuals' affinity groups within a network, that can or cannot support learning (Gee, 2004).

2. Application of Social network analysis and early school leaving

Through educational research we can investigate the factors that can interfere with the school participation of at-risk students (Boaler, Staples, 2008; O'Connor, Michaels, 1996), contrasting the tendency to



explain early school leaving as a phenomenon reducible to individual traits, or to a “culture of poverty” (Gutierrez, Rogoff, 2003).

The factors which explain early school leaving are complex and include a variety of situations: failure to attend school and drop out before the conclusion of studies; repetition, irregular frequency in attendance; and delays with respect to school age. Moreover, there is also a “covert dispersion”, that include frequent delays and absences from involvement during lessons. In Italy between the school years 2015/2016 and 2016/2017, 1.35% of students attending first grade of secondary school and 4.31% of students attending the second grade of secondary school, abandoned their schooling (see The National Student Registry, established by Legislative Act 76 on 15 April 2005).

Scholastic outcomes can be influenced by several factors; according to a Bioecological perspective (Brofrenbrenner, 1995), success or failure in school is never determined by a single factor. Rather the outcome in the students’ learning pathways is determined by interaction between child, family, school and social context, and their characteristics. Potentially at risk students can nevertheless achieve academic success against the odds, not only thanks to their personal characteristics, but also thanks to the contributions of people around them (Siraj and Mayo, 2014). Students are embedded within social networks, inside and outside school, that can affect in positive or negative way their learning pathway (Borgatti, Halgin, 2011).

The relational context at school, assumes different forms that can be studied in order to identify students at potential risk of abandonment. To contrast learning failure, is important to stimulate the building of strongly connected interactional networks, that favor the construction of learning communities (Brown, Campione, 1990). Belonging to a community constitutes a part of an individual’s identity: a positive sense of community emerges when individuals feel both important and needed in that group. It is enhanced by promoting interdependence between individuals, namely the collaboration between students with the specific aim of learning from each other (Sarazin, 2017). Researchers argue that interdependence is important in promoting social group cohesion and learning in educational settings (e.g., Osterman, 2000).

Through the application of Social Network Analysis (Dominguez, Hollstein, 2014), it is possible to highlight relational structures that are established during school activities, and identify possible situations of risk or isolated students. On the basis of SNA results, teachers and schools can design activities that promote cohesion among students and knowledge exchange to promote the implementation of interdis-



ciplinary projects, and encourage cooperative connections and collaboration between students (Haythornthwaite, De Laat, 2010).

Starting from these considerations, in the following we draw on an SNA exploratory study to analyze the phenomenon of early school leaving, developed inside the FAMI-IMPACT FVG 2018-2020 project. The exploratory inquiry (Lumbelli, 1989) is addressed to examine the existing network configuration of a class members, in order to define variables or categories of observation, to control and use for future investigations.

We investigated the structure of the social networks in a group of students attending the first year of a vocational school: we analyzed the network of their external relationships (in order to explore the support received by the students outside of school), and the internal network in the class; moreover, we wondered if being at risk of school dropout affects individual school members' likelihood of having relationships in networks.

Therefore, the following research questions are:

RQ1: How is structured the supportive network of the risk students outside the school?

RQ2: How is structured the social network of the risk students, inside the class?



3. Exploratory inquiry

3.1. *Sample*

The sample consisted of 17 students attending their first year of a vocational school in Trieste (15 males, 2 females; age $M=17$; $SD= 1.28$). Students came from a variety of countries of origin: 9 students came from extra UE countries (Foreign students); 7 had Italian citizenship (Italian students); and 1 had Italian citizenship with one of their parents coming from extra UE countries (Migration background). Only 4 students were not repeating the year; the majority failed during the previous school years.

Teachers provided information on the school context; in general, the students enrolled came from previous negative school experiences, and presented lacking motivation. However, there were also some students which showed motivation and interest for studying, although their previous school difficulties; these students were well engaged in their educational path, considered as a possibility for their future job

placement. The teachers also reported the difficulty to positively involve the families of the students, which in many cases were scarcely engaged in the school pathway of their sons.

3.2. *Measures*

Through a survey distributed to the students, we collected general information and relational data. The questions explored the following aspects:

1. General information: age, citizenship, sex
2. Risk of leaving school: “Have you ever thought of leaving school?”
3. Sociometric questions:
 - 3.1 Ego-centric network-external support received for schooling (from family, or other reference adults):
 - (a) “Who help you with your homework?”
 - 3.2 Complement network design within the class group:
 - (a) Network existing before beginning school: “Did you know your classmates before starting this school?”
 - (b) Network after beginning school:
 - B1 “With which of your classmates do you hang out with during break or talk about personal things?”
 - B2 “Which of your classmates do you hang out with even outside school?”.

To answer these socio-metric questions, respondents were provided with a list of their classmates. In order to contribute to ensuring anonymous analysis of the data, the list contained a letter code for each actor.

Students were asked to indicate this letter code by completing the survey; there was no limitation to the number of classmates a respondent could indicate.

3.3. *Data analysis*

Social network properties were calculated using the software package UCINET 6.0 (Borgatti et al., 2002, 2009). Regarding question 3.2(b), the dependent variable is defined as the existence or absence of a relationship between two students (a dyad). For every pair of schoolmates, a value of 1 represents a relationship between them; a value of 0 indicates the absence of a tie between the two members.



We identified the following SNA indexes: network density, network degree centralization and homophily. These indexes provide basic information about the global structure of the network and the activity of the group members.

The density of a network describes the general level of the links between the points in a graph. The more nodes that are directly connected to each other, the denser a graph is. The density index assumes a value that varies from 0 to 1 (= density of a complete graph when all the nodes are adjacent to each other). Centrality helps to describe the power relations that are established within it. A high degree of centrality can mean that the node occupies a position of prestige compared to the others, and can be interpreted as the percentage of relationships that school team members maintain within the whole network.

Homophily is a measure of the tendency of the individuals to be more likely to have ties with others who are similar to themselves on specific attributes (for example, age, gender, education). Similar backgrounds increase the likelihood that students possess shared experiences and knowledge (Reagans, McEvily, 2003). The E-I index is used as the measure of homophily; the value of the index ranges from 1 to -1 (1 being totally heterophilous and -1 totally homophilous).

Moreover, we used a graphical visualization of the networks to represent a major feature of SNA: through sociograms. These are useful to show graphically relevant information in the network (such as sub-groups, relevant positions, more and less prominent members).



3.4. Results

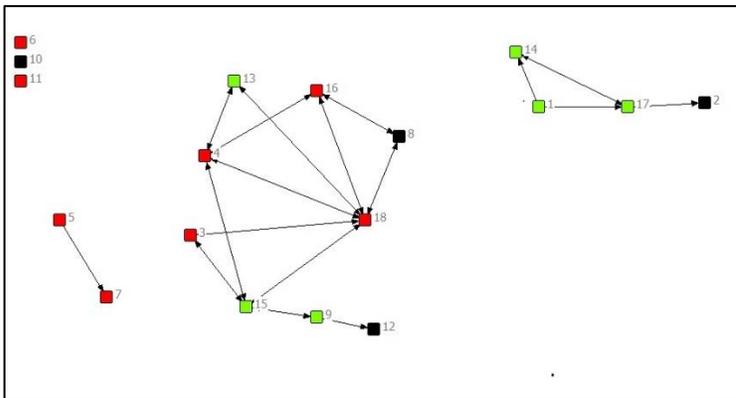
Risk of school leaving: “Have you ever thought of leaving school?”

Sixty-five percent (65%) of students declared that they had thought about abandoning school. Students justified their response as being due to relational difficulties in the classroom, demotivation and low consideration of the usefulness of schooling for their life or to find a job.

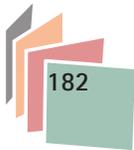
1. Egocentric network: External support received for school:

In general, students received low support at home from their family during homework, and was particularly lacking in students at risk of school dropout. Students most frequently received support from mothers and educators (graph 1).

The network constituents after beginning school are represented in Graph 3.



Graph 3: Network after beginning school: red circle = student at risk; green = student not at risk; black= student who didn't give a response to question 3.2 (b)



3.5. Discussion of results

RQ1: How is structured the supportive network of the risk students outside the school?

Our analysis was conducted in a class where the majority of students were not on track with their school attendance and most had experienced school failure in previous years; looking at the results shown in Graph 1, it is possible to see how the structure of the support network outside the school is very dispersed in general, and how it is almost entirely absent for students who consider leaving school. This result is in line with previous research, which highlight that students at risk of school failure have usually poor social support (Richman, Bowen, 1997). Family has a very strong impact on the possible scholastic success or failure of the children; Richman et al. (1998), have analyzed the social support provided to high school students, identified by school as at risk of school failure. Researchers found that when students at risk of failure received from adult caretakers a positive support, they were able to reach good school performance, with positive consequences for their learning outcome. The risk factor was therefore mitigated by the presence of functional and supportive external social network.

RQ2: How is structured the social network of the risk students, inside the class?

The results show that most of the boys did not know each other before starting school, and therefore the social network currently formed in the class was built after the start of school. The network shows that relationships between the students are not close, connections are not high, there are subgroups and some students are isolated. There are some more popular students who have more numerous connections, but in general there are no positions of prestige or leadership. The results therefore show a group where members have low interdependence and lack cohesion within a group. Prior research who examined the social relationship of the students at risk inside high school, highlights that the social integration was a protective factor for the drop out: students well integrated with peers, with dense and more centralized social network, were less likely to drop out; by contrast students isolated were more likely to early school leaving (Staff, Kreager, 2008; South et al., 2007).

Furthermore, our results show that students who have considered leaving school show a slightly tendency to associate outside school with classmates who have the same idea. This result highlights just a trend that should be deeper investigated. Making a comparison with previous research who have analyzed the homophily in student social networks, our result is in line with the evidences that at-risk students tend to bond with other students at risk, increasing significantly the like hood of drop out of high school (Ellenbogen, Chamberland, 1997; Ream, Rumberger, 2008).

Conclusion

Social Network Learning Analytics has been used to study the networks inside groups of students, detecting relevant influences on at-risk students, analyzing how they are connected into the network. The combination of data collection on a social network, with statistics and visualizations of connections through graphs, can increase the understanding of the learning networks, highlighting what constitutes a learning bond and how learning links are activated and supported (Dawson, 2008).

Social Network Analysis (Dominguez, Hollstein, 2014) makes it possible to identify structures of affiliation and reciprocity that are established during school activities, in order to develop action research projects that elaborate activities that promote cohesion, and the ex-



change of knowledge among students. Through social network analysis, it is possible to investigate how students are affected by or use their social networks in activities related to education.

Our preliminary results highlight as the “structure of the network” is a variable of analysis that can provide important information for the study of the phenomenon of early school leaving, and address educational intervention. This variable should be deeper in the future research, in order to collect comparative data from students with different background and students coming from different educational institutes.

Social networks can be an important source for students to support themselves, affecting individuals’ behaviors and attitudes. However, networks do not always have a positive function. When networks are poorly connected, and a strong relational structure is missing, the sense of belonging to a learning community is lost; if the external networks in the social context are lacking, there is no drive and containment for children in difficulty who run the risk of losing their way. If the relationships within the network are dysfunctional, there is the risk that nonfunctional learning behaviors will be enhanced and strengthened.

Practitioners and teachers should include social network information to plan intervention strategies to support learning pathway of students at risk. It is important to understand the relational structure of the student social network, and work subsequently to build cohesive group, by stimulating a positive interdependence (Brown, Campione, 1990). This suggests a rethinking of teaching practices and the design of learning paths that can foster functional network structures to achieve academic success, through the active involvement of students. Moreover, educational intervention should develop strategies to increase the effectiveness of resources presented in the external environment system, individuating the source available and the reference caretakers, even if this process is often difficult.

References

- Boaler J., Staples M. (2008). Creating mathematical futures through an equitable teaching approach: The case of Railside School. *Teachers College Record*, 110(3), 608-645.
- Borgatti S.P., Mehra A., Brass D.J., Labianca G. (2009). Network analysis in the social sciences. *Science*, 323(5916), 892-895.
- Borgatti S.P., Halgin D.S. (2011). On network theory’. *Organization Science*, 22 (5), 1168-81.
- Borgatti S.P., Everett M.G., Freeman L.C. (2002). *Ucinet for Windows: Software for Social Network Analysis*. Harvard, MA: Analytic Technologies.



- Bronfenbrenner U. (1995). The bioecological model from a life course perspective: Reflections of a participant observer. In P. Moen, G. H. Elder, Jr., K. Lüscher (Eds.), *Examining lives in context: Perspectives on the ecology of human development* (pp. 599-618). Washington, DC, US: American Psychological Association.
- Brown A.L., Campione J.C. (1990). Communities of learning and thinking, or a context by any other name. *Human Development*, 21, 108-125.
- Bruner J. (1986). *Actual minds. Possible worlds*. Cambridge: Harvard University Press
- Buckingham Shum S. (2012). *Learning analytics. UNESCO policy brief*. Retrieved from <http://iite.unesco.org/pics/publications/en/files/3214711.pdf>
- Dawson S. (2008). A study of the relationship between student social networks and sense of community. *Journal of educational technology & society*, 11(3), 224-238.
- Dawson S. (2010). "Seeing" the learning community: An exploration of the development of a resource for monitoring online student networking. *British Journal of Educational Technology*, 41(5), 736-752.
- Dawson S., McWilliam E. (2008). Investigating the application of IT generated data as an indicator of learning and teaching performance. *Queensland University of Technology and the University of British Columbia*, 41.
- Domínguez S., Hollstein B. (2014). *Mixed methods social networks research: Design and applications* (Vol. 36). Cambridge University Press.
- Ellenbogen S., Chamberland C. (1997). The peer relations of dropouts: a comparative study of at-risk and not at-risk youths. *Journal of Adolescence*, 20 (4), 355-367.
- Ferguson R. (2014). Learning Analytics: fattori trainanti, sviluppi e storie. *Italian Journal of Educational Technology*, 22(3), 138-147.
- Ferguson R., Shum S.B. (2012). Social learning analytics: Five approaches. In *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge (LAK '12)*, 29 April–2 May 2012, Vancouver, BC, Canada (pp. 23-33). New York: ACM.
- Gee J. (2004). *Situated Language and Learning: A Critique of Traditional Schooling*. Routledge.
- Gutierrez K.D., Rogoff B. (2003). Cultural ways of learning: Individual traits or repertoires of practice, *Educational Researcher* 32 (5), 19-25.
- Haythornthwaite C., De Laat M. (2010). Social networks and learning networks: Using social network perspectives to understand social learning. In *Proceedings of the 7th international conference on networked learning* (pp. 183-190). Aalborg, Denmark: Lancaster University.
- Legislative Act 76 of 15 April 2005 "Definizione delle norme generali sul diritto-dovere all'istruzione e alla formazione ai sensi dell'articolo 2, comma 1, lettera c, della legge 28 marzo 2003, n. 53".
- Lumbelli L. (1989). Pedagogia sperimentale e ricerca esplorativa, *Scuola e città*, 3, 97-104.
- Moreno J.L. (1934). *Who shall survive: A new approach to the problem of human interrelations*. Washington, DC, US: Nervous and Mental Disease Publishing Co.
- O'Connor M.C., Michaels S. (1996). Shifting participant frameworks: Orches-



- trating thinking practices in group discussion. In Hicks (Ed.), *Child Discourse & Society Learning: An Interdisciplinary Perspective* (pp. 63-103). Cambridge: Cambridge University Press.
- Osterman K.F. (2000). "Students' need for belonging in the school community". *Review of Educational Research*, 70 (3), 323-67.
- Reagans R., McEvily B. (2003). Network structure and knowledge transfer: The effects of cohesion and range. *Administrative science quarterly*, 48(2), 240-267.
- Ream R., Rumberger R. (2008). Student engagement, peer social capital, and school dropout among Mexican American and non-Latino white students. *Sociology of Education* 81, 109-139.
- Richman I.M., Bowen G.L. (1997). School failure: An ecological-interactional-developmental perspective. In M.W. Fraser (Ed.), *Risk and resilience in childhood: An ecological perspective* (pp. 95-116). Washington, DC: Press.
- Richman J.M., Rosenfeld L.B., Bowen G.L. (1998). Social support for adolescents at risk of school failure. *Social work*, 43(4), 309-323.
- Robins G., Pattison P., Kalish Y., Lusher D. (2007). An introduction to exponential random graph models for social networks. *Social Networks* 29, 173-191.
- Sarazin M. (2017). Can student interdependence be experienced negatively in collective music education programmes? A contextual approach. *London Review of Education*, 15, 488-504.
- Siemens G. (2005). Connectivism: A learning theory for the digital age. *International journal of instructional technology and distance learning* 2, 1, 3-10.
- Siemens G. (2013). Learning analytics: The emergence of a discipline. *American Behavioral Scientist*, 57(10), 1380-1400.
- Siemens G., Long P. (2011). Penetrating the fog: Analytics in learning and education. *Educause Review* 46, 5, 30-32.
- Siraj I., Mayo A. (2014). *Social class and educational inequality: The impact of parents and schools*. Cambridge University Press.
- Skyler J., Bruce A. (2010). Inferential Network Analysis with Exponential Random Graph Models. *Society*, 11(3), 2008, 224-238.
- South S., Haynie D., Bose S. (2007). Student mobility and school dropout. *Social Science Research*, 36, 68-94.
- Staff J., Kreager D. (2008). Too cool for school? Peer status and high school dropout. *Social Forces*, 87, 445-471.
- Van Steen, M. (2010). An Introduction to Graph Theory and Complex Networks. *Copyrighted material*.
- Wasserman S., Faust K. (1994). *Social network analysis: Methods and applications*. New York: Cambridge University Press.

