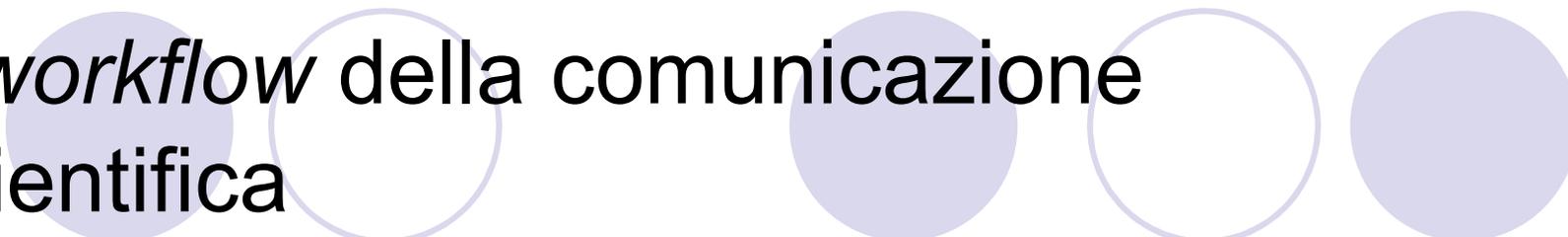# Le nuove frontiere della ricerca scientifica: le nuove forme di documento: dal *data/text mining* al web semantico

Maria Cassella

Università di Torino

maria.cassella@unito.it

Il *workflow* della comunicazione scientifica

- Gather = raccolta

- Create = (ri-uso) e creazione

- Share = condivisione

  Lorcan Dempsey - Reconfiguring the Library Systems Environment - Portal: libraries and the academy , 2008

In altre parole … il profilo del ricercatore nella **Scholarship 2.0**

- Ha molte risorse a disposizione, ma l'attenzione è scarsa (Lorcan Dempsey)
- Ha un approccio sempre più collaborativo alla ricerca … dai medici agli umanisti !
  - Affermazione di discipline transnazionali e "cross-fertilizzate"
- Ha bisogno di integrare e ri-utilizzare automaticamente una grande mole di dati e di informazioni rilevanti (*data/text mining*)
  - Data repositories http://oad.simmons.edu/oadwiki/Data_repositories
  - Data journal : The Journal of Meteorological Data
  - Nuovo ruolo per i bibliotecari: Data Curator
- Rielabora e condivide *preprint*, dati, video, mappe, immagini, *podcast*, commenti, bibliografie, valutazioni sulle piattaforme di *social*, *academic* and *professional networking*

nature.com   Login »

nature network

HOME   FORUMS   GROUPS   PEOPLE   BLOGS   HUBS

Nature Network: connecting scientists

Connect

Keep in touch with colleagues. Make new contacts.

Discuss

Talk about research and scientific issues.

Explore

Read the blogs.
Browse the forums and profiles.

Search

[   ] Search

Get Started

Sign up

What's new on Nature Network

Scientia pro Publica

Most commented blog posts

Most active forums

Recently active people

Popular tags

→ science
→ bioinformatics
→ biology

FACULTY OF 1000
BIOLOGY
MAJOR ADVANCES. EXPERT OPINION.

BIOLOGY REPORTS
NEW from Faculty of 1000
Fast reports on hot topics

email address   ••••••••
☐ Remember me   Forgotten your password?

Search all evaluated articles

ome   **Browse**   **Advanced Search**   **Subscribe**   **Free Trial**   **Hidden Jewels**   **About**   **Faculty Members**   **Associate FMs**   **F1**

The expert guide to the most important advances in biology

Explore the site to see which papers our Faculty of over 2000 world leading scientists have selected.

Sign up for a free 3-week trial to read their expert insights.
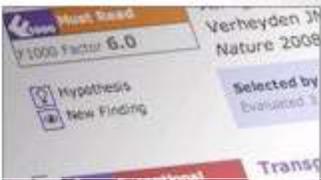
See how it all works on the
F1000 Biology tour

Free Search/Browse
See also our sister service comprising over 2400 Medicine experts

Global faculty of experts ›
Top scientists identify and explain the most important articles

Saving you time ›
Stay on top of the key papers in your chosen subject areas

Customized service ›
Adapt the site to your key areas of interest, store search results and receive regular email alerts

Unique ranking system ›
Faculty Members provide independent assessment of individual papers

Join Now | Sign In

sermo

Home   **Features**   **Tour**   **Blog**   **News**   **About**   **Become a Client**

Have questions?

Get answers from over 112,000 colleagues
for *the ultimate curbside*.

Join Now

Free! Physicians Only

Curbside colleagues today for instant feedback.

1. Join Sermo

Create Account

Please complete our simple

First Name

2. Ask Questions

Q

3. Get Answers

Curbsides

What's New

Recent News

athenahealth/Sermo Partner with
Common Goal: Enable MDs to
Leverage Technology to Improve
Patient Care

Sermo Named to Fast Company's List
of Most Innovative Companies
Recognized as 'Political Force'
Behind Healthcare Reform Efforts

Get Started
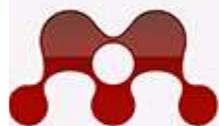
MENDELEY

RESEARCH NETWORKS BETA 0.9

e-mail: _____   password: _____   **Sign in**

☐ Remember me   Forgot your password?

