



La validazione delle registrazioni in IRIS

QUALITÀ DEI DATI NELL'INSTITUTIONAL REPOSITORY



Premessa (1)

Crescita di richiesta di *accountability* rende fondamentale poter disporre di dati certificati, robusti, affidabili, comparabili

Solo a partire da dati aventi queste caratteristiche possono essere costruiti indicatori affidabili

Il CRIS, rispetto alle basi di dati bibliometriche offre la possibilità di raccogliere informazioni e costruire indicatori **su tutte le tipologie di lavori di ricerca** (anche monografie, capitoli ecc.) e quindi di analizzare e monitorare tutte le aree di ricerca

I dati del CRIS sono autodichiarati o importati dalle banche dati (In entrambi i casi non c'è certezza sulla loro correttezza)

Per questo motivo la validazione dei dati su cui costruiamo analisi e monitoraggio non può essere lasciata al caso, né a personale che non ha competenza e sensibilità rispetto ai metadati bibliografici e al loro utilizzo per le diverse campagne di valutazione



Premessa (2)

Un ricercatore non è un esperto di metadati

Spesso l'inserimento di un punto o l'omissione di uno spazio vengono ritenuti ininfluenti, ma possono fare una grande differenza rendendo gli indicatori non utilizzabili o falsando le informazioni.

Se il CRIS diventa lo strumento per fare analisi, per monitorare l'attività di ricerca e per la costruzione di indicatori è necessario garantire (nel limite del possibile) l'affidabilità dei dati e il loro trattamento in maniera uniforme e coerente

Se centralmente (ANVUR o MIUR) i dati provenienti dai CRIS vengono utilizzati per la costruzione di indicatori a livello nazionale è necessario che i fornitori dei dati (atenei ed enti di ricerca) abbiano standard simili di certificazione dei dati forniti

Che un CRIS fornisca al sistema dati certificati è inutile a livello di sistema se ciò non viene garantito da tutti i CRIS

Alla ricerca di uno standard

A formulated rule for common and **voluntary** use, decided by one or several people and organizations (D. Seidl, STI 2015)

Chi propone lo standard da questa proposta ricava legittimazione che è accresciuta in presenza di expertise.

Ciò che i focus groups stanno facendo per lo sviluppo del CRIS IRIS è la ricerca di uno standard.

Ciò avviene quando mancano le leggi e l'interesse politico verso un certo tema.

Lo standard permette inoltre una maggiore flessibilità e possibilità di modifica rispetto alla legge e dunque la sua adozione rappresenta un vantaggio se applicato alle tematiche della ricerca per loro natura in continuo mutamento.



E' importante che a livello nazionale:

Ci siano criteri condivisi rispetto a

- Qualità
- Ricchezza
- Completezza
- Validazione
- Quantità
- Trattamento dei dati esposti

OGGETTO dell'incontro

- ✓ Suggerimenti per la creazione di un gruppo di lavoro a supporto della assicurazione della qualità dei dati
- ✓ Suggerimenti per la predisposizione di strumenti a supporto del gruppo di lavoro
- ✓ Suggerimenti per l'attivazione di un workflow di validazione
- ✓ Indicazione di alcuni problemi aperti su cui è necessario che gli atenei concentrino i loro sforzi al fine di trovare soluzioni condivise (o standards)

Tutto ciò ha senso....

.... Se a livello di amministrazione e di governance il CRIS e l'Archivio istituzionale diventano LO strumento che rappresenta la produzione scientifica dell'Ateneo, per gli organi interni e per gli organi di valutazione esterna.

Se non si crea questa consapevolezza sarà difficile convincere Direttori generali e Rettori dell'importanza di investire nella qualità dei dati

E' necessario un investimento in termini di personale e di formazione

Autorevolezza del gruppo di lavoro

Viene garantita da decisioni che devono essere prese dagli organi dell'ateneo.

OR, Prorettore ricerca, Senato accademico

Tali decisioni – declaratoria tipologie, definizione di rilevanza internazionale, definizione di authorship, definizione di cosa è peer reviewed – devono essere trasparenti e accessibili a tutti.

