





## **WISER**

Un motore di ricerca semantico per «expertise» Paolo Ferragina, Università di Pisa







## **OPEN SCIENCE & INNOVATION**

|25 febbraio 2019 - Palazzo Turati - Milano|

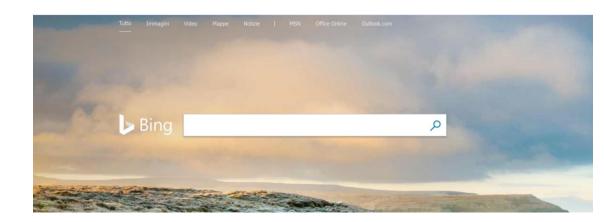






Cerca con Google

Mi sento fortunato



# Il nostro obiettivo

Come possiamo rendere più efficiente ed efficace la ricerca di «esperti accademici» da parte di colleghi, innovatori e aziende ?

	Web	Accademia
archivio	Documenti	Esperti
interrogazione	Parole Chiave	Argomenti, Ricercatori, Dipartimenti



## Welcome to SciVal

SciVal offers quick, easy access to research performance of more than 12,100 research institutions and their associated researchers from 230 nations worldwide.

Don't have access? Request a consultation ¬



Find collaborators to spur innovative solutions to complex problems.

### Demonstrate

Demonstrate my impact for promotion and funding applications.



## Discover

Discover relevant cross-disciplinary areas of research.





### machine learning pisa

Data Mining Logic Program Complex Network Social Network Knowledge Discovery

Q

♣ Follow 2

₽+ Follow | 4

♣ Follow | 4







Dino Pedreschi

h-index: 44 | #Paper: 250 | #Citation: 8649

Professor

♥ KDD Lab, University of Pisa

Similar



### Floriana Esposito

h-index: 34 | #Paper: 474 | #Citation: 4811

Full Professor

University of Bari Dpt. Computer Science

Similar

Machine Learning Knowledge Bases Knowledge Discovery Data Mining

Knowledge Engineering

4042 views



Similar

### Fosca Giannotti 🕝

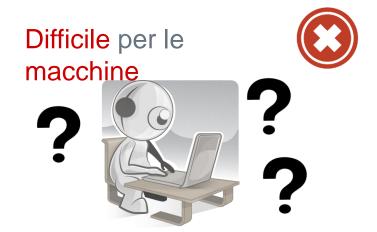
h-index: 42 | #Paper: 252 | #Citation: 9467

senior researcher, Director

Istituto di Scienza e Tecnologie dell'Informazione, National Research Council of Italy (ISTI-CNR)

Data Mining Query Language Data Analysis Data Privacy Movement Data





... ciò diventerà ancora più difficile nei nuovi sistemi di «Conversational AI»







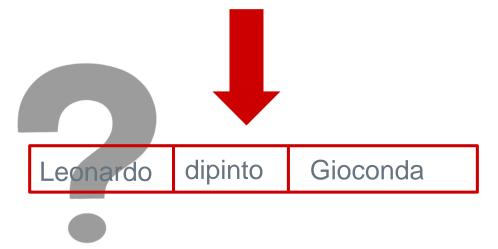
Tradizionalmente le machine hanno rappresentato il testo tramite *l'insieme delle parole* che lo compongono

Leonardo ha dipinto la Gioconda

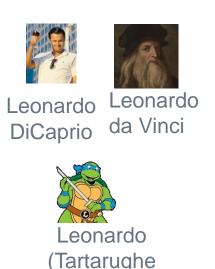


Tradizionalmente le machine hanno rappresentato il testo tramite *l'insieme delle parole* che lo compongono

Leonardo ha dipinto la Gioconda



Tradizionalmente le machine hanno rappresentato il testo tramite *l'insieme delle parole* che lo compongono





## Google Inside Search

Home

Tips & Tricks

Features

Search Stories

Playground

Blog

Help

Search plus your world

Flight Search

Related Search Previews

Search by Image

Voice Search

Google Instant





#### leonardo da vinci



Books

Settings

About 156,000,000 results (0.68 seconds)

### Leonardo da Vinci - Wikipedia

https://en.wikipedia.org/wiki/Leonardo\_da\_Vinci ▼

Leonardo di ser Piero da Vinci more commonly Leonardo da Vinci or simply Leonardo, was an Italian polymath of the Renaissance whose areas of interest ...