Get Mendeley   How it works   About us   Support   Blog

Organize, share and discover **research papers**

Like iTunes™ for research papers. Mendeley is a free research management tool for desktop & web.

Share papers   cite & create bibliographies

in Word & OpenOffice   collaborate

search & annotate PDFs   sync across PCs

Get Mendeley
FREE DOWNLOAD

Available for Windows, Mac OS X and Linux

Parole chiave della **Scholarship 2.0**

- **Ri-uso**

- **Integrazione**

  - Web semantico, *text mining/data mining,* mappe concettuali, grafici ... (Gargiulo 2009)
  - L'Open Access è un prerequisito

ChemSpider: il *data/text mining* accademico

- http://www.chemspider.com/

- "ChemSpider is a free access service providing a structure centric community for chemists. Providing access to millions of chemical structures and integration to a multitude of other online services, ChemSpider is the richest single source of structure-based chemistry information"

DPedia: il Web semantico applicato a Wikipedia [1]

- **Dbpedia.org** is a community effort to extract structured information from Wikipedia and to make this information available on the Web. Dbpedia allows you to ask sophisticated queries against Wikipedia and to link other datasets on the Webto Wikipedia data.

- The **Dbpedia.org** project uses the Resource Description Framework (RDF) as a flexible data model for representing extracted information and for publishing it on the Web.

DPedia: il Web semantico applicato a Wikipedia [2]

Gli e-journals nella **Scholarship 2.0**

○ Gli e-journals non sono più il principale punto di riferimento della **Scholarship 2.0**

○ Il periodico elettronico è una delle molteplici forme attraverso le quali si "**solidifica**" la comunicazione scientifica:

- Blogs
- Wikis
- Forums
- Podcasts, videocasts
- Professional hubs

Gli e-journals nella **Scholarship 2.0**

○ Continua ad avere una funzione rilevante per la carriera accademica

○ Conserva la funzione di certificazione, ma …

○ l'attenzione degli *stakeholders* della comunicazione scientifica si concentra sull'**articolo**

○ Editori commerciali e OA sperimentano **nuovi servizi a livello dell'articolo**

- Inseriscono commenti, riferimenti, valutazioni
  PLoSONE

- Incorporano i dati primari della ricerca
  - Environmental Research Letters (*Datument*, Murray-Rust, Rzepa, 2004)

- Attivano link a video, *podcast*, mappe, immagini ecc.

Nature

Elsevier article of the future

PLoSONE Article level metrics

Gli **e-journals** nella **Scholarship 2.0**

- La distinzione con altre tipologie di documento sfuma …
  - E-journals/e-books
  - E-journals/repositories [eScholarship: University of California](#)
  - E-journals/blogs [Open Medicine Blog](#)

  - L'informazione si scompone in oggetti digitali modulari

Oltre gli **e-journals:** gli oggetti digitali compositi. Il set di standard **OAI-ORE**

- L'informazione si scompone in oggetti digitali modulari, riutilizzabili, ricomponibili

- OAI- Object Reuse and Exchange è un set di standard, evoluzione di **OAI-PMH**

  "The goal of these standards is to expose the rich content in these aggregations to applications that support authoring, deposit, exchange, visualization, reuse, and preservation" (citazione tratta dal sito OAI-ORE)

- OAI-ORE da' origine ad "un **oggetto digitale composito** formato da un insieme di risorse distribuite nel web indipendentemente dalla loro allocazione (ovvero dai *repositories* che le contengono)"

  (Pasqui, 2009)

Progetti sperimentali su **OAI-ORE**

- [eSciDoc](#) Max Planck Society
  - "Will implement an infrastructure for its researchers which combines different types of information objects in one scientific knowledge space" (Hogenaar, 2009)
- **SCOPE** (Scientific Compound Object Publishing and Editing)
  - Offers authoring tools to researchers to enable them to author, edit and publish scientific compound objects

In **conclusione**
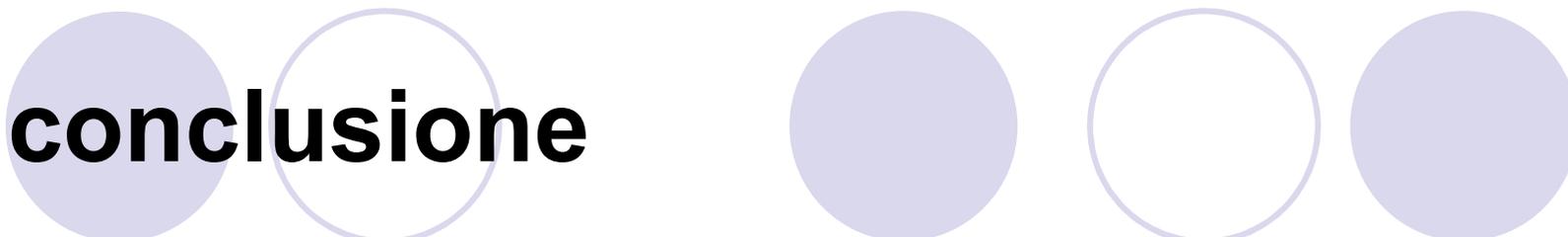
○ I documenti digitali tendono ad abbandonare la prospettiva "cartocentrica" che ha dominato fino ad oggi e si trasformano

○ Gli autori potranno **comporre**, **aggregare**, **riutilizzare** a proprio piacimento oggetti digitali e parti di oggetti digitali

○ Quale ruolo per gli editori?