Una volta che queste decisioni sono prese il gruppo di lavoro dispone degli strumenti necessari per agire in nome dell'Ateneo.

Formazione del gruppo di lavoro: Analisi preliminare

Analisi di quante registrazioni vengono inserite ogni anno nella anagrafe

Definizione di un tempo medio di validazione per ogni registrazione (comprensivo di un tempo standard per la validazione dei PDF)

*Nella definizione del tempo medio occorre tenere conto delle importazioni dalle banche dati (di solito le registrazioni importate sono mediamente più corrette)

Sulla base delle informazioni acquisite definizione del numero di persone necessario per il supporto dell'archivio (validazione+helpdesk via mail, telefonico, in presenza)

Il gruppo di lavoro dovrebbe essere stabile. Se distribuito in strutture diverse deve comunque esserci un responsabile che funga da punto di riferimento

Ragionamenti fatti per UNIMI

Analisi di quante registrazioni vengono inserite ogni anno nella anagrafe

Circa 9000 registrazioni di cui 3/5 importabili dalle banche dati

Definizione di un tempo medio di validazione per ogni registrazione (comprensivo di un tempo standard per la validazione dei PDF)

Circa 20 minuti

*Nella definizione del tempo medio occorre tenere conto delle importazioni dalle banche dati (di solito le registrazioni importate sono mediamente più corrette)

Sulla base delle informazioni acquisite definizione del numero di persone necessario per il supporto dell'archivio (validazione+helpdesk via mail, telefonico, in presenza)

Per UNIMI 3+1 tenendo conto di 5 ore al giorno dedicate esclusivamente alla validazione a cui si aggiungono i rapporti con i docenti l'helpdesk via mail e telefonico. Una persona (a rotazione) è dedicata invece al controllo dei passaggi al sito docente, alla deduplicazione e a specifiche campagne di validazione

Carico di lavoro in fase di avvio e a regime

Molti atenei partono da una situazione in cui la maggior parte delle registrazioni non sono state validate da personale esperto di metadati bibliografici



Nella fase di avvio e fino al raggiungimento della situazione «a regime» sarà necessario prevedere un rafforzamento dell'ufficio che si occupa delle validazioni

Analisi preliminare - strumenti



Analisi preliminare - strumenti

Creazione di una casella di helpdesk e definizione dei tempi di risposta (è importante che le risposte arrivino entro la giornata, meglio se entro qualche ora)

Tutto ciò che può aiutare ad una migliore qualità dei dati inseriti da parte di docenti e ricercatori è buona cosa e rende meno gravoso il lavoro di revisione, strumenti utili sono:

- brevi tutorials sui diversi argomenti: recupero dati dalle banche dati, caricamento del PDF, applicazione della policy, <http://portalevideo.unimi.it/media?mid=464> (qui un esempio di video), linee guida per l'utilizzo del desktop
- corsi a richiesta (ad esempio nei dipartimenti, o per particolari categorie di personale: dottorandi ecc.) – e' una buona cosa pubblicizzare/offrire corsi ogni mese a date fisse e poi tenerli se ci sono iscritti
- brevi interventi nei consigli di dipartimento per rispondere alle eventuali domande degli autori o per illustrare tematiche complesse

Import sì o import no?



In genere le banche dati forniscono dati un po' più affidabili e completi di quelli inseriti dagli autori, tuttavia anche nelle banche dati ci sono errori e il fatto che una registrazione sia importata non garantisce sulla sua qualità e richiede in ogni caso un controllo