Analytics dei testi riflessivi scritti dai docenti neoassunti nel portfolio digitale

Newly Qualified Teachers' reflective writing: analytics and preliminary results

Valentina Della Gala

Istituto Nazionale Documentazione, Innovazione, Ricerca Educativa (INDIRE), v.dellagala@indire.it

Giulia Chiriatti

Università di Pisa, giuliachiriatti@gmail.com

Felice Dell'Orletta

Istituto di Linguistica Computazionale "Antonio Zampolli" (ILC-CNR) - ItaliaNLP Lab, felice.dellorletta@ilc.cnr.it

Maria Chiara Pettenati

Istituto Nazionale Documentazione, Innovazione, Ricerca Educativa (INDIRE), mc.pettenati@indire.it

Giulia Venturi

Istituto di Linguistica Computazionale "Antonio Zampolli" (ILC-CNR) - ItaliaNLP Lab, giulia.centuri@ilc.cnr.it

187

In this paper, we present the preliminary results of the analysis performed on 50,000 texts written by the Newly Qualified Teachers to reflect on their practice. The Italian Induction Programme for NQT includes online activities to be carried out in the E-portfolio managed by Indire. As part of the training monitoring with the aim to evaluate our support and to improve instructional scaffolding to teachers, we analysed teachers' reflective writing mixing a new developed framework for teachers' reflective writing and Natural Language Processing (NLP) techniques. In providing a case study of a large-scale analysis of textual teachers' data in an institutional setting, this paper wants to contribute to the field of learning analytics. We will describe the context, the characteristics of the data analysed, how all of this has shaped the analytics design. We will describe its implementation and therefore the procedures, tools and metrics developed to represent the content of the teachers' reflective writing, and we will try to evaluate if and how these preliminary results can help us to reach the research goals and to understand the strengths and weaknesses of different methodologies.

Keywords: Teacher professional development; reflective writing; content analytics; learning content.

Presentiamo i risultati preliminari e l'analisi svolta su circa 50.000 testi scritti dai docenti neo nominati in ruolo per riflettere su due attività didattiche svolte con gli studenti, nel contesto del percorso dell'anno di formazione e prova 2016/17. Il percorso prevede attività in presenza e attività a distanza completate sul portfolio digitale, ospitato nell'ambiente online gestito dall'Indire. Nell'ambito del monitoraggio della formazione, con il fine di ottimizzare gli strumenti e il supporto fornito, abbiamo interrogato i dati testuali prodotti dai docenti nell'interazione con l'ambiente per capire se i testi presentassero evidenze riconducibili alle scritture riflessive. Obiettivi dell'indagine sono stati la definizione di uno schema per la classificazione dei testi sulla base del livello di riflessività evidenziato e l'im-

Valentina Della Gala è autrice dei §§ 1, 3, 4, 5, 6, 6.1 e 7; Giulia Chiriatti, Felice Dell'Orletta e Giulia Venturi dei §§ 5.1 e 6.2; Maria Chiara Pettenati del § 2.

piego di strumenti di Trattamento Automatico del Linguaggio (TAL) per l'analisi dell'intero corpus testuale prodotto dai docenti. Descriveremo il contesto scientifico e progettuale, le caratteristiche dei dati analizzati, come questo abbia determinato il disegno d'indagine; descriveremo inoltre la sua implementazione e dunque le procedure, gli strumenti e le metriche adottate o elaborate per rappresentare il contenuto dei dati; infine discuteremo i primi risultati e alcuni vantaggi e limiti dell'approccio adottato.

Parole chiave: formazione professionale dei docenti; scrittura riflessiva; analytics del contenuto; contenuto di apprendimento.

1. Introduzione



Dalla sua istituzione sperimentale nell'a.s. 2014/15, il percorso di formazione e prova ha coinvolto ad oggi circa 194.000 docenti neo nominati in ruolo, la cui interazione con l'ambiente *online* ha prodotto un'enorme quantità di dati di tracciamento delle azioni e di contenuto elaborato dai docenti, oltre al contenuto dell'ambiente stesso, che è variato, pur di poco, di edizione in edizione. La sovrabbondanza di dati è un fenomeno ampiamente discusso da anni in molti campi di ricerca. Nell'ambito educativo l'analisi di grandi quantità di dati è stata riconosciuta come una strada per nuovi modelli di scoperta (Siemens, 2013), con importanti implicazioni ad ogni livello, da quello istituzionale a quello del singolo studente. L'analisi del *learning content*, cioè dei dati, principalmente testuali ma in potenza anche in formato video o audio, prodotti dagli studenti (saggi, discussioni ecc.), dagli educatori (sillabi, documenti ecc.) o dagli editori (libri di testo ecc.), è una sotto area dei LA definita *content analytics* (CA), ad indicare i metodi automatici per esaminare, valutare, indicizzare, filtrare, raccomandare e visualizzare «different forms of digital learning content, regardless of its producer (e.g., instructor, student) with the goal of understanding learning activities and improving educational practice and research» (Kovanović et al., 2017). Con l'intenzione di ottenere informazioni sull'uso degli strumenti a supporto dei docenti in anno di formazione e prova e fornire evidenze per ottimizzarli e migliorare il supporto alla loro esperienza formativa, abbiamo analizzato il contenuto dei testi scritti da circa 25.000 docenti per riflettere su 2 attività didattiche svolte con gli allievi. Nei paragrafi che seguono presentiamo gli obiettivi dell'anno di formazione e prova, il disegno di indagine, le analisi compiute, gli strumenti impiegati e i risultati preliminari ottenuti.

2. Contesto: il percorso, l'attività oggetto di analisi e il ruolo della riflessione nel modello formativo dei neoassunti

Il sostegno agli insegnanti durante la fase iniziale della carriera o *induction*, è un elemento di indirizzo strategico per tutti i paesi, evidenziato anche nella Comunicazione della Commissione europea sullo sviluppo scolastico e l'eccellenza nell'insegnamento (Commissione Europea, 2018). L'Italia, come la grande maggioranza dei sistemi europei, mette in atto ogni anno un programma nazionale obbligatorio della durata di un anno per l'accompagnamento all'ingresso in servizio dei nuovi docenti. Fino al 2013/14 il modello italiano di *induction* prevedeva un impianto centrato su incontri formativi in presenza e un'attività online dedicata alla fruizione di contenuti in auto-formazione e alla redazione di un documento di sintesi finale, dal 2014/15 il modello è stato revisionato e consolidato nel suo nuovo formato con il DM 850/15. Recependo i risultati della ricerca educativa nazionale ed internazionale sulla formazione dei docenti, con particolare attenzione all'avvio alla professione, il nuovo modello poggia su alcuni principi fondanti (D'Amico, 2018), che possono essere così riassunti:

- *governance* distribuita tra autorità centrale (MIUR, che stanziava i finanziamenti, definisce l'impianto, le priorità, attribuisce ruoli, funzioni e tempi), gli Uffici Scolastici Regionali (che recepiscono le indicazioni nazionali e le contestualizzano, attuando un governo territoriale della formazione, col supporto degli Uffici di Ambito Territoriale), le Scuole Polo (che implementano l'offerta formativa in presenza prevista dal modello di formazione secondo le priorità tematiche stabilite a livello nazionale e i bisogni formativi locali), la scuola di presa di servizio dei nuovi docenti (che accoglie e guida l'inserimento del docente nella comunità educante), Indire (Istituto Nazionale di Documentazione, Innovazione e Ricerca Educativa) che implementa e gestisce l'LMS a supporto all'anno di formazione e prova.
- accompagnamento guidato nella scuola di presa di servizio con il Dirigente Scolastico chiamato a stipulare un Patto Formativo col nuovo docente, osservarne l'attività didattica e nominare un Tutor accogliente, che svolge un'azione di affiancamento tra pari (*peer to peer*) che comprende una riflessione sulle competenze in ingresso, una co-progettazione dell'attività didattica, l'osservazione reciproca in classe.
- alternanza tra immersione e distanziamento dalla pratica didattica guidata da una serie di "dispositivi" che confluiscono nel portfolio formativo digitale (Rossi et al., 2015) implementato da Indire.



Uno dei dispositivi implementato nel portfolio formativo è quello relativo alle “Attività Didattiche” (le attività sono 2 identiche), oggetto di approfondimento nel presente studio. Attraverso questo dispositivo si dà supporto alla progettazione-documentazione-riflessione guidati dell’azione didattica del docente nel nuovo contesto educativo (Rossi et al., 2015), al fine di supportarne l’acquisizione di un atteggiamento di riflessività che – adottato sistematicamente come *habitus* professionale – favorisca la crescita professionale dell’insegnante in un’ottica di formazione continua, oltre l’anno di prova (Rossi et al., 2018).

3. Il corpus documentale analizzato

Il dispositivo tramite il quale al docente è stato richiesto di documentare 2 attività didattiche, si componeva nell’a.s. 2016/17 di 5 sezioni, (progettazione; riflessione sulla progettazione; documentazione; narrazione; riflessione post-azione). Analizziamo i testi scritti nella sezione sulla riflessione post-azione in cui era richiesto di riflettere sull’attività didattica svolta rispondendo, se lo si riteneva opportuno, ad 8 domande stimolo. La scelta di fornire un solo *form* per rispondere a un intero gruppo di domande è dovuta alla volontà di lasciare il docente libero di scrivere potendo scegliere a quali domande rispondere, in quale ordine o se avvalersi o meno delle domande. Il risultato sul piano della struttura dei dati è stato un data-set in cui ad ogni codice identificativo del docente corrispondono il numero dell’attività (1 e 2) e l’intero testo composto nella sezione.

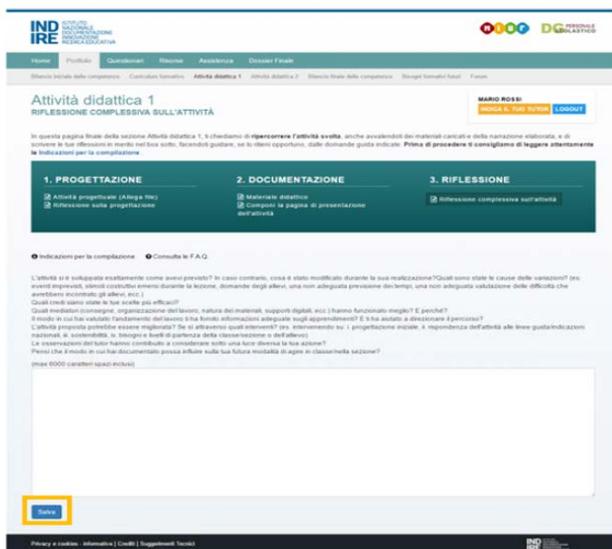
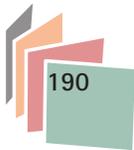


Fig. 1: Screenshot della sezione del Portfolio digitale dedicata alla scrittura di un testo riflessivo post-azione

4. Analisi del contenuto e disegno dell'indagine

L'articolazione dell'indagine origina dalla *content analysis* (Krippendorff, 2004), di un campione di circa 200 testi che ci ha permesso di individuare alcuni tratti ricorrenti. I testi del campione sono risultati:

- composti da un numero di caratteri molto inferiore ai 6000 (limite massimo consentito);
- contenere risposte alle domande tendenzialmente disposte in modo ordinato;
- non contenere sempre le risposte a tutte le domande;
- presentare livelli diversi di complessità sintattica;
- presentare livelli diversi di ricchezza lessicale.

La complessità sintattica e l'aderenza alla realtà dell'esperienza si sono rivelati criteri discriminanti per distinguere tra segmenti di testo (o risposte) in cui vi si può riconoscere un ripensamento autentico dell'esperienza e segmenti non coerenti con le aspettative. Abbiamo dunque ipotizzato che questi, insieme ad altri attributi relativi al coinvolgimento cognitivo (Garrison, 2001) ed emotivo (Dewey, 1933; Chung, 2007), potessero fornirci indicazioni automatizzabili per indagare tutti i testi.

Una scelta critica in questi casi è quella dell'unità da considerare per l'analisi. Noi abbiamo scelto di ricorrere alla suddivisione, in prevalenza già presente nel *corpus*, in testi di risposta alle 8 domande stimolo. Scelta che ci consentiva anche di valutare singolarmente l'efficacia di ogni domanda. Abbiamo, quindi, strutturato l'indagine in 2 macro azioni per indagare:

1. cosa i docenti hanno scritto, cioè a quale sollecitazione/*topic* avanzata dalle domande stimolo hanno risposto.
2. come lo hanno scritto, i testi si risolvono in compilazioni burocratiche o mostrano evidenze di un ripensamento circostanziato dell'esperienza didattica e delle caratteristiche condivise in letteratura come proprie di un approccio riflessivo? Quali indicazioni forniscono sull'adeguatezza degli strumenti? Alcuni hanno funzionato meglio?

5. Segmentazione dei testi per argomento: individuazione delle unità di analisi

La prima azione di questa fase è stata l'annotazione manuale di un campione di testi per risposta. Cioè ad ogni segmento abbiamo attribuito



il codice della domanda cui risponde (Tab. 1). Il campione così annotato è andato a costituire il *corpus* per l'addestramento (*training corpus*) degli algoritmi di classificazione.