Works: Mona Lisa; The Last Supper, Salvator ... Known for: Art, science Movement High Renaissance

List of works by Leonardo da ... - Science and inventions - Portal:Leonardo da Vinci

#### Leonardo da Vinci - Wikipedia

https://it.wikipedia.org/wiki/Leonardo da Vinci \* Translate this page

Leonardo di ser Piero da Vinci (Anchiano, 15 aprile 1452 - Amboise, 2 maggio 1519) è stato un inventore, artista e scienziato italiano. Uomo d'ingegno e ...

Codici di Leonardo da Vinci : Vita personale di Leonardo da ... : Uomo vitruviano

### Top stories



Leonardo da Vinci's Geological Secrets Revealed By Forensic Science

Forbes - 1 day ago



miti da sfatare su Leonardo da Vinci

Wired · 2 days ago



Leonardo da Vinci masterpieces to be displayed in

Derbyshire Live · 2 days ago

→ More for leonardo da vinci

People also ask

What was Leonardo da Vinci's greatest achievement?

What is so great about da Vinci?

https://www.leonardodavinci.net/ -

Who was Leonardo da Vinci apprenticed to?

Why was Leonardo da Vinci famous?

Leonardo da Vinci - Paintings, Drawings, Quotes, Facts, & Biography

Leonardo da Vinci was a true genius who graced this world with his presence from April 15, 1452 to May 2, 1519. He is among the most influential artists in ...



### Leonardo da Vinci



Leonardo di ser Piero da Vinci, more commonly Leonardo da Vinci or simply Leonardo, was an Italian polymath of the Renaissance whose areas of interest included invention, drawing, painting, sculpting, ... Wikipedia

Born: April 15, 1452, Anchiano

Died: May 2, 1519, Château du Clos Lucé, Amboise, France

On view: Louvre Museum, Royal Collection Trust, Uffizi Gallery,

Periods: High Renaissance, Early renaissance, Renaissance, Italian Renaissance, Florentine painting

Known for: Art, science

Buried: Chapel of Saint-Hubert, Amboise, France

Quotes

View 7+ more

Simplicity is the ultimate sophistication.

Art is never finished, only abandoned.

When once you have tasted flight, you will forever walk the earth with your eyes turned skyward, for there you have been, and there you will always long to return.

#### Artworks





People also search for







View 15+ more







Feedback











Google **Faculty Research Awards** 

due edizioni: 2010 & 2013

# Capire il testo mediante le sue «entità»

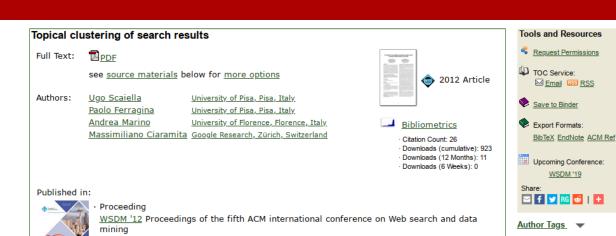




# Documenti Abstract di pubblicazioni







Email BS RSS

WSDM '19



Source Materials

Pages 223-232

ACM New York, NY, USA @2012

Seattle, Washington, USA — February 08 - 12, 2012

References

table of contents ISBN: 978-1-4503-0747-5 doi>10.1145/2124295.2124324

Cited By



Search results clustering (SRC) is a challenging algorithmic problem that requires grouping together the results returned by one or more search engines in topically coherent clusters, and labeling the clusters with meaningful phrases describing the topics of the results included in them. In this paper we propose to solve SRC via an innovative approach that consists of modeling the problem as the labeled clustering of the nodes of a newly introduced graph of topics. The topics are Wikipedia-pages identified by means of recently proposed topic annotators [9, 11, 16, 20] applied to the search

Index Terms

results, and the edges denote the relatedness among these topics computed by taking into account the linkage of the Wikipedia-graph.

Publication

Reviews

Comments

Table of Contents

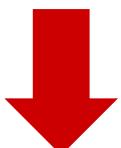
We tackle this problem by designing a novel algorithm that exploits the spectral properties and the labels of that graph of topics. We show the superiority of our approach with respect to academic state-of-the-art work [6] and well-known commercial systems (CLUSTY and LINGO3G) by performing an extensive set of experiments on standard datasets and user studies via Amazon Mechanical Turk. We test several standard measures for evaluating the performance of all systems and show a relative improvement of up to 20%.