Analisi preliminare – Competenze necessarie



Competenze necessarie

- competenze biblioteconomiche con particolare riferimento ai metadati e al Dublin Core
- conoscenza delle diverse banche dati bibliografiche e bibliometriche,
- conoscenza delle tematiche di base per la valutazione della ricerca e dei diversi usi fatti dei dati dell'anagrafe
- conoscenza delle principali questioni legate all'accesso aperto, alla questione degli embarghi e del trattamento delle diverse versioni di un lavoro (pre-print, post-print, versione editoriale)
- conoscenza della policy di ateneo
- conoscenza dell'albero delle tipologie previste dalla anagrafe e del funzionamento del sito loginmiur
- Conoscenza di cosa sia un CRIS e dei flussi di informazioni al suo interno, conoscenza del modello entità relazione

Analisi preliminare – formazione del personale

E' difficile trovare nei nostri atenei personale con competenze così ampie ma anche così specifiche, per cui:

Una volta selezionato il personale andranno previsti dei momenti di formazione sulle diverse tematiche:

- Struttura e funzionamento del CRIS e dell'archivio con particolare riferimento ai metadati descrittivi
- Qualità dei dati
- Uso dei dati ai fini della valutazione, conoscenza delle principali procedure locali e nazionali
- Cenni sull'OA
- Diritti d'autore in relazione alle pubblicazioni scientifiche

Si tratta di formazione continua, perché le caratteristiche delle pubblicazioni si modificano, così come si modificano le diverse richieste a livello nazionale e locale, quelle degli enti finanziatori della ricerca e degli editori.

Caratteristiche del personale che lavora alla validazione



Flessibilità. Intesa come capacità di cogliere e accogliere i mutamenti all'interno dell'editoria scientifica

Analisi preliminare – collezioni specifiche

Nel caso ci siano collezioni dell'archivio che richiedono competenze specialistiche nella validazione (ad esempio i brevetti) è possibile pensare che queste collezioni abbiano validatori diversi e quindi un WF separato

Analisi preliminare – organizzazione del gruppo di lavoro

Il GDL lavora in costante collaborazione sia nell'affrontare le diverse problematiche sia per quanto riguarda l'aggiornamento delle competenze e delle conoscenze (di ciò si deve tener conto quando si effettua la scelta fra validazione centralizzata o decentrata)

Assegnazione delle responsabilità (??)

E' necessario decidere se ci debba essere un responsabile della qualità dei dati

E' necessario decidere se tutti i validatori eseguiranno tutte le operazioni:

Validazione, cancellazione, modifica, deduplicazione, controllo fallimenti invio al MIUR

Oppure se le diverse attività saranno attribuite a persone diverse.

La scelta consigliata è quella che tutti devono essere in grado di fare tutto. Una persona che sia poi responsabile del funzionamento generale e dei controlli a campione è necessaria (prove sulle diverse installazioni)

Predisposizione di linee guida per la validazione

Può essere utile predisporre un manuale di validazione che serva ai validatori come linea guida e strumento di consultazione. Potrebbe essere anche un manuale in forma di FAQ

Il manuale può essere aggiornato a mano a mano che si presentano nuove tematiche (ad esempio l'inserimento di un nuovo campo a seguito di richieste dell'ateneo o a livello nazionale) indicando le relative modalità di controllo

La base delle linee guida sarà la declaratoria delle tipologie approvata dall'Ateneo

Rapporti con gli autori e gestione dei conflitti

Il gdl deve diventare l'authority di ateneo per quanto riguarda l'archivio.

Le regole devono essere fissate da un responsabile scientifico (Prorettore, osservatorio, delegato ecc.) ma l'applicazione pratica è in mano agli operatori.

In particolare ci sono alcune tematiche che vanno risolte prima dell'avvio della validazione:

- l'archivio accoglie preprint? (cosa succede alle registrazioni quando il preprint diventa articolo? Vien fatto un qualche legame?)
- l'archivio accoglie gli in stampa? (la registrazione viene aggiornata? Da chi? Si prevedono controlli periodici su questo materiale?)
- l'archivio accoglie ahead of print? (**PROBLEMI**)

E' importante prevedere un wf interno al gdl che permetta l'aggiornamento dei dati (alla fine di ogni anno?)

- declaratoria delle tipologie presenti nell'archivio: cosa intendo per review, o chi considero autore (la differenza fra gruppo autore unico e autore che fa parte di un gruppo che è coautore).