Codice	DOMANDE	Codice rivisto
12.1	L'attività si è sviluppata esattamente come avevi previsto? In caso contrario, cosa è stato modificato durante la sua realizzazione?	12.1.2
12.2	Quali sono state le cause delle variazioni? (es: eventi impreveduti, stimoli costruttivi emersi durante la lezione, domande degli allievi, una non adeguata previsione dei tempi, una non adeguata valutazione delle difficoltà che avrebbero incontrato gli allievi, ecc.)	
12.3	Quali credi siano state le tue scelte più efficaci?	12.3.4
12.4	Quali mediatori (consegne, organizzazione del lavoro, natura dei materiali, supporti digitali, ecc.) hanno funzionato meglio? E perché?	
12.5	Il modo in cui hai valutato l'andamento del lavoro ti ha fornito informazioni adeguate sugli apprendimenti? E ti ha aiutato a direzionare il percorso?	12.5
12.6	L'attività proposta potrebbe essere migliorata? Se sì attraverso quali interventi? (es. intervenendo su: i. progettazione iniziale, ii. rispondenza dell'attività alle linee guida/indicazioni nazionali, iii. sostenibilità, iv. bisogni e livelli di partenza della classe/sezione o dell'allievo)	12.6
12.7	Le osservazioni del tutor hanno contribuito a considerare sotto una luce diversa la tua azione?	12.7
12.8	Pensi che il modo in cui hai documentato possa influire sulla tua futura modalità di agire in classe/nella sezione?	12.8

Tab. 1: Elenco delle domande stimolo con il codice attribuito per l'analisi e quello rivisto dopo i primi cicli di annotazione

5.1 Classificazione con strumenti di TAL

Allo scopo di riconoscere all'interno di ogni testo i singoli segmenti di risposta scritti dai docenti alle domande stimolo, sono stati adottati algoritmi di classificazione automatica supervisionata (*machine learning*) basati su strumenti di TAL. Tali algoritmi sono capaci di associare ad ogni testo in analisi la classe di appartenenza (nel nostro caso la singola domanda: 12.1, 12.2 ecc.), assegnando al testo un punteggio di probabilità di appartenenza ad ogni classe. La classe con punteggio massimo è quella che l'algoritmo associa al segmento di testo. Il classificatore è capace di apprendere la strategia di classificazione "imparando" dal *training corpus*. Il classificatore è capace nella fase di addestramento di apprendere le caratteristiche specifiche dei testi appartenenti ad una

classe e di associare ad ognuna di esse un peso, che rappresenta il contributo della specifica caratteristica nel predire la classe di appartenenza del testo. Terminata la fase di addestramento, in fase di analisi il classificatore è capace di estrarre da un nuovo testo preso in esame (quindi sconosciuto al classificatore fino a quel momento) le varie caratteristiche che, associate ai pesi definiti in fase di addestramento, determinano la classe di appartenenza del testo. Pertanto, i passaggi dell'intero processo di classificazione sono stati: *i*) la scelta della dimensione del *training corpus*; *ii*) la selezione delle caratteristiche per classificare in modo adeguato i segmenti di testo.

In letteratura, è ben noto come una maggiore quantità di testi contenuti nel *training corpus* permette di raggiungere accuratezze di analisi maggiori. Questo è dovuto al fatto che più esempi annotati a mano costituiscono una base di apprendimento più ampia. Dal momento però che la costruzione del *training corpus* è un'operazione manuale, essa richiede un notevole investimento di tempo e di esperti del dominio. Per questo abbiamo deciso di procedere per gradi, incrementando progressivamente la quantità di segmenti di testo annotati manualmente. In questo processo, l'effetto di ogni nuovo incremento è stato verificato ripetendo la classificazione automatica e misurando le nuove accuratezze di volta in volta raggiunte. Siamo partiti da un numero iniziale di 762 frasi e siamo arrivati a un totale di 3.322 frasi. Da notare come oggetto delle nostre analisi siano state le singole frasi e non l'intero segmento. La motivazione di questa scelta è legata al fatto che i docenti non hanno sempre risposto in maniera sequenziale alle domande stimolo. La classificazione della singola frase ha pertanto permesso di superare questo problema, andando a rintracciare le diverse frasi che, a volte sparse nel testo, rispondevano alle domande poste. Un ulteriore aspetto di organizzazione interna del *training corpus* ha riguardato il numero di domande rispetto a cui classificare il *corpus*. Le 8 domande stimolo richiedevano risposte diverse. Tuttavia, i primi esperimenti di classificazione automatica hanno rivelato che la distinzione tra le risposte alle domande 12.1 e 12.2, e 12.3 e 12.4 era difficilmente individuabile. In entrambe le coppie, il contenuto della risposta richiesta dalla domanda è molto simile, poiché le due domande hanno uno stretto legame logico tra di loro. La mancanza di caratteristiche chiaramente distintive portava ad accuratezze di classificazione molto basse. È stato pertanto scelto di accorpare le domande, riducendo il numero delle classi a 6 (Tab. 1, "codice rivisto").

La selezione di quali siano le caratteristiche che contribuiscono a classificare le frasi contenute nei testi dei docenti è stato un secondo aspetto considerato. Le caratteristiche scelte includono elementi lin-



guistici estratti sia dalle risposte sia dalle domande stimolo, oltre ad elementi di contiguità testuale. Inizialmente, l'intero *corpus* è stato analizzato impiegando la catena di strumenti di annotazione linguistica automatica sviluppata dai ricercatori del gruppo ILC-CNR¹. Ogni risposta e ogni domanda è stata segmentata in singole frasi e all'interno di ogni frase sono state individuate le unità ortografiche che la compongono (*token*). Successivamente è stata fatta l'analisi morfo-sintattica e la lemmatizzazione del testo *tokenizzato*, infine l'analisi della struttura sintattica della frase in termini di relazioni di dipendenza sintattica. Una serie di esperimenti preliminari hanno permesso di individuare l'insieme di caratteristiche che hanno il maggiore potere discriminatorio. Le caratteristiche estratte dalle risposte fornite dai docenti includono: *i*) sequenze di unigrammi e bigrammi di caratteri, *token* e lemmi e *ii*) la rappresentazione vettoriale e compatta della semantica di ogni parola contenuta nelle risposte utilizzando informazioni sul contesto linguistico in cui è essa inserita. Per estrarre dal testo questa seconda caratteristica sono stati adottati metodi allo stato dell'arte noti come metodi di estrazione di *word embeddings* (Mikolov, 2013). In aggiunta, è stato considerato in fase di classificazione il lessico contenuto nelle domande stimolo fornite. L'intuizione è stata che alcune delle parole impiegate nelle domande fossero state utilizzate anche nelle risposte e che potessero dunque contribuire a classificarne il contenuto. Infine, gli esperimenti preliminari condotti hanno messo in luce che l'informazione sulla classe attribuita alla frase precedente o successiva a quella in analisi, in casi di dubbi sull'attribuzione, permette di migliorare l'accuratezza della classificazione. L'idea è che se il contenuto di una frase verte su di un determinato argomento di riflessione è molto probabile che la frase successiva tratti dello stesso argomento e condivida con essa il livello di riflessione. Al termine dell'intero processo, sono stati ottenuti i risultati di classificazione riportati nella Tab. 2, secondo i quali l'accuratezza complessiva della classificazione è del 70%, con estremi nelle risposte alle prime domande, che sono risultate anche le più lunghe e articolate, e alla domanda 12.7, sola domanda ad avere per oggetto il tutor, e per questo facilmente individuabile². Al contrario la

1 <http://www.italianlp.it/demo/linguistic-annotation-tool/>

2 In questo contributo l'accuratezza delle due tipologie di esperimenti condotti è stata calcolata come l'F-score, la media armonica tra precisione (calcolata come il numero di veri positivi diviso il numero di tutti i risultati positivi) e recupero (calcolato come il numero di veri positivi diviso il numero di tutti i test che sarebbero dovuti risultare positivi).

domanda 12.6 è quella classificata con meno accuratezza perché presenta elementi meno caratterizzanti. Inoltre i dati confermano il buon esito della scelta di accorpare le domande 12.1 e 12.2, la cui attribuzione è notevolmente migliorata.

Sotto-risposta	Accuratezza
12.1.2	80,9
12.3.4	70,3
12.5	61,4
12.6	58,1
12.7	83,5
12.8	65,8
TOTALE	70

*Tab. 2: Risultati della fase finale di classificazione automatica
Distribuzioni delle attribuzioni corrette per la classe della domanda.*



6. Analisi dei testi per livello di riflessione

A partire dall'analisi delle risposte e dagli studi sulle caratteristiche e il ruolo del pensiero riflessivo nell'ambito della formazione docente, abbiamo identificato 4 categorie in cui classificare i testi; selezionato gli attribuiti riconosciuti come specifici di uno scritto riflessivo ed elaborato una matrice che incrocia categorie e attributi per validare l'affidabilità della classificazione attraverso una codifica da parte di più valutatori. Il *corpus* codificato manualmente è stato poi analizzato con strumenti di TAL.

6.1 Cosa s'intende per riflessione e perché è stata categorizzata

«There is no such thing as the definition of reflection, but there is an on-going discourse about what constitutes reflection» (Ullman, 2015). Nell'ambito della formazione e della professione docente, riflettere ha un'accezione specifica e l'educazione alla riflessività un ruolo sostanziale. La letteratura sull'argomento è talmente ampia da coincidere quasi con l'intero ambito scientifico della formazione degli insegnanti. Tutti gli studi consultati fanno risalire a Dewey la prima esplicita e generatrice associazione tra pensiero riflessivo ed educazione. Dewey distingue il processo mentale riflessivo dagli altri (credenze, fantasticherie, flusso di coscienza), identificandolo in un processo del pensiero che a partire da una situazione percepita come ambigua, problematica, non nota,

indaga la realtà e la relaziona ad esperienze pregresse e ulteriori termini, al fine di eliminarne l'ambiguità e sciogliere la situazione problematica con l'azione e l'acquisizione di una nuova conoscenza (Dewey, 1933). Il pensiero riflessivo prende le mosse dall'esperienza, quindi, e si configura come un processo di costruzione di significati, attraverso la creazione di relazioni tra le dimensioni di un'esperienza, tra esperienze diverse e tra esperienze e conoscenze. Il fine di questo processo è trovare una soluzione a un problema/ambiguità e/o aumentare la propria base di conoscenze (Rodgers, 2012). La postura riflessiva è propria dei professionisti esperti. Una pratica in cui essa agisce è foriera di conoscenza professionale e risolve quello che Schon chiama il «dilemma tra rigore e pertinenza» (Schön, 2006), cioè consente l'integrazione tra ricerca scientifica e pratica didattica. Quando il processo riflessivo, ponendo in discussione i presupposti del problema, guida quello decisionale o l'azione, allora si configura come apprendimento dall'esperienza (Mezirow, 1990). Il ruolo della riflessione è stato analizzato in contrasto con gli obiettivi tecnico-procedurali delle formazioni docenti come progettate fino agli anni '80 negli Stati Uniti, in questo senso la "scoperta" della riflessività ha messo al centro dell'attenzione degli educatori il ruolo delle capacità cognitive, affettive, comunicative e critiche/emancipatorie del docente e dello studente nel processo di insegnamento/apprendimento. Sono stati dimostrati i benefici di una postura riflessiva (Boud, Walker 1991), e descritti metodi e strumenti per sostenerne lo sviluppo (p.e. *reflective conversation, storytelling, micro-teaching, action research* ecc.). Tra questi, la scrittura è il *medium* più ampiamente sperimentato. Come conseguenza negli ultimi 25 anni sono stati proposti molti schemi per valutare le "scritture riflessive". Alcuni hanno avuto l'obiettivo di valutare la profondità della riflessione; altri l'ampiezza, cioè i temi affrontati (Moon, 2006). L'approccio e i criteri di valutazione utilizzati sono tuttavia molto dipendenti dagli obiettivi dell'attività entro cui è analizzata, da cosa s'intende per riflessività in quel determinato contesto, dal contesto stesso e dagli input forniti. Al contrario di gran parte degli studi sull'argomento, nel nostro caso la redazione dei testi è avvenuta in un contesto "debole", di cui non è dato conoscere completamente il tipo di supporto e le indicazioni fornite al singolo docente, perché distribuite tra soggetti diversi; inoltre ha riguardato un numero altissimo di soggetti. Da questo punto di vista le sfide che ha presentato l'analisi sono state più vicine a quelle che affrontate nell'ambito dei *reflective writing analytics* (Gibson, Kitto 2015; Buckingham Shum et al., 2016; Buckingham Shum et al., 2017; Liu et al., 2019), che a quelle degli studi sulla valutazione della riflessività nella formazione docente, che generalmente coinvolge numeri bassi.

Da quest'ultimo ambito abbiamo però derivato le dimensioni che compongono il concetto di riflessività, mentre dal primo la legittimità per estendere l'esame all'intero *corpus* con metodi automatici.

Abbiamo assunto la complessità sintattica e l'aderenza alla realtà come criteri distintivi (Par. 4), e individuato le seguenti 4 classi per discriminare tra livelli di impegno/coinvolgimento diversi, e di cui solo in una, nella classe "riflessiva", c'è una compresenza delle caratteristiche considerate in letteratura come proprie di uno scritto riflessivo.

1. Non riflessiva: risponde con una semplice parafrasi della domanda e/o una descrizione dell'accaduto. Non viene fatto alcun riferimento esplicito o implicito a dubbi o problemi; non presenta evidenze di un approccio di ricerca. Assenza (o scarsa presenza e non significativa) di preposizioni subordinate.
2. Considerazione generale: il segmento mostra un legame debole con la realtà dell'esperienza didattica. I riferimenti ad essa, se ci sono, sono deboli. Si tratta di considerazioni generali o generiche che possono essere anche articolate in enunciati complessi.
3. Riflessiva descrittiva: presenta considerazioni direttamente connesse alla realtà dell'azione didattica; è possibile l'esplicitazione di conoscenze ed esperienze pregresse ma non emerge con chiarezza alcun problema o dubbio o questo non è portato alle sue conseguenze ultime, cioè non è individuata alcuna soluzione. L'enunciato è articolato.
4. Riflessiva: la questione affrontata è considerata problematica; sono riferiti dubbi relativi all'azione didattica; c'è evidenza di un cambiamento nell'attitudine del docente o che egli ha acquisito nuove conoscenze o a modificato la sua azione didattica. L'enunciato è articolato.

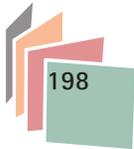
Abbiamo valutato l'affidabilità dello schema attraverso una procedura di validazione intersoggettiva. A questo fine abbiamo definito gli attributi di uno scritto riflessivo entro il nostro contesto di riferimento e sulla base della letteratura. Per ognuno abbiamo anche indicato come si può manifestare in forma scritta. Un segmento di testo è classificato come riflessivo se:

- I. contiene riferimenti espliciti all'esperienza didattica svolta. Presenza di termini concreti, ricchezza lessicale.
- II. include almeno un riferimento a: esperienze precedenti; conoscenze pedagogiche, di contenuto o di altro tipo con il fine di sostenere l'interpretazione e l'analisi dell'esperienza da parte del docente. Il risultato è una struttura sintattica articolata con presenza di verbi al gerundio, di avverbi e di preposizioni.



- III. discute un problema, un dilemma, un dubbio; reagisce allo stimolo.
- IV. genera un *output* (un risultato), che può essere la ri-modulazione dell'intervento didattico nel corso dell'azione, nel qual caso il testo si configura come la narrazione di una "riflessione in azione". Oppure può essere la previsione di possibili azioni didattiche modificate dalla conoscenza acquisita (in questo caso si ha una "riflessioni sull'azione"); la soluzione ad un dilemma ecc.
- V. presenta di verbi mentali.
- VI. presenta pronomi in prima persona.