Abstract Source Materials Authors References Cited By Index Terms Publication Reviews Comments Table of Contents

Search results clustering (SRC) is a challenging algorithmic problem that requires grouping together the results returned by one or more search engines in topically coherent clusters, and labeling the clusters with meaningful phrases describing the topics of the results included in them.

In this paper we propose to solve SRC via an innovative approach that consists of modeling the **problem** as the *labeled clustering* of the nodes of a newly introduced *graph of topics*. The topics are **Wikipedia**-pages identified by means of recently proposed topic annotators [9, 11, 16, 20] applied to the search results, and the edges denote the relatedness among these topics computed by taking into account the linkage of the **Wikipedia**-graph.

We tackle this problem by designing a novel algorithm that exploits the spectral properties and the labels of that graph of topics. We show the superiority of our approach with respect to academic state-of-the-art work [6] and well-known commercial systems (CLUSTY and LINGO3G) by performing an extensive set of experiments on standard datasets and user studies via Amazon Mechanical Turk. We test several standard measures for evaluating the performance of all systems and show a relative improvement of up to 20%.













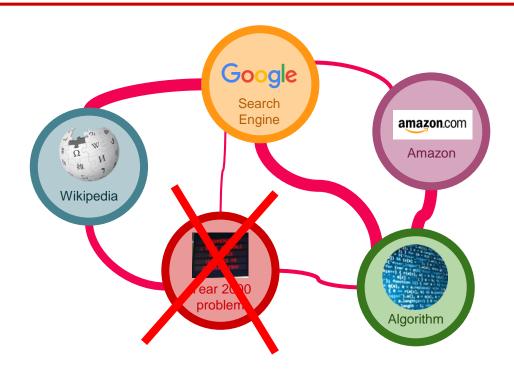


Abstract Source Materials Authors References Cited By Index Terms Publication Reviews Comments Table of Contents

Search results clustering (SRC) is a challenging algorithmic problem that requires grouping together the results returned by one or more search engines in topically coherent clusters, and labeling the clusters with meaningful phrases describing the topics of the results included in them.

In this paper we propose to solve SRC via an innovative approach that consists of modeling the problem as the labeled clustering of the nodes of a newly introduced graph of topics. The topics are Wikipedia-pages identified by means of recently proposed topic annotators [9, 11, 16, 20] applied to the search results, and the edges denote the relatedness among these topics computed by taking into account the linkage of the Wikipedia-graph.

We tackle this problem by designing a novel algorithm that exploits the spectral properties and the labels of that graph of topics. We show the superiority of our approach with respect to academic state-of-the-art work [6] and well-known commercial systems (CLUSTY and LINGO3G) by performing an extensive set of experiments on standard datasets and user studies via Amazon Mechanical Turk. We test several standard measures for evaluating the performance of all systems and show a relative improvement of up to 20%.







Search results dustering (SRC) is a challenging algorithmic problem that requires grouping together the results returned by one or more search engines in topically coherent dusters, and labeling the clusters with meaningful phrases describing the topics of the results included in them.

In this paper we propose to solve SAC via an innovative approach that consists of modeling the problem as the labeled obstening of the nodes of a newly introduced graph of topics. The topics are Wikipedia-pages identified by means of recently proposed topic annotators [9, 11, 16, 20] applied to the search results, and the edges denote the relatedness among these topics computed by taking into account the linkage of the Wikipedia-graph.

We tackle this problem by designing a novel algorithm that exploits the spectral properties and the labels of that graph of topics. We show the superiority of our approach with respect to academic state-of-the-art work [6] and well-known commercial systems (CLUSTY and LINGOSE) by performing an extensive set of experiments on standard datasets and user studies via Amazon Mechanical Turk. We test several standard measures for evaluating the performance of all systems and show a relative improvement of up to 20%.