Dati obbligatori e dati facoltativi

Cosa controllare?

I dati obbligatori non sono molti, quelli facoltativi sono importanti per analisi a granularità fine.

Da decidere cosa fa il validatore: integra il dato nei limiti del possibile / rimanda sempre la registrazione all'autore? Può effettuare delle modifiche senza avvisare l'autore?

Helpdesk



La casella di helpdesk deve comprendere più di una persona in modo che ci sia sempre qualcuno che possa rispondere.

Devono essere fissati e dichiarati i tempi di risposta (in generale la risposta deve essere data entro la giornata)

Una risposta va sempre data anche se è di tipo interlocutorio e non risolutivo

Le persone che risponderanno alle domande lo faranno utilizzando uno stile uniforme, cercando di dare risposte coerenti (da qui ad esempio l'importanza di lavorare insieme. Dare risposte non coerenti rende difficile accreditarsi presso i ricercatori)

Tempi di validazione

Ci sarà una fase di recupero del pregresso in cui i tempi potrebbero essere più lunghi

A regime andrà dichiarato qual è il tempo previsto per la validazione



3 giorni è una buona misura

L'importanza degli identificativi

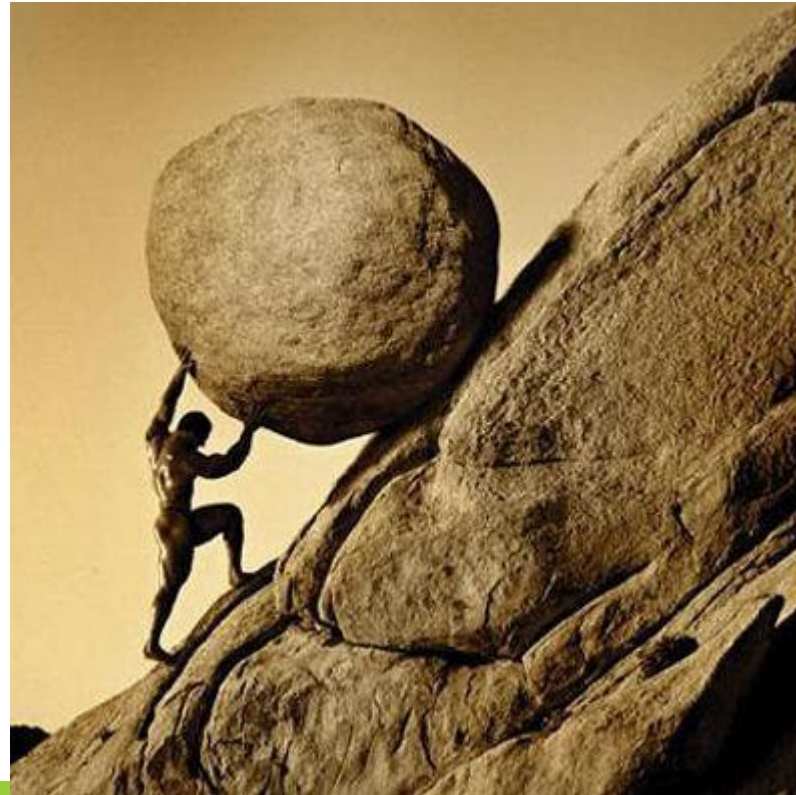


ISBN 10: 1-932698-18-3
ISBN 13: 978-1-932698-18-3



Criticità

Le modifiche al sistema interno o definite dall'agenzia nazionale possono avere un forte impatto rispetto all'archivio.



Modificare la struttura dei dati a volte è necessario e richiesto anche da mutamenti fisiologici delle pubblicazioni scientifiche, o dalla nascita di nuove tipologie, ma le modifiche hanno una risonanza su tutto l'archivio per cui è bene pensare a cosa succede al progresso

Criticità

Chi ha l'ultima parola rispetto all'inserimento in una tipologia o in un'altra? rispetto al carattere internazionale o nazionale di una pubblicazione?