Abbiamo elaborato una matrice per indicare la presenza o assenza nei segmenti di testo degli attributi. Come mostrato nella Tab. 3, combinazioni definite di presenza o assenza delle variabili indicano l'appartenenza a una delle 4 categorie di "riflessività" elaborate.



Livelli di riflessività/ attributi	Pronomi personali	Presenza di riferimenti	Esperienza	Termini astratti	Problema	Output	Verbi mentali
Non riflessiva	0, 1	0	0,1	0,1	0	0	0, 1
Considerazione generale	0,1	1	0	1	0,1	0,1	0, 1
Riflessiva descrittiva	0,1	1	1	0,1	0,1	1,0	0, 1
Riflessiva	1	1	1	0,1	1	1	1

Tab. 3: Matrice per la classificazione manuale dei testi sulla base dello schema elaborato

Sulla base dello schema e della matrice, 3 esperti di formazione docente e un membro del gruppo di progetto hanno annotato 75 segmenti di testo. Su 44 di questi c'è stato un accordo alto (3 su 4) o totale (4 su 4). L'accordo intersoggettivo misurato sulla base del Fleiss K (una misura statistica per valutare l'affidabilità dell'accordo nell'attribuzione di categorie, tra più di due soggetti) ha dato esiti positivi (Fleiss K: 0,66). Di seguito a titolo esemplificativo, un segmento di testo di risposta alle domande accorpate 12.1.2 (Tab.1), classificato unanimemente come riflessivo. Abbiamo omesso per ragioni di *privacy* i brani con riferimenti riconoscibili all'esperienza.

Quella di [...] è stata la prima fiaba prevista dal progetto proposta ai bambini in un momento dell'anno (fine settembre - ottobre) dove non avevo ancora la piena conoscenza della sezione e dove muovevo i primi passi nella scuola [...], realtà totalmente

diversa a quella a cui ero abituata nella [...]. Quindi il procedere delle varie attività non sempre si è sviluppato come avevo previsto e progettato inizialmente. Ma, uno dei punti essenziali di una buona didattica è proprio la sua flessibilità e soprattutto la possibilità, o meglio, il dovere di poter correggere il tiro anche in itinere. Proprio per questo ho modificato le attività che avevo previsto ma che si sono rivelate poco idonee non avendo valutato la tenera età dei bimbi e le loro ancora giovani competenze: ho ingrandito disegni predisposti inizialmente troppo piccoli, sostituendo la tecnica inizialmente proposta del colore a pastello o a pennarello al riempimento delle immagini con varie tecniche manipolative [...]. Un'altra cosa che non avevo valutato bene sono stati i tempi: infatti mi sono resa conto che, nonostante la sezione abbia un'alta percentuale di presenza scolastica, essendo bambini così piccoli ci sono stati dei picchi di assenze dovuti a malattie stagionali che non mi hanno permesso di portare a termine tutte le attività progettate nel tempo previsto. Ho operato quindi una scelta prediligendo le attività dove il bimbo si riusciva ad esprimere meglio e dove c'era più partecipazione, entusiasmo e cooperazione (la narrazione con la loro partecipazione attiva, la drammatizzazione, psicomotricità musicale legata alla storia). [...]



La matrice ha funzionato come strumento di orientamento e confronto per il giudizio umano, cui in molti casi spetta di interpretare la presenza e/o di valutare le eventuali diverse gradazioni di ogni attributo, un'abilità, quest'ultima, spesso decisiva nell'attribuzione del segmento di testo a una delle 4 categorie. Abbiamo annotato manualmente in cicli diversi di annotazione oltre 600 risposte, il *corpus* annotato è stato poi classificato con strumenti di TAL.

6.2 *Classificazione automatica del livello di riflessività*

Come per la classificazione rispetto al *topic* (Par. 5.1), anche in questo caso sono stati adottati algoritmi di classificazione supervisionata basati su strumenti di TAL. Primi esperimenti in questa direzione sono descritti in Chiaritti (2018), dove gli autori riportano i risultati condotti a partire da un nucleo iniziale di risposte impiegate come *training corpus*. L'ampliamento di tale nucleo ha permesso di migliorare progressivamente le accuratezze del sistema automatico originariamente sviluppato. I risultati così raggiunti sono contenuti nelle Tabelle 4 e 5 che riportano rispettivamente l'accuratezza nella distinzione tra risposte riflessive e non riflessive, e tra i diversi gradi di "riflessività".

	Accuratezza
Risposte non riflessive	81,06
Risposte riflessive	82,01
TOTALE	81,53

Tab. 4: Accuratezza della classificazione tra segmenti di testo (risposte) riflessivi e non riflessivi

	Accuratezza
Risposta non riflessiva	59,67
Considerazione generale	64,65
Risposta riflessiva descrittiva	37,86
Risposta riflessiva	58,89
TOTALE	55,27

Tab. 5: Accuratezza della classificazione per grado di "riflessività"



Nel primo caso, abbiamo usato 642 risposte annotate a mano come *training corpus* (bilanciato rispetto al numero di risposte categorizzate come riflessive e non riflessive); nel secondo caso, il *training* era composto da 158 risposte categorizzate a mano per ognuno dei tre livelli e un numero uguale di risposte non riflessive, per un totale di 632 risposte. Come ci potevamo aspettare, il compito di classificazione in due classi era più semplice e infatti abbiamo ottenuto risultati più alti. Distinguere le risposte rispetto al livello di riflessività è invece un compito più complesso e l'accuratezza raggiunta dipende da quanto sono distinte le classi considerate.

Come introdotto nel Par. 5.1, la metodologia di classificazione automatica da noi messa a punto si basa sul testo linguisticamente annotato rispetto a più livelli di descrizione linguistica. Questo ci ha permesso di classificare il livello di riflessività delle risposte tenendo in considerazione in fase di addestramento una ricca serie di caratteristiche linguistiche. Sono state in particolare considerate caratteristiche che catturano diversi aspetti dello stile di scrittura del testo:

- caratteristiche di base che includono la lunghezza *i*) della risposta, calcolata come numero medio di parole, *ii*) della frase, calcolata come numero medio di parole, *iii*) delle parole, calcolata come numero medio di caratteri;
- caratteristiche lessicali che includono *i*) la distribuzione del lessico delle risposte rispetto al «Vocabolario di Base» redatto da Tullio De Mauro (De Mauro, 2000) e ai repertori d'uso in cui è organizzato (Fondamentale, Alto Uso, Alta disponibilità) e *ii*) la ricchezza lessicale delle risposte calcolata come *Type/token ratio*;

- caratteristiche morfosintattiche che includono *i*) la distribuzione delle categorie morfosintattiche riconosciute in fase di annotazione morfosintattica dei testi e *ii*) la distribuzione dei verbi rispetto al tempo, modo e persona in cui sono coniugati nei testi;
- caratteristiche sintattiche: che includono una serie di informazioni estratte dal testo annotato, adottando il formalismo di rappresentazione sintattica a dipendenze, che permette di ricostruire le relazioni di dipendenza sintattica tra i singoli *token* di una frase. Tali informazioni includono *i*) l'altezza dell'intero albero sintattico di una frase; *ii*) la lunghezza delle relazioni di dipendenza sintattica, calcolata in numero di *token* presenti tra la testa sintattica e la dipendenza coinvolte nella relazione; *iii*) la distribuzione dei tipi di dipendenza sintattica (es. soggetto, oggetto, modificatore); *iv*) la profondità media e la distribuzione di sequenze di complementi preposizionali che modificano un sostantivo; *v*) una serie di informazioni che riguardano il comportamento sintattico di predicati verbali (es. l'arità verbale media di una frase, cioè il numero di dipendenti di un predicato verbale, e la distribuzione di predicati con soggetto ellittico); *vi*) l'uso della subordinazione (es. la distribuzione di frasi principali e subordinate e il grado di incassamento di frasi subordinate); *vii*) l'ordine canonico degli elementi all'interno di una frase (es. la posizione di soggetto e oggetto rispetto al verbo da cui dipendono e delle subordinate rispetto alla principale).



L'analisi di quali tra le caratteristiche considerate contribuisce di più alla classificazione del grado di riflessività delle risposte fornite ha permesso di condurre uno studio sullo stile distintivo della scrittura riflessiva. Un testo riflessivo rispetto a uno non riflessivo risulta essere innanzitutto più lungo (in termini di numero di parole e frasi) e lessicalmente meno ricco, rispetto ai valori di *Type/token ratio*. È contraddistinto da caratteristiche tipicamente associate in letteratura ad una maggiore complessità sintattica, quali: uso maggiore della modificazione verbale (distribuzione maggiore di ausiliari e verbi modali), strutture verbali complesse (verbi caratterizzati da un valore di arità maggiore), uso esteso della subordinazione, ordine non canonico degli elementi della frase (distribuzione maggiore di oggetti pre-verbali e soggetti post-verbali).

Sono in particolare le frasi classificate come “riflessiva descrittiva” e “riflessiva”, ad essere caratterizzate da elementi di complessità sintattica. Sono, ad esempio, mediamente più lunghe, contengono strutture verbali complesse, incassamenti profondi di frasi subordinate ecc. Al contrario, esse contengono un lessico più semplice (una maggiore quantità

di parole contenute nel “Vocabolario di Base” e in particolare nel repertorio delle parole più frequenti e più semplici, quelle cioè “Fondamentali”) e sono caratterizzate da una minore varietà lessicale, con valori di *Type/token ratio* inferiori. Interessante infine notare come il grande uso di pronomi dimostrativi e determinativi rifletta uno degli aspetti principali della prosa riflessiva, il riferirsi cioè ad esperienze di vita diretta.

7. Conclusioni



L'integrazione di metodi manuali e automatici ha comportato molteplici e impegnative fasi di analisi per l'ottimizzazione degli strumenti, di cui l'ultima è ancora in corso, ma ha mostrato senz'altro di poter fornire indicazioni per rispondere alle domande di ricerca. La classificazione dei testi sulla base delle risposte (con un'accuratezza del 70%) ha fornito alcune conferme dei risultati delle analisi manuali (le domande “preferite” dai docenti, ad esempio). Inoltre il classificatore automatico, restituendoci i gradi di “somiglianza” tra risposte, deducibili dall'analisi degli errori di attribuzione, ci ha confermato l'opportunità di ristrutturare le prime due domande, mentre ha confutato la scelta di accorpare le altre due domande (12.3 e 12.4), azione che non ha comportato un significativo miglioramento delle prestazioni del classificatore. Anche nel caso della classificazione sulla base del “livello di riflessività” l'analisi degli errori di attribuzione ci ha fornito evidenze utili a progettare interventi per una più efficace azione di *scaffolding* con una rielaborazione della struttura e del contenuto delle domande. I bassi valori di accuratezza nell'attribuzione delle classi “riflessiva” e “riflessiva descrittiva” suggeriscono, invece, l'opportunità di fare ricorso a metodi e strumenti per l'estrazione del contenuto semantico dei testi: un'azione che potrebbe contare sulla buona *performance* del classificatore nella distinzione tra testi riflessivi e non riflessivi e quindi su un *corpus* già annotato linguisticamente a diversi livelli di granularità (frase, segmento di testo/risposta e testo), per argomento (le domande), e per categoria di “riflessività”.

Riferimenti bibliografici

- Boud D., Walker D. (1991). *Experience and learning: reflection at work*. Victoria: Deakin University.
Buckingham Shum S., Sándor Á., Goldsmith R., Bass R., McWilliams M.

- (2017). Towards reflective writing analytics: Rationale, methodology and preliminary results. *Journal of Learning Analytics*, 4, 58-84.
- Buckingham Shum S., Sándor Á, Goldsmith R., Wang X., Bass R., McWilliams M. (2016). Reflecting on reflective writing analytics: Assessment challenges and iterative evaluation of a prototype tool. *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge*. New York: ACM.
- Chiriatti G., Della Gala V., Dell'Orletta F., Montemagni S., Pettenati M.C., Saggi M.T., Venturi G. (2018). A NLP-based Analysis of Reflective Writings by Italian Teachers. *Proceedings of the Fifth Italian Conference on Computational Linguistics (CLiC-it 2018)*. Torino. <http://ceur-ws.org/Vol-2253/>
- Chung C., Pennebaker J.W. (2007). The psychological functions of function words. *Social communication*, 1, 343-359.
- Commissione europea/EACEA/Eurydice (2018). *Teaching Careers in Europe: Access, Progression and Support. Rapporto Eurydice*. Lussemburgo: Ufficio delle pubblicazioni dell'Unione europea.
- D'Amico D., Mangione G., Pettenati M.C. (2018). Governing a State-wide Induction Program: Characteristics and Success Conditions of the Italian Model. *Journal of e-Learning and Knowledge Society*, 14, 15-37.
- De Liddo A., Buckingham Shum S., Quinto I., Bachler M., Cannavacciuolo L. (2011). Discourse centric learning analytics. *LAK 2011: Proceedings of the first International Conference on Learning Analytics & Knowledge*. New York: ACM.
- Dewey J. (1933). *How we think: a restatement of the relation of reflective thinking to the educative process*. Boston: Heath.
- Kember D., McKay J., Sinclair K., Wong F.K.Y. (2008). A four-category scheme for coding and assessing the level of reflection in written work. *Assessment & Evaluation in Higher Education*, 33, 369-379.
- Kovanovic V., Joksimovic S., Gasevic D., Hatala M., Siemens G. (2017). Content Analytics: The Definition, Scope, and an Overview of Published Research. *Handbook of Learning Analytics. First Edition*. Society for Learning Analytics Research (SoLAR). <https://www.solaresearch.org/hla-17/>
- Krippendorff K. (2004). *Content Analysis: An Introduction to Its Methodology*. Thousand Oaks: Sage Publications.
- Garrison D.R., Anderson T., Archer W. (2001). Critical thinking, cognitive presence, and computer conferencing in distance education. *American Journal of Distance Education*, 15, 7-23.
- Gibson A, Kitto K. (2015). Analysing Reflective Text for Learning Analytics: An Approach Using Anomaly Recontextualisation. *Proceedings of the Fifth International Conference on Learning Analytics and Knowledge, Association for Computing Machinery*. New York: ACM.
- Liu M., Buckingham Shum B., Mantzourani E., Lucas C. (2019). Evaluating Machine Learning Approaches to Classify Pharmacy Students' Reflective Statements. *Conference proceedings International Conference on Artificial Intelligence in Education*. Chicago: AIED.
- Mezirow J. (1990). *Fostering critical reflection in adulthood: a guide to transformative and emancipatory learning*. San Francisco: Jossey-Bass Publishers.
- Mikolov T., Chen K., Corrado G., Dean J. (2013). Efficient estimation of word representations in vector space. *Proceedings of ICLR Workshops Track*. arxiv.org/abs/1301.3781

- Moon J.A. (2006). *Learning Journal: A Handbook for reflective practice and Professional Development*. New York: Routledge.
- Rodgers C. (2002). Defining Reflection: Another Look at John Dewey and Reflective Thinking. *Teachers College Record*, 104, 842-866.
- Rossi P.G., Pettenati M.C., Mangione G.R., Magnoler P., Giannandrea L., Fedeli L. (2018). Italian Induction Model Since 2014: Bridging Two Sides of Teachers Professional Development. *Proceedings of the AERA Annual Meeting April 13-17*.
- Rossi P.G., Magnoler P., Giannandrea L., Mangione G.R., Pettenati M.C., Rosa A. (2015). Il Teacher Portfolio per la formazione dei neo-assunti. *Pedagogia oggi*, 2, 223-242.
- Schön D.A. (2006). *Formare il professionista riflessivo. Per una nuova prospettiva della formazione e dell'apprendimento nelle professioni*. Milano: Franco Angeli.
- Siemens G. (2013). Learning Analytics: The Emergence of a Discipline. *American Behavioral Scientist*, 57, 1380-1400.
- Ullmann TD. (2015). *Automated detection of reflection in texts. A machine learning based approach*. PhD Thesis, The Open University. <http://oro.open.ac.uk/45402/>



Learning Analytics e bisogni speciali di apprendimento: un connubio possibile

Learning Analytics and special learning needs: a possible combination

Toto Giusi Antonia

Università di Foggia, giusy.toto@unifg.it

A theme strongly developed in the last five years is the relationship between Learning Analytics and disability (and more generally special needs). The methodological need of this new union stems from the necessity to collect a large amount of quantitative evidence with standardized tools to measure learning in special education, today very focused on qualitative analysis. These studies, pioneering, are tracing significant trajectories that describe the macro-themes in future research between Learning Analytics and special needs pedagogy: 1) Learning Analytics can help us to ascertain to what extent the accessibility of online courses contribute to inclusion and success training for disabled students; 2) Learning analytics brings out the relationship between the use of the online course and well-being, indicated by higher values in the dimension of hope supported, thanks also to the greater familiarity of students with learning disabilities towards assistive technology; 3) Dedicated Learning Analytics tools to evaluate and monitor the academic and scholastic success of students with disabilities.