Introducing a new dataset with human judgments of entity relatedness, we present a thorough study of all entity relatedness measures in recent Iterature based on Wikipedia as the knowledge graph. No dear dominance is seen between measures based on textual similarity and graph proximity. Some of the better measures involve expensive global graph computations. We then propose a new, space-efficient, computationally lightweight, twostage framework for relatedness computation. In the first stage, a small weighted subgraph is dynamically grown around the two query entities; in the second stage, relatedness is derived based on computations on this subgraph. Our system shows better agreement with human judgment than existing proposals both on the new dataset and on an established one. We also plug our relatedness algorithm into a state-of-the-art entity linker and

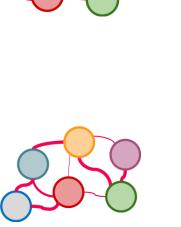
Abstract Authors References Cited By Index Terms Publication Reviews Comments Table of Contents

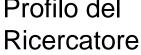
observe an increase in its accuracy and robustness.













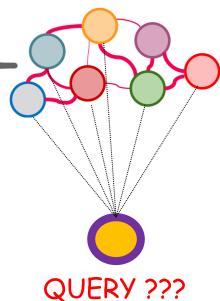




Broth is an open source general-purpose data compressor introduced by Google in late 2013 and now adopted in most known browsers and Web seriers. It is publicly available on Gibblo and its data format was submitted as RFC 7932 in July 2016. Broth is based on the Lempel-Ziv compression scheme and planned as a generic replacement of Guip and ZLib. The main coal in its design was to compress data on the Internet, which meant optimizing the resources used at decoding time, while achieving maximal compression density.

This article is intended to provide the first thorough, systematic description of the Broth format as well as a detailed computational and experimental analysis of the main algorithmic blocks underlying the current encoder implementation, together with a companison against compressors of different families constituting the state-of-the-art either in practice or in theory. This treatment will allow us to raise a set of new algorithmic and software engineering problems that deserve further attention from the scientific community.







Search for expertise... Q

Search by Expertise Search by Name Search by Department

- > ~1.5K autori
- > ~65K documenti
- > ~35K argomenti (topic)
- Adottato dall'Ufficio Trasferimento Tecnologico di UniPI







machine learning Q

Search by Expertise

Search by Name

Search by Department

0 100 results in 1.513s

### Alessandro Lenci

Professor

DIPARTIMENTO DI FILOLOGIA, LETTERATURA E LINGUISTICA

### **Davide Bacciu**

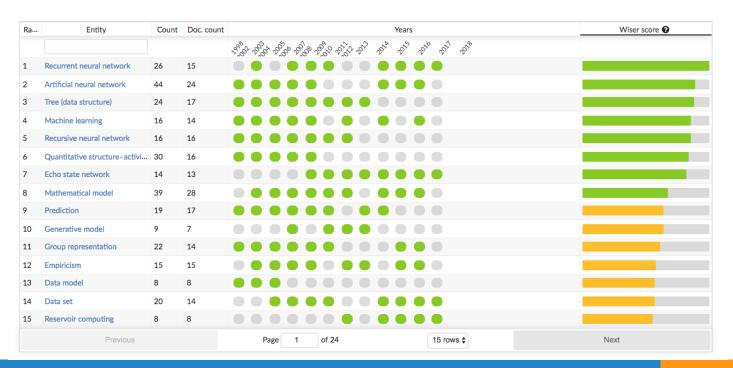
Researcher

DIPARTIMENTO DI INFORMATICA

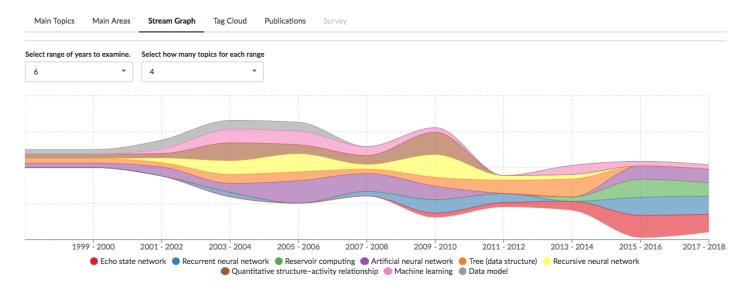


Main Topics Main Areas Stream Graph Tag Cloud Publications Surve











Search for expertise		Q
Search by Expertise	Search by Name	Search by Department



Research Infrastructure http://wiser.d4science.org



http://acube.di.unipi.it