L'ufficio che si occupa della validazione dei dati deve difendere la propria professionalità ed essendo in ultima istanza responsabile della qualità dei dati deve poter prendere la decisione finale ottenendo il consenso dell'autore.



Problemi



Authorship



Chi è l'autore di una pubblicazione?

THE AUTHOR LIST: GIVING CREDIT WHERE CREDIT IS DUE

The first author
Senior grad student on the project. Made the figures.

The third author
First year student who actually did the experiments, performed the analysis and wrote the whole paper. Thinks being third author is "fair".

The second-to-last author
Ambitious assistant professor or post-doc who instigated the paper.

Michaels, C., Lee, E. F., Sap, P. S., Nichols, S. T., Oliveira, L., Smith, B. S.

The second author
Grad student in the lab that has nothing to do with this project, but was included because he/she hung around the group meetings (usually for the food).

The middle authors
Author names nobody really reads. Reserved for undergrads and technical staff.

The last author
The head honcho. Hasn't even read the paper but, hey, he got the funding, and his famous name will get the paper accepted.

JORGE CHAM © 2005

www.phdcomics.com

Quanti autori può avere una pubblicazione?

The image shows a screenshot of a web page from the journal Nature. The page features a dark red header with the 'nature' logo and the tagline 'International weekly journal of science'. A search bar is located in the top right corner. Below the header is a navigation menu with links for Home, News & Comment, Research, Careers & Jobs, Current Issue, Archive, Audio & Video, and For Authors. A secondary navigation bar shows a breadcrumb trail: News & Comment > News > 2015 > August > Article. The main content area displays the article title 'Physics paper sets record with more than 5,000 authors' and a sub-headline 'Detector teams at the Large Hadron Collider collaborated for a more precise estimate of the size of the Higgs boson.' The author's name, Davide Castelvecchi, and the date, 15 May 2015, are listed below. A 'Rights & Permissions' button is visible in the bottom left. On the right side, there are social media sharing options (E-alert, RSS, Facebook, Twitter) and a featured article titled 'Summer in the city' with a cityscape image and the sub-headline 'How cities can beat the heat'. At the bottom right, there is a partial view of another article titled 'The 2015 Mid-Year'.

nature International weekly journal of science

Search Go [Advanced search](#)

Home | News & Comment | Research | Careers & Jobs | Current Issue | Archive | Audio & Video | For Authors

News & Comment > News > 2015 > August > Article

NATURE | NEWS

Physics paper sets record with more than 5,000 authors

Detector teams at the Large Hadron Collider collaborated for a more precise estimate of the size of the Higgs boson.

Davide Castelvecchi

15 May 2015

[Rights & Permissions](#)

E-alert RSS Facebook Twitter

Summer in the city

How cities can beat the heat

Rising temperatures are threatening urban areas, but efforts to cool them may not work as planned.

The 2015 Mid-Year

Cosa si fa quando il nostro autore fa parte di un gruppo?

Quando ad esempio il gruppo è unico autore: ATLAS group (intorno a 3000 autori)

Quando il gruppo collabora ad una ricerca e l'autore del nostro ateneo è fra i collaboratori

Quando il gruppo è autore ma vengono nominati alcuni autori on behalf of..... e l'autore del nostro ateneo fa parte del gruppo, non degli autori nominati

...

Linee guida

Defining the Role of Authors and Contributors

PAGE CONTENTS

1. Why Authorship Matters
2. Who Is an Author?
3. Non-Author Contributors

1. Why Authorship Matters

Authorship confers credit and has important academic, social, and financial implications. Authorship also implies responsibility and accountability for published work. The following recommendations are intended to ensure that contributors who have made substantive intellectual contributions to a paper are given credit as authors, but also that contributors credited as authors understand their role in taking responsibility and being accountable for what is published.