If it is true that an important didactical model followed in the technologies applied to teach, it is also true that the Self directed in learning pursues the autonomy and independence of the student. This model is still far from being reached by disabled students, for who it is still assumed the presence of a tutor who supports them in learning with technology.

Keywords: Special needs; Learning Analytics; disability, ICT.

Un tema in forte sviluppo nell'ultimo quinquennio è il rapporto fra Learning Analytics e disabilità (e più in generale special needs). L'esigenza metodologica di questo inedito connubio nasce dalla necessità di raccogliere una grossa mole di evidenze quantitative mediante strumenti standardizzati per misurare gli apprendimenti nella pedagogia speciale, oggi ancora molto focalizzata su analisi qualitative. Tali studi, pionieristicamente, stanno tracciando delle traiettorie significative che descrivono i macro-temi nella ricerca futura tra Learning Analytics e pedagogia speciale: 1) il Learning analytics permetta di accertare in quale misura l'accessibilità a corsi online contribuisca all'inclusione e al successo formativo degli studenti disabili; 2) il Learning analytics fa emergere la relazione tra l'utilizzo del corso online e l'autoefficacia formativa il benessere, indicata da valori più alti nella dimensione della speranza, incrementata anche dalla maggiore familiarità degli studenti con disabilità dell'apprendimento verso la tecnologia assistiva; 3) Strumenti dedicati di Learning Analytics per valutare e monitorare il successo accademico e scolastico degli studenti con disabilità.

Se è vero che un importante modello pedagogico seguito nelle tecnologie applicate alla didattica è il Self directed in learning che persegue l'autonomia e l'indipendenza dello studente, è altrettanto vero che tale modello è ancora lontano da raggiungere per gli studenti con disabilità, per i quali si presuppone ancora la presenza di un tutor che li supporti negli apprendimenti mediati dalla tecnologia.

Parole chiave: Bisogni speciali di apprendimento; Learning Analytics; disabilità; ICT.

1. Introduzione

Secondo una recente revisione della letteratura condotta da Seale et al. (2018) relativa al rapporto tra tecnologia e disabilità, esistono principi comuni di progettazione didattica e strumentale per supportare le difficoltà di apprendimento degli studenti con bisogni educativi speciali. Le dimensioni compromesse nelle disabilità sono le più varie dalla memoria alla percezione, dalla cognizione alla comunicazione, pertanto un progetto didattico innovativo e tecnologicamente avanzato può e deve facilitare l'inclusione psico-sociale dei soggetti. I principi di design per le tecnologie didattiche sono riassunte, in questa revisione, in 4 principali categorie principali (tab. 1): 1. il *Learning Support* che si concentra sul miglioramento dell'apprendimento, 2. l'*Accessibility*, ottenuta attraverso la leggibilità, la prevedibilità e l'operazionabilità delle azioni, 3. l'*Usability*, cioè la capacità di rendere semplice da usare la tecnologia hardware e software, e 4. l'*Agency*, la capacità che rende l'utente parte attiva del proprio apprendimento.



Principi	Area di applicazione
Learning Support	Miglioramento dell'apprendimento e strumenti di help.
Accessibility	Leggibilità, la prevedibilità e l'operazionabilità delle azioni nei programmi o piattaforme di apprendimento.
Usability	Semplificazione di utilizzo della tecnologia hardware e software.
Agency	Ruolo attivo dello studente nel proprio processo di apprendimento.

Tab. 1: Principi di design per le tecnologie didattiche

Il Learning Analytics è una fonte di dati che plasmereà il futuro dell'istruzione superiore (Long, Siemens, 2011), anche nel settore della pedagogia speciale. Romero e Ventura (2010) annoverano di diritto il Learning Analytics all'interno del più ampio settore dell'Educational Data Mining (EDM), facendo emergere il legame fra i dati sugli apprendimenti e i processi educativi correlati. Secondo il celebre modello di Pope e Brandt (1997) la disabilità deriva da tre dimensioni concorrenti: l'individuo, l'ambiente e l'interazione tra individuo e ambiente. Gli individui, infatti, in relazione a fattori di rischio o abilitanti nella relazione con l'ambiente circostante hanno un diverso accesso all'inclusione sociale. Tra questi fattori, anche l'istruzione e l'accesso alle tec-

nologie determinano un supporto o una limitazione alle barriere funzionali tra individuo e ambiente, assumendo nuove difese. La dimensione strutturale dell'ambiente in generale e quella tecnologica nello specifico, dunque, può incidere positivamente o negativamente sulle dinamiche della disabilità ed essere un'interessante variabile da considerare nella progettazione educativa. Ferguson (2014), infatti riferisce la definizione del Learning Analytics riportata nella prima conferenza internazionale sul Learning Analytics and Knowledge (LAK, 2011) e adottata dalla Society for Learning Analytics Research (SoLAR), secondo la quale il Learning Analytics misura, analizza e presenta dati sugli studenti e sui loro contesti, ai fini della comprensione e dell'ottimizzazione dell'apprendimento e degli ambienti in cui ha luogo. Rispetto a questa iniziale enunciazione anche il rapporto fra il Learning Analytics e gli special needs si configura all'interno dello stesso orizzonte semantico definito da Pope e Brandt: individuo, contesto e relazioni mediate dal digitale.

A questo proposito il Learning Analytics, può svolgere un importante ruolo di strumento di monitoraggio delle scelte adottate dal soggetto in formazione e restituire un feedback sia al docente che allo studente per progettare e riparametrare il piano di lavoro. Tra le obiezioni mosse all'uso della tecnologia digitale utilizzata da alunni con bisogni speciali vi è l'eccessivo isolamento che produrrebbe. Un isolamento questo, prodotto dall'assenza del confronto fra pari e dall'interazione vis a vis con gli insegnanti, che si riflette retroattivamente anche sulla professionalità del docente priva dei necessari feedback di apprendimento da parte degli studenti (Mazza, Dimitrova, 2007). Il ruolo del docente nell'apprendimento online, dunque, sarebbe sommerso da una quantità notevole di dati che seppur quantitativamente validi, impediscono una valutazione qualitativa dell'apprendimento e della partecipazione attiva dello studente. La sfida proposta al Learning Analytics applicati ai bisogni speciali di apprendimento rimane quella di ottimizzare le possibilità di apprendimento online (Siemens, 2013), ad oggi anche in contesto europeo ancora relegata alla formazione superiore. Il concetto di Bes nella formazione online assume ancora una volta, la forma di una categoria ombrello che ingloba in sé studenti con disabilità e DSA, e i relativi studi e sperimentazioni sono ancora in fase di formulazione.



2. Stato dell'arte e metodologia della ricerca

Per sostenere e migliorare le revisioni della letteratura su *Learning analytics and special learning needs* proposte in questo contributo è stato utilizzato un approccio teorico e argomentativo. Sono stati usati i seguenti metodi per raggiungere lo scopo della ricerca e ottenere dati oggettivi: analisi teorica e generalizzazione dei dati della letteratura scientifica e utilizzo di metodi statistici. Una revisione sistematica della letteratura esistente è stata condotta per la ricerca di articoli nei principali database internazionali (Google Scholar, PsycInfo e Scopus) sul tema del Learning Analytics and special needs, usando solo le parole chiave in inglese e italiano “Learning Analytics”, “special needs”, “disability” e “ICT”. Inoltre, alla ricerca bibliografica ha fatto seguito una valutazione dei titoli dei contributi e degli abstract basata sull'idea di ricerca che gli studi sul Learning Analytics in questi ultimi anni si stiano indirizzando verso i bisogni specifici di apprendimento, le disabilità e, in generale verso la pedagogia speciale (da un punto di vista metodologico è stato usato il modello di revisione di Moher et al., 2009). La letteratura scientifica internazionale suggerisce che la disponibilità di una grossa mole di studi possa dare maggiore linfa alla ricerca applicata alle disabilità soprattutto legata agli ambienti di apprendimento virtuali, alla formazione a distanza e al Life Long Learning (identificazione) (Kuzilek et al., 2017). Per quanto riguarda i criteri di inclusione ed esclusione, gli articoli sono stati selezionati in riviste peer review, libri o capitoli di libri in inglese o italiano che mirano a descrivere o valutare dimensioni e variabili espresse nella suddetta idea di ricerca (screening). Tutte le pubblicazioni riguardanti specificatamente la tecnologia assistiva e compensativa e le pubblicazioni il cui formato completo (pertinenza) non era stato trovato sono stati esclusi. Il tempo limite per l'anno di pubblicazione è stato fissato agli ultimi 10 anni, quindi gli articoli sono stati selezionati dal 2009. Per l'inclusione dei contributi, è stato anche fornito un riepilogo qualitativo delle informazioni più pertinenti condotto con confronti tra varie pubblicazioni senza effettuare un'analisi quantitativa nel formato di meta-analisi. Dopo l'eliminazione dei duplicati, la ricerca ha identificato i macro-filoni di ricerca che saranno trattati in modo esaustivo nella seguente sezione “risultati”.



3. Risultati

Questa sezione presenta i principali contenuti tematici della letteratura sull'argomento. Da questa analisi emergono i macro-temi nella ricerca futura tra il Learning Analytics e la pedagogia speciale: 1) il Learning

Analytics permette di accertare in quale misura l'accessibilità dei corsi online contribuisca all'inclusione e al successo formativo degli studenti con disabilità; 2) il Learning analytics fa emergere la relazione tra l'utilizzo del corso online e l'efficacia formativa, indicata da valori più alti nella dimensione della speranza, incrementata anche dalla maggiore familiarità degli studenti con disabilità dell'apprendimento verso la tecnologia assistiva; 3) gli strumenti dedicati al Learning Analytics per valutare e monitorare il successo accademico e scolastico degli studenti con disabilità. Fra gli studi si selezionano tre sperimentazioni di software che riportano importanti risultati in termini di efficace sinergia fra bisogni speciali e Learning Analytics.

Dall'analisi di ulteriori studi (Ifenthaler, 2018), emergono tre principali prospettive di progettazione del Learning Analytics design, fase preliminare alla fruizione dei corsi in modalità e-learning: 1) la prospettiva sommativa (cioè alla fine del corso) della progettazione è una metodologia che riflette sui risultati, essa infatti, misura l'impatto degli interventi, i progressi e i risultati in termini di obiettivi di apprendimento. Il modello pedagogico su cui si basa questa prospettiva mira a descrivere l'allineamento o il disallineamento tra intervento e comportamento dello studente; 2) la prospettiva simultanea o real-time (durante lo svolgimento del corso) permette di fornire risorse, aiuto e supporto agli studenti durante il corso online. Questo monitoraggio *in itinere* permette di individualizzare gli apprendimenti riparametrando compiti e fornendo materiali adattati ai risultati raggiunti; 3) la prospettiva predittiva infine, attraverso le caratteristiche e i comportamenti degli studenti e agli algoritmi relativi ai percorsi di apprendimento è possibile prevedere bisogni, scelte e difficoltà degli utenti e ideare strumenti per superarli.

Una prima sperimentazione (Heiman, Shemesh, 2012) ha riguardato studenti universitari con difficoltà di apprendimento (DSA) iscritti a corsi online, ha dimostrato come questi studenti accedono più spesso ai siti del corso, e consultano e partecipano ai forum tra pari. Lo studio dimostra altresì, che tali studenti hanno più familiarità con la tecnologia assistiva rispetto al gruppo di controllo costituito da studenti senza difficoltà di apprendimento. Inoltre, il monitoraggio degli studenti con learning disabilities ha rilevato che essi raggiungono punteggi più alti nella scala della speranza (Hope scale), hanno una forte spinta motivazionale e desiderio di trovare percorsi diversi per raggiungere i loro obiettivi formativi e il loro benessere soggettivo è superiore rispetto a quello degli altri studenti.

Un secondo esempio è lo sviluppo del software *PADA: a Dashboard of Learning Analytics for University Students with Dyslexia*, prodotto a



partire dal 2012 da Mejia Corredor et al. (2012) per studenti dislessici adulti per supportare la comprensione, adattarsi agli stili di apprendimento, migliorare le difficoltà di lettura e monitorare l'autoregolazione. Il software nella prima versione ha prodotto risultati incoraggianti che hanno reso necessaria una revisione per renderlo maggiormente funzionale agli scopi della ricerca (Mejia Corredor et al., 2015). Esso risulta costituito da 1) un database per raccogliere dati sulle attività degli utenti (ad es. demografia) e sui loro risultati, e una serie di strumenti fra cui l' ADDA (per le difficoltà di lettura) e l'ADEA (per il riconoscimento degli stili di apprendimento), e lo strumento BEDA (per i deficit cognitivi); 2) gli aggregatori, per correlare i dati dal database al piano sociale di apprendimento (es. studente, coetanei o classe) e il livello richiesto (ad es. studente, insegnante o esperto); 3) l'estrazione dei dati e l'analisi statistica; 4) il Learning Analytics che ingloba overview analytics, reading difficulties analytics, learning styles analytics, e cognitive processes analytics, e 5) un server web che memorizza i dati analizzati e permette la comunicazione tra gli utenti e PADA tramite un browser (fig. 1).

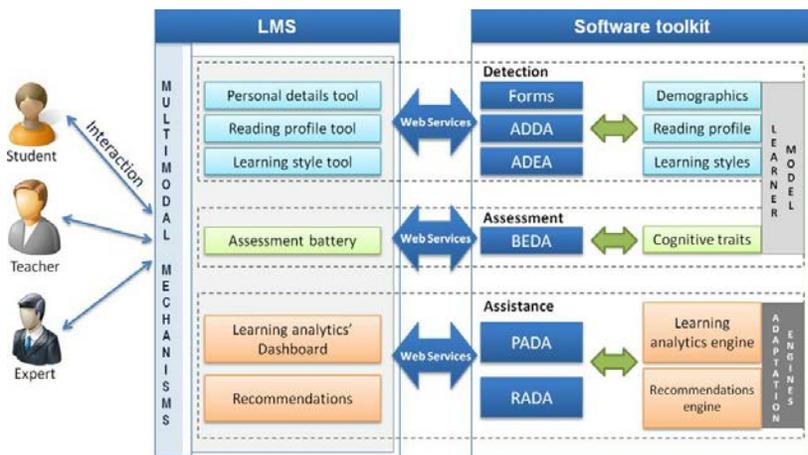


Fig. 1: Framework components, fonte: Mejia Corredor, 2013

Un'interessante meta-analisi condotta da Tsikinas e Xinogalos, (2019), su database internazionali nel periodo 2005-2018, mostra che i 104 studi inclusi analizzano il rapporto fra Serious game e disabilità intellettive e autismo. Gli studi mostrano che i Serious game e le attività di simulazione in generale, sono progettati per migliorare le competenze sociali e comunicative negli studenti autistici e le capacità concettuali e cognitive negli alunni con disabilità intellettive. Pur

riportando risultati positivi in termini di successo del processo di apprendimento, gli autori segnalano che sono assenti ricerche sull'acquisizione di capacità pratica da utilizzare nel quotidiano. È ormai ampia la tradizione di studi che asserisce il potenziale educativo dei giochi digitali sugli allievi con disabilità e come le conoscenze costruite nel mondo virtuale possano essere trasferite in abilità funzionali al mondo reale. Pertanto, il progetto di Torrente ed altri (2012) rappresenta il terzo esempio di questa proficua relazione e propone due serious game per studenti con disabilità con esplicite finalità di inclusione professionale. All'interno di questa sperimentazione inoltre, i docenti attraverso uno strumento definito eAdventure possono monitorare e valutare le attività degli studenti; eAdventure supporta lo sviluppo di giochi di avventura *punta e clicca*, personalizzandoli attraverso un aumento o diminuzione della complessità. I due giochi educativi proposti all'interno del progetto "Il mio primo giorno di lavoro" e "la mia prima festa" mirano a facilitare l'inclusione di un lavoratore con disabilità cognitiva in una nuova azienda. Nel primo gioco il lavoratore al suo primo impiego deve acquisire competenze e abilità (utilizzo attrezzatura, messaggistica, ecc.) necessarie per svolgere il proprio lavoro. Nel secondo il lavoratore deve partecipare a una cena sociale organizzata dalla compagnia per cui lavora, risolvendo una serie di problemi dall'abbigliamento, al trasporto fino alle regole di comportamento durante un party.



4. Discussioni

La definizione di Siemens (2010, p. 1) indirizza *Il Learning Analytics verso l'uso di dati intelligenti, di dati prodotti dallo studente e di modelli di analisi per scoprire informazioni e connessioni sociali, e per predire e dare consigli sull'apprendimento*. Nella fase di progettazione, ad esempio, degli Educational Game per disabilità intellettive il percorso ideativo deve essere compiuto a ritroso, cioè bisogna identificare un profilo ideale dell'utente poi sperimentarlo nell'interazione uno a uno con il tutor e, infine, generalizzarlo ad un numero considerevole di utenti (Cano et al., 2016). A questo proposito la tabella 1 esemplifica la prima fase di strutturazione delle caratteristiche di un utente-tipo con disabilità intellettive al quale proporre un'attività di apprendimento online.

Intelligence, Memory and Perception
- Visual perception and visual retention better than hearing
- Procedural memory good enough for playing
- Short Memory limitation. Limitation in the number of instructions/numbers that they can
Handle at the same time (3 sequential instructions and 3–6 digits)
- Difficulty sustaining attention during long periods of time
- Mild intellectual deficiency
- Problems of understanding the information
- Difficulty in the process of abstraction, conceptualization, generalization and learning transfer
Personality
- Limited initiative
- Persistence of behavior and resistance to change
- Few response capacity and reaction to the environment
- Poor social and collaboration skills (depending on the disability)
Biological and Motor Skills
- Listening and sight problems
- Heart diseases
- Clumsiness in motor skills (gross and fine)
- Poor coordination

Tab. 1: Caratteristiche dell'utente con disabilità intellettive relative all'intelligenza, alla memoria, alla percezione, personalità, abilità motorie. Fonte: Cano et al., 2016

La contemporanea progettazione di game per l'educazione non si concentra soltanto sul supporto al singolo utente, ma anche su sessioni di apprendimento collettivo, anche quest'ultima una ricca fonte di dati per codificare stili e risultati degli apprendimenti. L'analisi dei dati può fornire spunti di riflessione sul fallimento e sul successo degli utenti ed essere la base per la realizzazione di un apprendimento predittivo, infatti, per elaborare una previsione sui risultati di apprendimento si utilizzano tecniche statistiche e analisi storiche di dati. Tale processo è necessario per raccogliere dati su una vasta popolazione di utenti, un'operazione resasi complessa per la specificità e singolarità delle caratteristiche degli stessi (Sedkaoui, 2018). Secondo la ricerca condotta da Ebner e Schön (2013) dell'Università di Graz sull'uso dei Learning Analytics applicati alla scuola primaria si propone uno schema sui requisiti che un software deve possedere per essere funzionale agli ambienti di apprendimento formali contemporanei: 1) il sistema deve valutare il grado di competenza dello studente; 2) il software deve essere

in grado di registrare e archiviare i dati di tutto ciò che è stato fatto, ad esempio esercitazioni, risultati dei test e competenze raggiunte dallo studente per avviarlo alla successiva sessione di studio. L'obiettivo è generare un database completo per monitorare la competenza degli studenti su ogni singolo compito e fornire informazioni agli insegnanti; 3) il sistema deve fornire compiti appropriati e in successione secondo il grado di competenza raggiunta dallo studente; 4) il sistema, ancora, deve garantire che gli esercizi conclusi correttamente possano essere ripetuti e praticati continuamente (rafforzamento dell'apprendimento); 5) il sistema in generale, deve motivare e rendere l'apprendimento ricreativo. I compiti, tuttavia devono comunque essere impegnativi e progressivamente più complessi; 6) la strutturazione di questo sistema può liberare gli insegnanti da alcune incombenze amministrative; 7) da un punto di vista tecnico, infine, il sistema deve funzionare sul web e su molti client. Il sistema deve essere fruibile, inoltre, sui dispositivi mobili (iPhone, iPad, Android, ecc.) diffusi tra gli studenti.

Il problema dell'accessibilità degli studenti con disabilità è una questione che riguarda tutto il settore della formazione online, non soltanto il Learning Analytics; secondo Hashey e Stahl (2014), in questa fase di profondo cambiamento e adattamento del mondo educativo virtuale, riveste un ruolo principale l'insegnante che ha la responsabilità di programmare obiettivi educativi personalizzati e fornire strumenti accessibili grazie alla tecnologia digitale disponibile (Office, Google suite, libri digitali ecc.). L'emergenza educativa è così incombente nel mondo dell'educazione speciale, che non riguarda più e non solo i contesti formali dell'istruzione, ma si è trasferita alla formazione professionale e all'esterno dell'ambiente scolastico.

5. Conclusioni

All'interno del Learning Analytics si stanno susseguendo una serie di sotto-specializzazioni, due delle quali fortemente correlate ai bisogni speciali dell'apprendimento: il Social Learning Analytics e la Disposition Analytics. La prima rappresenta la lettura dei dati in relazione agli studi sulla formazione permanente in una molteplicità di contesti; la seconda specializzazione, invece, si concentra sull'esperienza, sulla motivazione e sulle intelligenze che influenzano le risposte agli stimoli apprenditivi di uno studente in relazione con il proprio tutor (Shum, Ferguson, 2012). Come ben espone Trentin (2019) sono molteplici le cause che impediscono l'accesso alla formazione scolastica per gli studenti con disabilità. Esse infatti, possono essere riassunte in tre princi-



pali categorie: disabilità fisico-sensoriali, disabilità cognitive/disagio psico-sociale, gravi patologie che impediscono l'accesso alla classe, per le quali è necessario studiare soluzioni atte a favorire la partecipazione alla vita scolastica.

Un apprendimento significativo è uno dei principali obiettivi di Learning Analytics in ambienti di apprendimento tecnologici, con un alto grado di autoregolazione. Lo studio autoregolato offre maggiori possibilità didattiche ai diversi stili di apprendimento (Leder Müller, Fallmann, 2017). Tale apprendimento secondo la concezione costruttivista è un processo attivo, produttivo, emotivo e autoregolato per l'ap-punto. L'acquisizione delle conoscenze non è affidata al docente, ma direttamente all'azione attiva degli studenti (Toto, Limone, 2019). Il modello pedagogico maggiormente prescritto nelle tecnologie applicate alla didattica è quello del Self directed in learning che persegue l'auto-nomia e l'indipendenza dello studente; tale modello è ancora lontano da raggiungere per gli studenti con disabilità, per i quali si presuppone ancora la presenza di un tutor che li supporti negli apprendimenti attraverso la tecnologia.

Un ulteriore settore di sviluppo fra Learning Analytics, bisogni speciali e formazione online è quello dei Serious game, cioè giochi digitali progettati per il divertimento e l'educazione sono ormai ampiamenti utilizzati nei settori più disparati, dalla salute all'economia finanche nell'istruzione. È proprio questa un'emergenza pedagogica da affrontare per le persone con disabilità, cioè il trasferimento di competenze e abilità trasversali anche ad altri ambiti di applicazione oltre a quello scolastico. In contesto educativo questo tipo di game sono spesso utilizzati soprattutto nei contesti formali dell'educazione e solo negli ultimi anni sono stati progettati specifici game per ragazzi con disabilità, soprattutto intellettive e affetti dallo spettro autistico (Bernardini, Porayska Pomsta, Smith, 2014). Lo scopo dei serious game per la disabilità è simulare e addestrare a comportamenti adattivi e a competenze funzionali. Nello specifico, si intendono sviluppare: 1) abilità concettuali quali l'alfabetizzazione, l'uso del denaro, la cognizione del tempo e l'autodirezione; 2) abilità sociali come la relazione interpersonale, l'auto-stima, la responsabilità sociale, il rispetto delle regole ecc., e inoltre, 3) abilità pratiche e quotidiane relativamente alla routine, sicurezza, benessere, trasporti, uso della tecnologia, ecc.

Secondo Cooper, Ferguson e Wolff (2016), infine, l'accessibilità è un tema critico nel rapporto tra disabilità e ambienti di apprendimento virtuale, molto trascurato dalla ricerca scientifica contemporanea. Gli studiosi definiscono l'accessibilità in relazione alla disabilità come: "Web accessibility means that people with disabilities can perceive, un-



derstand, navigate, and interact with the Web, and that they can contribute to the Web” (p. 100). Questi autori hanno dimostrato nelle loro ricerche che gli studenti con disabilità anche se raggiungono percentuali inferiori di completamento dei compiti rispetto ai coetanei, il divario aumenta quando la variabile accessibilità compromette lo svolgimento del percorso formativo online. La sfida per la progettazione futura è quella di ridurre questo divario attraverso un rafforzamento del supporto e dell’accessibilità mediante l’apporto dei dati rivenienti dal Learning Analytics.

Riferimenti bibliografici

- Bernardini S., Porayska Pomsta K., Smith T. J. (2014). ECHOES: An intelligent serious game for fostering social communication in children with autism. *Information Sciences*, 264, 41-60.
- Cano A.R., Fernández-Manjón B., García-Tejedor Á. J. (2016). GLAID: Designing a game learning analytics model to analyze the learning process in users with intellectual disabilities. In *International Conference on Serious Games, Interaction, and Simulation* (pp. 45-52). Cham: Springer.
- Cooper M., Ferguson R., Wolff A. (2016). What Can Analytics Contribute to Accessibility in e-Learning Systems and to Disabled Students’ Learning? In *6th International Learning Analytics and Knowledge (LAK) Conference*, ACM, (pp. 99-103).
- Designing serious games for adult students with cognitive disabilities. In *International Conference on Neural Information Processing* (pp. 603-610). Berlin, Heidelberg: Springer.
- Ebner M., Schön M. (2013). Why learning analytics in primary education matters. *Bulletin of the Technical Committee on Learning Technology*, 15(2), 14-17.
- Ferguson R. (2014). Learning Analytics: fattori trainanti, sviluppi e sfide. *TD Tecnologie didattiche*, 22 (3), 138-147.
- Hashey A.I., Stahl S. (2014). Making online learning accessible for students with disabilities. *Teaching exceptional children*, 46(5), 70-78.
- Heiman T., Shemesh, D.O. (2012). Students with learning disabilities in higher education: Use and contribution of assistive technologies and website courses and their correlation to students' hope and well-being. *Journal of Learning Disabilities*, 45(4), 308-18.
- Ifenthaler D. (2018). Learning analytics design. In L. Lin, M. Spector (eds.), *The sciences of learning and instructional design: Constructive articulation between communities* (pp. 202-211), New York and London: Routledge.
- Kuzilek J., Hlostá M., Zdrahal Z. (2017). Open university learning analytics dataset. *Scientific data*, 4, 170-171.
- Ledermüller K., Fallmann I. (2017). Predicting learning success in online learning environments: Self-regulated learning, prior knowledge and repetition. *Zeitschrift für Hochschulentwicklung*, 12(1), 79-99.
- Long P., Siemens G. (2011). Penetrating the Fog: Analytics in Learning and Edu-



- cation. *EDUCAUSE Review Magazine*, 46(5), 31-40.
- Mazza R., Dimitrova V. (2007). CourseVis: A graphical student monitoring tool for supporting instructors in web-based distance courses. *International Journal of Human-Computer Studies*, 65(2), 125-139.
- Mejía Corredor C. (2013). *Framework for detection, assessment and assistance of university students with dyslexia and/or reading difficulties*. Doctoral Thesis.
- Mejía Corredor C., Fábregatb Gesa R., Salas Álvarez D. (2015). Integration of a framework with a learning management system for detection, assessment and assistance of university students with reading difficulties. *Revista EAN*, 79, 98-115.
- Mejía C., Bull S., Vatrappu R., Florian B., Fabregat R. (2012). PADA: a Dashboard of Learning Analytics for University Students with Dyslexia. *Last ScandLE Seminar*, 3-5.
- Moher D., Liberati A., Tetzlaff J., Altman D.G., Prisma Group (2010). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *International Journal of Surgery*, 8(5), 336-341.
- Pope A.M., Brandt Jr, E.N. (Eds.). (1997). *Enabling America: Assessing the role of rehabilitation science and engineering*. Washington, DC: National Academies Press.
- Romero C., Ventura S. (2010). Educational Data Mining: A Review of the State of the Art. *IEEE Transactions on Systems, Man, and Cybernetics – Part C. Applications and Reviews*, 40(6), 601-618.
- Seale J., Garcia-Carrizosa H., Rix J., Sheehy K., Hayhoe S. (2018). A proposal for a unified framework for the design of technologies for people with learning difficulties. *Technology and Disability*, 30(1-2), 25-40.
- Sedkaoui S. (2018). *Data analytics and big data*. London: John Wiley & Sons.
- Serrano-Lagunaa Á., Torrente J., Maneroa B., del Blancoa Á., Borro-Escribanoa B., Martínez-Ortiz I., ... Fernández-Manjóna B. (2013, October). Learning analytics and educational games: lessons learned from practical experience. In *International Conference on Games and Learning Alliance* (pp. 16-28). Cham: Springer.
- Shum S.B., Ferguson R. (2012). Social learning analytics. *Journal of educational technology & society*, 15(3), 3-26.
- Siemens G. (2013). Learning Analytics: The Emergence of a Discipline. *American Behavioural Scientist*, 57(10), 1380-1400.
- Siemens G. (2010). What are learning analytics. *Retrieved March, 10*, 1-8.
- Torrente J., Del Blanco Á., Moreno-Ger P., Fernández-Manjón B. (2012, November).
- Toto, G. A., Limone P. (2019). L'evoluzione epistemologica del Self Direction in learning tra esperienze empiriche e formulazioni teoriche. *Formazione, Lavoro, Persona*, 9(26), 20-25.
- Trentin G. (2019). Tecnologie e inclusione: come fare di necessità virtù. In P.C. Rivoltella, P.G. Rossi (eds), *Tecnologie per l'educazione* (pp. 57-68). Milano: Pearson Italia.
- Tsikinas S., Xinogalos S. (2019). Studying the effects of computer serious games on people with intellectual disabilities or autism spectrum disorder: A systematic literature review. *Journal of Computer Assisted Learning*, 35(1), 61-73.