Because authorship does not communicate what contributions qualified an individual to be an author, some journals now request and publish information about the contributions of each person named as having participated in a



Upcoming Events

2015 WEBINAR 3: TAKING AUTHOR INSTRUCTIONS TO THE NEXT LEVEL »

2016 CSE ANNUAL MEETING »

White Paper on Publication Ethics

CSE's White Paper on Promoting Integrity in Scientific Journal Publications, 2012 Update

(approved by the CSE Board of Directors on March 30, 2012)

Download a PDF of the [entire White Paper](#)

1.0 INTRODUCTION

2.0 ROLES AND RESPONSIBILITIES IN PUBLISHING

2.1 [Editor Roles and Responsibilities](#)

2.1.1 [Editorial Freedom](#)

2.1.2 [Confidentiality](#)

[Its of Interest](#)



ROARS

Return On Academic Research


[CHI SIAMO »](#)
[ARGOMENTI »](#)
[IDEE E PROSPETTIVE »](#)
[APPROFONDIMENTI »](#)
[SEGUICI »](#)
[ROARS IN ENGLISH »](#)
[SOSTIENICI](#)
[APPROFONDIMENTI / BIBLIOMETRIA](#)

Se l'autore è un ospite oppure un fantasma

Posted by Paola Galimberti on 1 settembre 2015 at 02:16

Negli ultimi anni si è assistito al fenomeno dell'aumento esponenziale del numero degli autori nelle pubblicazioni scientifiche. La presenza di grandi collaborazioni internazionali è senza dubbio un fattore positivo (e talvolta necessario) per lo sviluppo delle ricerche in alcune aree scientifiche (ad esempio la fisica), ma ciò ha da un lato sollevato il problema della authorship, e di cosa significhi essere un autore all'epoca delle grandi collaborazioni, dall'altro ha anche posto la questione di quale possa essere il reale contributo portato da un autore di 40 lavori in un anno. L'attribuzione di un lavoro ad un autore segue spesso percorsi diversi da quelli del credito e, oltre



Related Posts

Credit where credit is due





[MEMBERS](#)

PROJECT CREDIT

Project CRediT

[Home](#) > [Standards](#) > [Subject Groups](#) > Project CRediT

The Subject Matter Experts Working Group (SME-WG) on [Project CRediT](#) have built on the work of the Wellcome-Harvard contributorship project by extending community participation in the initiative, to include a wider range of publishers, researchers, funding agencies, and academic administrators.

[More Information](#)

Subscribe

[SUBMIT](#)

CASRAI Events

[CASRAI Reconnect15](#)

October 27, 2015 - October 28, 2015 - Toronto,

La data di pubblicazione

Il problema della data di pubblicazione è ora molto urgente per via delle ultime indicazioni di Anvur rispetto al tema (VQR 2011-2014 – Bando definitivo)

E' un problema per gli ahead of prints ma anche per quelle pubblicazioni che escono con forte ritardo per cui la stampa è parecchi anni successiva alla data sul fascicolo.

Lo si può risolvere indicando due date, di cui una obbligatoria, quella di inserimento in fascicolo. L'altra, quella dell'ahead of print resta sia per evitare che la stessa pubblicazione venga utilizzata in campagne di valutazione diverse, ma anche per evitare problemi come quelli creati dalle ultime indicazioni di Anvur.

Dal punto di vista di chi valuta questo risulta essere un problema perché i due campi vanno riempiti correttamente e occorre porre molta attenzione durante il controllo

Chi inserisce per conto di altri

Chi inserisce per conto del proprio docente deve conoscere il sistema ed essere in grado di interpretare i campi da inserire. Queste persone, se non sono autorizzate a inserire nell'archivio devono essere abilitate dopo breve formazione. Questo per facilitare gli scambi via mail nel caso di respingimenti

Registrazioni riaperte

Spesso restano lì sospese. E' necessario periodicamente fare un controllo e sollecitare le chiusure

Rilevanza

Il problema della rilevanza va affrontato a livello di ateneo e deve essere fornita al gruppo di lavoro e a chi lo coordina una definizione che possa servire da riferimento in caso di contestazioni.