Promuovere il successo formativo: quale il contributo del Learning Analytics per la ricerca educativa?

Promote educational success: what is the contribution of Learning Analytics to educational research?

Luisa Pandolfi

Università degli Studi di Sassari
lupandolfi@uniss.it

The notion of educational success is broad and multidimensional and refers to the full and authentic personal and social realization of students. The European Council in 2018 suggest the importance of providing support to all students for the development of their skills and resources, including those in disadvantaged conditions or with specific needs. But in Europe and, specifically, in Italy can we talk about educational success or failure? The rates of early school leaving are still very high; the scientific literature focuses in adopting systemic models and interpretations of the phenomenon. The importance of having data and observable and measurable indicators emerges and, at the same time, to bring out the perception and the meaning attributed by the protagonists to their experience, paying attention to the contextual dimensions. The paper explores some points of intersection between the strategies of intervention and contrast to the early school leaving, the educational research and the current challenges of Learning Analytics, in order to outline possible directions of analysis and empirical investigation to improve the learning processes and teaching, educational practices and contexts.

Keywords: educational success; evaluation; improvement; educational research; learning analytics.

La nozione di successo formativo è ampia e multidimensionale e fa riferimento alla piena ed autentica realizzazione personale e sociale degli studenti. Il Consiglio Europeo nel 2018 ribadisce l'importanza di fornire sostegno a tutti gli studenti per lo sviluppo delle proprie competenze e delle proprie risorse, compresi quelli in condizioni svantaggiate o con bisogni specifici. Ma in Europa e, nello specifico, in Italia possiamo parlare di successo o insuccesso formativo? I tassi di dispersione scolastica sono tuttora molto elevati; la letteratura scientifica converge nell'adottare modelli e chiavi di lettura del fenomeno di tipo sistemico. Emerge l'importanza di poter disporre di dati e di indicatori osservabili e misurabili e, allo stesso tempo, di far emergere la percezione ed il significato attribuito dai protagonisti alla loro esperienza, prestando attenzione alle dimensioni contestuali. Il contributo esplora alcuni punti di intersezione tra le strategie di intervento e di contrasto all'abbandono scolastico, la ricerca educativa e le attuali sfide del Learning Analytics, al fine di tracciare possibili direzioni di analisi e di indagine empirica per migliorare i processi di apprendimento e insegnamento, le pratiche ed i contesti educativi.

Parole chiave: successo formativo; valutazione; miglioramento; ricerca educativa; learning analytics.



217

1. Successo o fallimento formativo?



Il recente dossier *L'Autonomia Scolastica per il Successo Formativo* (Miur, 2018)¹, realizzato dal gruppo di lavoro istituito presso il Dipartimento per il sistema educativo d'istruzione e formazione, individua, sia in ambito organizzativo che metodologico-didattico, strategie di innovazione, ricerca e sperimentazione dell'autonomia scolastica finalizzate alla promozione del successo formativo. In particolare, il documento mette in luce l'importanza dell'autonomia scolastica per rispondere alle esigenze educative con strumenti flessibili di progettazione organizzativa e didattica, evidenziando il ruolo centrale della valutazione formativa, al fine di garantire percorsi inclusivi, basati sul raggiungimento di risultati di apprendimento adeguati alle potenzialità di ognuno. Elementi altrettanto rilevanti riguardano la relazione educativa ed il clima della classe, soprattutto nel punto in cui si sottolinea come la qualità della gestione della classe non corrisponde prioritariamente con il controllo della disciplina, ma «comprende tutto ciò che i docenti possono realizzare per promuovere interesse e partecipazione e soprattutto il riconoscimento dell'altro come persona» (Ivi, p. 8). In tale prospettiva, si può notare come attualmente il concetto di successo formativo abbia integrato e quasi sostituito quello di successo scolastico, maggiormente focalizzato sul percorso di studi e sul raggiungimento di una adeguata preparazione scolastica. La nozione di successo formativo è, invece, più ampia e multidimensionale e fa riferimento alla piena ed autentica realizzazione personale e sociale, in base a cui, come sottolineano Lisimberti e Montalbetti (2015, p. 80): «l'impegno e la fatica scolastica assumono significato non solo ai fini di un voto ma in un cammino di crescita personale e di progressiva costruzione di un progetto di sé». In tal senso, anche la raccomandazione del 22 maggio 2018 del Consiglio d'Europa, relativa alle competenze chiave per l'apprendimento permanente, ribadisce l'importanza di fornire sostegno a tutti i discenti, compresi quelli in condizioni svantaggiate o con bisogni specifici, affinché esprimano appieno le proprie potenzialità, al fine di migliorare la capacità di gestire la propria vita in modo attento alla salute e orientato al futuro.

Alla luce di tali considerazioni, il titolo del presente paragrafo è volutamente provocatorio, in quanto intende soffermarsi sugli ostacoli

1 Reperibile al seguente link: <http://istruzioneer.gov.it/wp-content/uploads/2018/09/lautonomia-scolastica-per-il-successo-formativo.pdf>

che la scuola ancora incontra nella concreta realizzazione del successo formativo.

Il dossier pubblicato da Tuttoscuola nel 2018² riporta, a tal proposito, dati allarmanti: dal 1995 al 2013-14 sono 3 milioni e mezzo i ragazzi italiani iscritti alle scuole superiori statali che non hanno completato il corso di studi, ossia il 30,6% degli oltre 11 milioni di studenti (11.430.218) che si erano iscritti in questo arco di tempo alle scuole superiori statali. I dati rilevati riguardano solo la scuola statale, in quanto non esistono dati pubblici che consentano di sapere quanti studenti tra quelli che abbandonano la scuola statale siano transitati nella scuola paritaria o non statale o siano passati alla formazione professionale. L'indicatore empirico utilizzato da Tuttoscuola si basa sul numero (e la percentuale) di abbandoni rispetto al complesso degli iscritti al primo anno di corso della scuola secondaria statale, facendo riferimento ai dati messi a disposizione dal Miur. Anche dando uno sguardo agli ultimi anni la situazione non migliora, infatti si rileva che dal 2013-14 al 2017-18, al primo anno gli studenti iscritti erano 612.675, mentre cinque anni dopo, al 5° anno, erano scesi a 461.120 unità: lungo il percorso hanno abbandonato anzitempo 151.555 studenti, cioè il 24,74% di quelli che erano partiti cinque anni prima. Praticamente ha abbandonato uno studente ogni quattro' (Ivi, p. 28). Oltre la metà degli studenti abbandona la scuola dopo il primo biennio.

Emerge, dunque, come il fenomeno della dispersione scolastica in Italia sia strutturale, con alcune differenze rispetto alle tipologie di scuola (il tasso più elevato si registra negli istituti professionali e a seguire in quelli tecnici) e alle aree territoriali (il 29% nelle isole, di cui il 33% in Sardegna, ma sopra la media nazionale si collocano anche regioni come la Toscana e la Lombardia).

La letteratura scientifica sul tema (Bølling et al., 2018; Alistair & Leathwood, 2013; Batini, 2014; Batini & Bartolucci, 2016; Colombo, 2010; 2015; Scales, 2015) converge nell'adottare modelli e chiavi di lettura del fenomeno di tipo sistemico, in quanto al di là delle caratteristiche oggettive degli indicatori a cui si fa riferimento³, le dinamiche



2 Reperibile al seguente link: <https://www.tuttoscuola.com/content/uploads/2018/09/La-scuola-colabrodo-Tuttoscuola-settembre-18.pdf>

3 L'indicatore Early School Leavers (il cui acronimo è ESL) è stato proposto dall'Eurostat ed è utilizzato in ambito europeo per l'analisi del fenomeno e come base di confronto tra i vari Paesi dell'UE; si riferisce alla quota dei giovani di età compresa fra i 18 e i 24 anni, in possesso della sola licenza media e che si trovano al di fuori dal sistema nazionale di istruzione e da quello regionale di istruzione e formazione professionale.

e le interazioni fra i molteplici elementi che entrano in gioco a più livelli sono differenti e variegati ed investono non solo il contesto scolastico e formativo (che includono l'organizzazione ed il clima scolastico, le modalità di svolgimento della didattica, la qualità delle relazioni interpersonali, le ripetenze), ma anche i fattori relativi alle caratteristiche psicologiche degli studenti (presenza di eventuali difficoltà di apprendimento e disabilità, di bisogni educativi speciali, un atteggiamento di rifiuto verso la scuola, il disimpegno, l'auto-percezione di inadeguatezza) e i fattori familiari e socio-economici (l'appartenenza a una categoria socialmente ed economicamente svantaggiata, uno scarso livello di istruzione dei genitori e di coinvolgimento di questi ultimi nella vita scolastica dei figli, l'instabilità del nucleo familiare).



2. Politiche e misure di intervento: la necessità di dati e di valutazione degli esiti

A livello europeo, la Raccomandazione 2011/C 191/01 del Consiglio dell'Unione Europea sulle politiche di riduzione dell'abbandono scolastico evidenzia la necessità di disporre di dati aggiornati sul fenomeno (tassi di abbandono, passaggi tra livelli di istruzione, tassi di iscrizione e di completamento nell'istruzione secondaria superiore, assenteismo, elusione dell'obbligo scolastico) che permettano un'analisi a livello locale, regionale e nazionale, al fine di definire con più precisione gli interventi da adottare, valutando, al contempo, l'efficacia e l'efficienza delle misure già adottate per migliorare le strategie e i programmi finanziati e/o da implementare. In tale direzione, sul piano delle politiche pubbliche diventa sempre più rilevante l'esigenza di capire, documentare ed ottimizzare gli effetti delle risorse finanziarie investite (Martini & Trivellato, 2011). Allo stesso tempo, la valutazione consente di produrre benefici non solo per i decisori politici, ma per tutti i vari attori coinvolti (tra cui docenti e studenti), in termini di apprendimento, miglioramento professionale, diffusione di metodologie e buone pratiche (Fraccaroli & Vergani, 2004).

Tali considerazioni assumono una particolare rilevanza nell'ambito delle strategie di azione impegnati nella lotta all'abbandono scolastico; ambito in cui sono stati molteplici i finanziamenti impegnati negli ultimi decenni nei tre assi principali indicati dal Consiglio d'Europa (2011):

- *prevenzione*: politiche e misure tese ad affrontare i problemi strutturali ed i fattori di rischio che possono causare l'abbandono precoce

e volte a ottimizzare l'offerta di istruzione e formazione per migliorare le prospettive di successo formativo;

- *intervento*: politiche e misure che rispondono ai primi segnali dell'abbandono scolastico, con la finalità di migliorare la qualità dell'istruzione e della formazione (dalla scuola della prima infanzia al secondo ciclo dell'istruzione secondaria) e di offrire un supporto mirato alle difficoltà incontrate dagli studenti a rischio;
- *compensazione*: politiche e misure rivolte a coloro che hanno abbandonato precocemente i percorsi di istruzione, offrendo loro nuove opportunità per ottenere una qualifica (ad esempio attraverso corsi di recupero, programmi d'istruzione di «seconda opportunità», supporto individuale specifico, ecc.).

In merito alle tre tipologie di politiche e misure di intervento indicate, sono state raccolte diverse evidenze empiriche per capire quali siano le effettive ricadute delle politiche pubbliche e degli interventi realizzati, mentre sul piano metodologico si rileva il ricorso ad indicatori di risultato, che non esclude, però, l'attenzione verso le dimensioni interpersonali, didattiche ed educative ed il significato che i percorsi hanno avuto per gli attori coinvolti (Araújo H.C. et al., 2014; Commissione europea, 2013).

Emerge l'importanza di poter disporre di indicatori osservabili e misurabili, al fine di rilevare i risultati raggiunti tramite procedure trasparenti e ripetibili, ma, ricostruendo, contemporaneamente, la visione, la percezione ed il significato attribuito dai protagonisti alla loro esperienza, prestando attenzione alle dimensioni contestuali e stimolando processi di autovalutazione e di autoriflessione mediante approcci di ricerca collaborativi e partecipativi (Lotti & Pedani, 2016; Pandolfi, 2016).

In tale prospettiva, è interessante il 'toolkit' per le scuole (Jasińska-Maciążek & Tomaszewska-Pękała, 2017) finalizzato a supportare a livello operativo docenti e professionisti educativi nell'identificazione e monitoraggio dei fattori di rischio di insuccesso e abbandono scolastico, elaborato nell'ambito del programma europeo 'Reducing early school leavers in Europe'⁴.

4 Il progetto ha previsto la partecipazione di nove Paesi membri dell'EU (Belgio, Regno Unito, Svezia, Portogallo, Olanda, Polonia, Spagna, Ungheria e Austria) ed ha realizzato sia una raccolta dati quantitativa su un campione di circa 1.500 studenti per ciascun Paese ed un'analisi qualitativa mediante interviste e focus group che hanno coinvolto studenti e professionisti. L'obiettivo era quello di sviluppare ed implementare politiche educative efficaci, anche mediante la trasferibilità delle buone pratiche specifiche rilevate in ciascun Paese.

Il toolkit, nello specifico, fornisce alcune indicazioni per la realizzazione del processo di raccolta e analisi sistematica dei dati; processo che si articola in sei fasi che rispondono a sei domande guida principali, come raffigurato nel seguente schema:

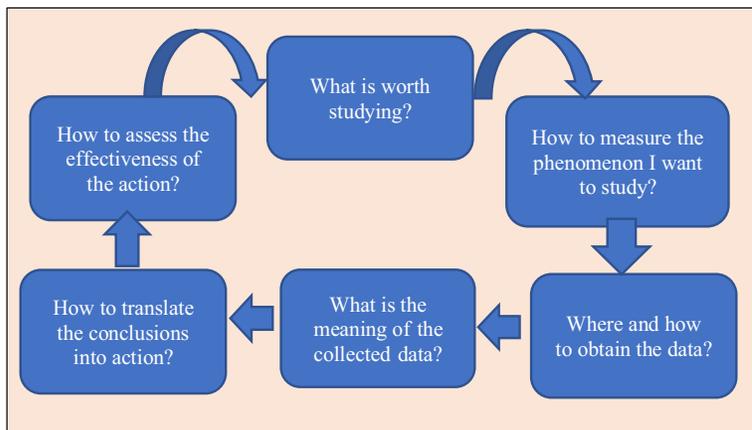


Fig. 1: Il processo di raccolta ed analisi sistematica dei dati (Jasińska-Maciązek & Tomaszewska-Pękała, 2017, p. 11)

Come si può notare, il processo inizia con una riflessione sugli elementi informativi che consentono di valutare quali studenti richiedono maggiore attenzione e supporto e, quindi, di individuare i principali indicatori predittivi di rischio, per poi proseguire con l'individuazione delle modalità e degli strumenti più adeguati per misurare il livello della performance e dell'impegno scolastico, oltre che dello stile di apprendimento degli studenti. Le fasi successive comprendono: la necessità di integrare i numerosi dati già disponibili all'interno delle istituzioni scolastiche con *surveys* e analisi qualitative che rilevino anche il punto di vista degli studenti; la necessità di accuratezza dei processi di elaborazione ed interpretazione dei dati raccolti (*'the meaning of the collected data'*); la declinazione dei dati raccolti in misure di intervento appropriate ed efficaci rispetto al gruppo target di studenti coinvolto nel processo in uno specifico contesto; valutazione dell'efficacia degli interventi messi in atto.

3. Quale il contributo delle tecniche di Learning Analytics?

Se si prendono in esame i punti chiave della definizione di Learning Analytics, ossia: misurazione, raccolta, analisi e presentazione dei dati

sugli studenti e sui loro contesti, con l'obiettivo di comprendere ed ottimizzare l'apprendimento e l'ambiente in cui si realizza (Ferguson, 2014), appare chiara la stretta connessione con le azioni strategiche indicate dalla letteratura e dalle indagini europee sul tema del successo/insuccesso formativo. Infatti, alcuni dei fattori che guidano lo sviluppo del Learning Analytics (Ivi, pp. 140-141) sono anche alla base delle questioni messe in evidenza nei paragrafi precedenti, in particolare:

- la crescente richiesta da parte delle istituzioni scolastiche e formative di misurare, dimostrare e migliorare i risultati raggiunti, anche al fine di ottimizzare la distribuzione delle risorse e la ricaduta/impatto dei dati raccolti sulla presa di decisioni nella risoluzione di problemi educativi specifici, come quello della dispersione scolastica e dell'insuccesso formativo;
- il riconoscimento di tre principali gruppi di interesse intorno al tema del miglioramento continuo dei processi di insegnamento e apprendimento, ovvero: i governi, le istituzioni educative e gli insegnanti/studenti; con la necessità di adattare e calibrare le tecniche di rilevazione e analisi dei dati, così come la diffusione e declinazione operativa dei risultati ai tre differenti target di riferimento.

Come afferma Galliani (2019), c'è bisogno di utilizzare i dati nelle decisioni educative e l'adozione di dispositivi di Learning Analytics negli ambienti formativi è utile per migliorare le pratiche di insegnamento e apprendimento, in modo contestualizzato alle specifiche situazioni rilevate dagli operatori (docenti, dirigenti, amministratori) durante le loro azioni formative.

In tale ottica, rivestono particolare interesse, per il tema in oggetto, alcune delle principali sfide, intese anche come traiettorie di ricerca da sviluppare, con cui attualmente il settore del Learning Analytics si confronta (Ferguson, 2014, pp. 144-145); nello specifico le più rilevanti:

- supportare la comprensione e l'ottimizzazione dell'apprendimento, infatti capire come si apprende consente di facilitarne le strategie, i modelli ed i processi, non in modo meccanico, bensì contestualizzato e prendendo in considerazione i fattori identitari ed emotivi;
- contribuire a rendere più efficaci e funzionali gli ambienti di apprendimento e la strutturazione delle attività e delle pratiche di insegnamento e apprendimento;
- sviluppare analisi basate sulle prospettive degli studenti, al fine di far emergere le loro esigenze, le loro aspirazioni, motivazioni e percezioni, anche rispetto al concetto di 'successo formativo';



- rimodulare il focus della valutazione, spostandolo da una valutazione sommativa e certificativa degli esiti verso una valutazione formativa dell'intero processo di apprendimento, che aiuti gli studenti a crescere e a riflettere sul proprio percorso formativo.

Nello specifico, i dati sulle esperienze formative degli studenti potrebbero essere utilizzati, in una forma a loro facilmente comprensibile, per fornire feed-back costanti sull'apprendimento e su eventuali difficoltà incontrate, ai fini di una revisione in itinere del percorso didattico, ma anche per sviluppare strategie di scaffolding e di coinvolgimento cognitivo attivo dello studente. Come diversi autori sottolineano (Hattie, 2009; Castoldi, 2011; Murray & Mitchell, 2016), monitorare l'attività dello studente, fornendo feed-back sulle sue prestazioni e stimolandone l'auto-riflessione e l'auto-valutazione, permette di incidere positivamente sull'acquisizione di una maggiore consapevolezza delle proprie risorse e dei propri limiti e, di conseguenza, sul rafforzamento dei livelli di autoefficacia ed autostima.

A tal proposito, Fulantelli e Taibi (2014) sostengono che tanto più i dati relativi ai percorsi di apprendimento degli studenti sono dettagliati e puntuali, tanto più potranno essere utilizzati per personalizzare i percorsi di studio e migliorare i processi di apprendimento.

Khalil e Ebner (2015) evidenziano come tra gli obiettivi del Learning Analytics via sia quello di prevenire l'abbandono scolastico, identificando gli studenti 'a rischio' e fornendo loro informazioni sul proprio sul processo di apprendimento, anche mediante il confronto delle loro prestazioni con quelle degli altri.

In tal senso, anche altre indagini (Vera, 2017) mostrano i vantaggi dei dati messi a disposizione dai dispositivi di Learning Analytics non solo per gli studenti, ma anche per gli insegnanti ed altre figure professionali, in particolare nella precoce individuazione degli alunni che necessitano di supporto; nel monitorare il rendimento degli studenti, ma anche gli stili di apprendimento, le loro attitudini, i loro talenti ed aspirazioni; nel definire attività e strategie didattiche più efficaci; nel migliorare l'individualizzazione dei percorsi e nel diversificare ed articolare le modalità di valutazione, anche nella direzione di incentivare processi di rielaborazione e di documentazione delle attività svolte e risultati conseguiti.

Si sottolinea, dunque, alla luce degli elementi esaminati, come i Learning Analytics possano orientare le fasi del processo di raccolta ed analisi sistematica dei dati proposto dal già citato toolkit europeo per l'identificazione ed il supporto degli studenti a rischio di dispersione scolastica (Jasińska-Maciąg & Tomaszewska-Pękala, 2017) e fornire



delle risposte alle domande guida precedentemente esaminate, nell'ambito di un dialogo costante con la ricerca educativa, anch'essa chiamata sempre di più a rispondere alla necessità di integrare, in una visione complementare, approcci e metodi quantitativi e qualitativi, nella consapevolezza che «i dati non sono dati, ma sono costruiti attraverso processi di operazionalizzazione di fattori, ossia elementi non direttamente osservabili che incidono su un sistema [...] Oltre al problema della costruzione del dato vi è poi il problema della sua interpretazione [...] Sia la costruzione del dato sia la sua interpretazione richiedono quindi una conoscenza di tipo qualitativo [...] Sul versante qualitativo si acquisisce consapevolezza che una scienza che studia 'casi unici' non serve a nulla, dato che non produrrebbe conoscenza esportabile [...] L'attenzione viene portata sulle condizioni che rendono possibile un dato risultato e sulla ricerca di regolarità ed invarianti che caratterizzano casi e contesti e che permettono di trasferire le conclusioni di uno studio svolto in un dato contesto in contesti analoghi, ad esso similari» (Trincherò & Robasto, 2019, p.11). Tali riflessioni metodologiche acquistano grande rilevanza di fronte a ricerche che indagano temi complessi, come quello, per l'appunto, della dispersione scolastica e relativi fattori di rischio; un tema che si configura come indicatore della qualità del sistema scolastico, in termini di equità, di opportunità e successo formativo. Un tema ed un ambito in cui il contributo del Learning Analytics può essere significativo per comprendere “cosa è accaduto”, “perché è accaduto” e “cosa si potrebbe fare per migliorare”, nella prospettiva di tracciare la reale entità del fenomeno e agire su più livelli, sapendo “leggere” le evidenze empiriche, interrogandole e traducendole in strategie di intervento efficaci.

In tale ottica, appare interessante fare riferimento alla revisione sistematica proposta da Papamitsiou e Economides (2014), integrata da meta-analisi che sintetizzano i dati e le evidenze empiriche di alcune ricerche prese in esame, nell'arco temporale compreso tra il 2008 e il 2013, con l'obiettivo di esplorare l'utilizzo e l'impatto del Learning Analytics nell'ambito della progettazione didattica ed educativa in riferimento a determinati ambiti di indagine, quali:

- modelli di comportamento degli studenti;
- indicatori di previsione delle prestazioni scolastiche;
- indicatori predittivi del *drop-out* e delle ripetenze;
- miglioramento della valutazione.

Per ciascun ambito d'indagine gli autori presentano alcuni esempi di utilizzo degli Learning Analytics nel campo della promozione del



successo formativo. Nello specifico, in relazione agli indicatori predittivi dell'abbandono scolastico, vengono sintetizzati e comparati i risultati emersi da studi di caso che hanno utilizzato tecniche di Learning Analytics nell'ambito della scuola secondaria, per indagare, da un lato, la relazione tra la motivazione dello studente e le prestazioni in un compito/esame finale e, dall'altro lato, l'identificazione di fattori che promuovono il successo scolastico, sia a livello di apprendimento formale che informale (Dekker et al., 2009; Romero-Zaldivar et al., 2012). Si riportano alcuni degli aspetti più rilevanti (vedi tabella 1).

Si evince come un efficace utilizzo della grande quantità di dati in possesso delle scuole (ormai disponibile anche in versione informatizzata, considerato il ricorso a strumenti digitali per la registrazione dei voti e di altre informazioni significative) possa essere finalizzato, attraverso l'impiego di tecniche specifiche, al miglioramento del percorso formativo degli studenti soprattutto quando, come nei casi esaminati, si va oltre la mera raccolta dei dati, ma l'interpretazione di questi ultimi diventa oggetto di condivisione tra i vari stakeholder interessati, all'interno di uno specifico contesto di apprendimento. Le potenzialità dei Learning Analytics rappresentano, dunque, nuove strade da percorrere per la ricerca educativa e per le istituzioni scolastiche, in una prospettiva di reciproca collaborazione.



Possibili modalità di utilizzo dei Learning Analytics nella scuola secondaria di secondo grado	Fattori che promuovono il successo formativo
<ul style="list-style-type: none"> – Monitoraggio dell'attività e delle carriere degli studenti mediante la visualizzazione dei file di registro degli studenti (principalmente votazioni ottenute nel primo semestre del primo anno e, di seguito, all'inizio di ciascun semestre) e conseguente estrazione ed aggregazione dei dati, al fine di rilevare gli studenti a rischio già in una fase iniziale; – Discussione dei dati raccolti insieme a studenti e docenti, con un focus su difficoltà incontrate e risorse dell'ambiente di apprendimento; – Messa a punto di strategie didattiche alternative e personalizzate. 	<ul style="list-style-type: none"> – Grado di soddisfazione della propria performance scolastica; – Riduzione del sovraccarico di studio; – Consolidamento delle competenze di base disciplinari e di quelle trasversali; – Autoefficacia percepita; – Didattica motivante; – Buon clima di classe; – Supporto didattico individualizzato e di gruppo; – Valutazione formativa.

Tab. 1: Classification of the results of LA case studies (pedagogical) in order to Prediction of dropout

(Papamitsiou, Economides, 2014; traduzione e sistemazione grafica a cura dell'autrice)

Riferimenti bibliografici

- Alistair R., Leathwood C. (2013). Problematising Early School Leaving. *European Journal of Education*, 3, 405-418.
- Araújo H.C., Magalhães A., Rocha C., Macedo E. (2018). *Policies on Early School Leaving in nine European countries: a comparative analysis*. University of Antwerp Belgium.
- Batini F., Bartolucci M. (eds.). (2016). *Dispersione scolastica. Ascoltare i protagonisti per comprenderla e prevenirla*. Milano: Franco Angeli.
- Batini F. (2014). *Drop-out*. Lavis: Fuorionda.
- Bølling M., Otte C. R., Elsborg P., Nielsen G., Bentsen P. (2018). The association between education outside the classroom and students' school motivation: Results from a one-school-year quasi-experiment. *International Journal of Educational Research*, v89, 22-35.
- Castoldi M. (2011). *Progettare per competenze. Percorsi e strumenti*. Roma: Carocci.
- Colombo M. (2010). *Dispersione scolastica e politiche per il successo formativo. Dalla ricerca sugli early school leaver alle proposte di innovazione*. Trento: Erickson.
- Colombo M. (2015). Abbandono scolastico in Italia: un problema serio, molti circoli viziosi e qualche strategia di prevenzione. *Scuola Democratica*, 2, 411-424.
- Commissione europea (2013). *Reducing early school leaving: key messages and policy support*.
- Dekker G.W., Pechenizkiy M., Vleeshouwers J.M. (2009). Predicting students drop out: a case study. In T. Barnes, M. Desmarais, C. Romero, S. Ventura (eds.), *Proceedings of the 2nd International Conference on Educational Data Mining* (pp. 41-50). EDM 2009, July 1-3, 2009. Cordoba, Spain.
- Ferguson R. (2014). Learning analytics: fattori trainanti, sviluppi e sfide. *TD Tecnologie Didattiche*, 22(3), 138-147.
- Fraccaroli F., Vergani A. (2004). *Valutare gli interventi formativi*. Roma: Carocci.
- Fulantelli G., Taibi D. (2014). Learning Analytics: opportunità per la scuola. *TD Tecnologie Didattiche*, 22(3), 157-164.
- Galliani L. (2019). Tecnologie e valutazione: bio-bibliografia di un intreccio. *Giornale Italiano della Ricerca Educativa*, XII.
- Hattie J. (2009). *Visible learning: a synthesis of over 800 meta-analyses relating to achievement*. London-New York: Routledge.
- Jasińska-Maciązek A., Tomaszewska-Pękała H. (2017). *Reducing early school leaving: toolkit for schools. How to identify and monitor students and schools in need of additional care and support*. Warsaw: Faculty of Education, University of Warsaw.
- Khalil M., Ebner M. (2015). Learning Analytics: Principles and Constraints. *Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications*, 1326-1336.
- Lisimberti C., Montalbetti K. (2015). Oltre il successo scolastico, verso il successo formativo. Presentazione di una ricerca empirica. *Giornale Italiano della Ricerca Educativa*, 15, 79-95.





- Lotti P., Pedani V. (2016). *Rapporto di monitoraggio e analisi dei prototipi di intervento territoriale*. Indire.
- Martini A., Trivellato U. (2011). *Sono soldi ben spesi? Perché e come valutare l'efficacia delle politiche pubbliche*. Venezia: Marsilio.
- Miur (2018). *L'autonomia scolastica per il successo formativo*. Reperibile al seguente link: <http://istruzioneer.gov.it/wp-content/uploads/2018/09/lautonomia-scolastica-per-il-successo-formativo.pdf>
- Murray S., Mitchell J. (2016). Teaching practices that re-engage early school leavers in further education: an Australian study. *Journal of Further and Higher Education*, 3, 372-391.
- Pandolfi L. (2016). Programmi e azioni di contrasto della dispersione scolastica: casi e metodi di valutazione. *Form@re*, 16, 3, 67-78.
- Papamitsiou Z., Economides A. (2014). Learning Analytics and Educational Data Mining in Practice: A Systematic Literature Review of Empirical Evidence. *Educational Technology & Society*, 17 (4), 49-64.
- Raccomandazione del Consiglio d'Europa del 22 maggio 2018 relativa alle competenze chiave per l'apprendimento permanente.
- Raccomandazione 2011/C 191/01 del Consiglio dell'Unione Europea, 28 giugno 2011. *On policies to reduce early school leaving*.
- Romero-Zaldivar V-A., Pardo A., Burgos D., Kloos C.D. (2012). Monitoring student progress using virtual appliances: a case study. *Computers & Education*, 58(4), 1058-1067.
- Scales H.H. (2015). Another look at the drop out problem. *The Journal of Educational Research*, 62, 339-343.
- Trincherò R., Robasto D. (2019). *I mixed methods nella ricerca educativa*. Milano: Mondadori.
- Tuttoscuola (2018). *La scuola colabrodo*. Roma: Tuttoscuola Srl.
- Vera V.D.G. (2017). Learning Analytics and Scholar Dropout: A Predictive Model. *Middle-East Journal of Scientific Research* 25 (7), 1414-1419.

