Microsoft Research

Microsoft Academic Graph

Alex Wade
Microsoft Research

Microsoft

@alexwade

SIGIR2015
SANTIAGO - CHILE

Microsoft Academic Graph

- Background
- Bing / Cortana
- The Data
- The Challenge (WSDM Cup)

Microsoft Academic Graph

- Background
  - Project Libra (2005 – 2008) Wei-Ying Ma, et al.
  - Microsoft Academic Search (2008 – 2012)
  - → Bing, Cortana (2015 –)
- Bing examples
- The Data
- The Challenge (WSDM Cup)

	Papers	Citations	Authors
Libra	~5M	~20M	~1M
MAS	~50M	~350M	~20M
Bing	~125M	~1B	~40M

SIGIR2015
SANTIAGO - CHILE

Microsoft Academic Search Beta

Fields of Study

Advanced Search

All Fields of Study
Computer Science
Algorithms & Theory
Artificial Intelligence
Bioinformatics & Computational Biology
Computer Education
Computer Vision
Data Mining
Databases
Human-Computer Interaction
Information Retrieval
Machine Learning & Pattern Recognition
Interaction
Information Retrieval
Machine Learning

Authors Publications Conferences Journals Keywords Organizations

Top auth






Authors Publications Conferences Journals Keywords Organizations

Top pu

Introdu
Indexin
Informa
Term-w
Lettters
Machin
The Ac
The co
Automa
A vecto

See more

Authors Publications Conferences Journals Keywords Organizations

Top conferences in information retrieval

SIGIR - Research and Development in Inform...
CIKM - International Conference on Informa...
TREC - Text REtrieval Conference
JCDL - ACM/IEEE Joint Conference on Digita...
DL - Digital Libraries
ECDL - European Conference on Digital Libr...
CLEF - Cross-Language Evaluation Forum
RIAO - Recherche d'Information Assistee pa...
ECIR - European Colloquium on IR Research
Multimedia Information Retrieval

See more

Authors (3516)

W. Bruce Croft
Jamie Callan
James Allan
Cheng-xiang Zhai
Norbert Fuhr
Mark Sanderson
Susan T. Dumais
Justin Zobel
Alistair Moffat
Ryen W. White

Keywords (2238)

Document Retrieval **Indexation** Information Need **Information Retrieval System** **Information Retrieval** Language Model Machine Learning Query Expansion Question Answering Relevance Feedback **Search Engine** **Test Collection** Text Retrieval Web Pages **Web Search**

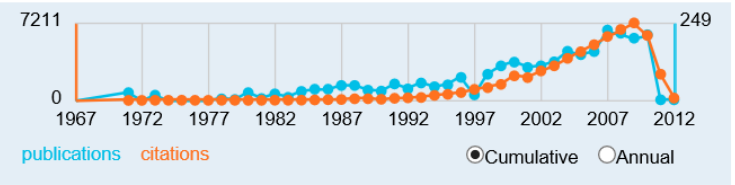
SIGIR - Research and Development in Information Retrieval

Publications: 2,717 | Citation Count: 72,105 (Self-Citation: 5,483)

Year Range: 1971-2012

Fields of study: Information Retrieval

Homepage



Publications (2717)

Sort by: Year

[Improving tweet stream classification by detecting changes in word probability]
Kyosuke Nishida, Takahide Hoshide, Ko Fujimura
Conference: Research and Development in Information Retrieval - SIGIR, pp. 971-980, 2012

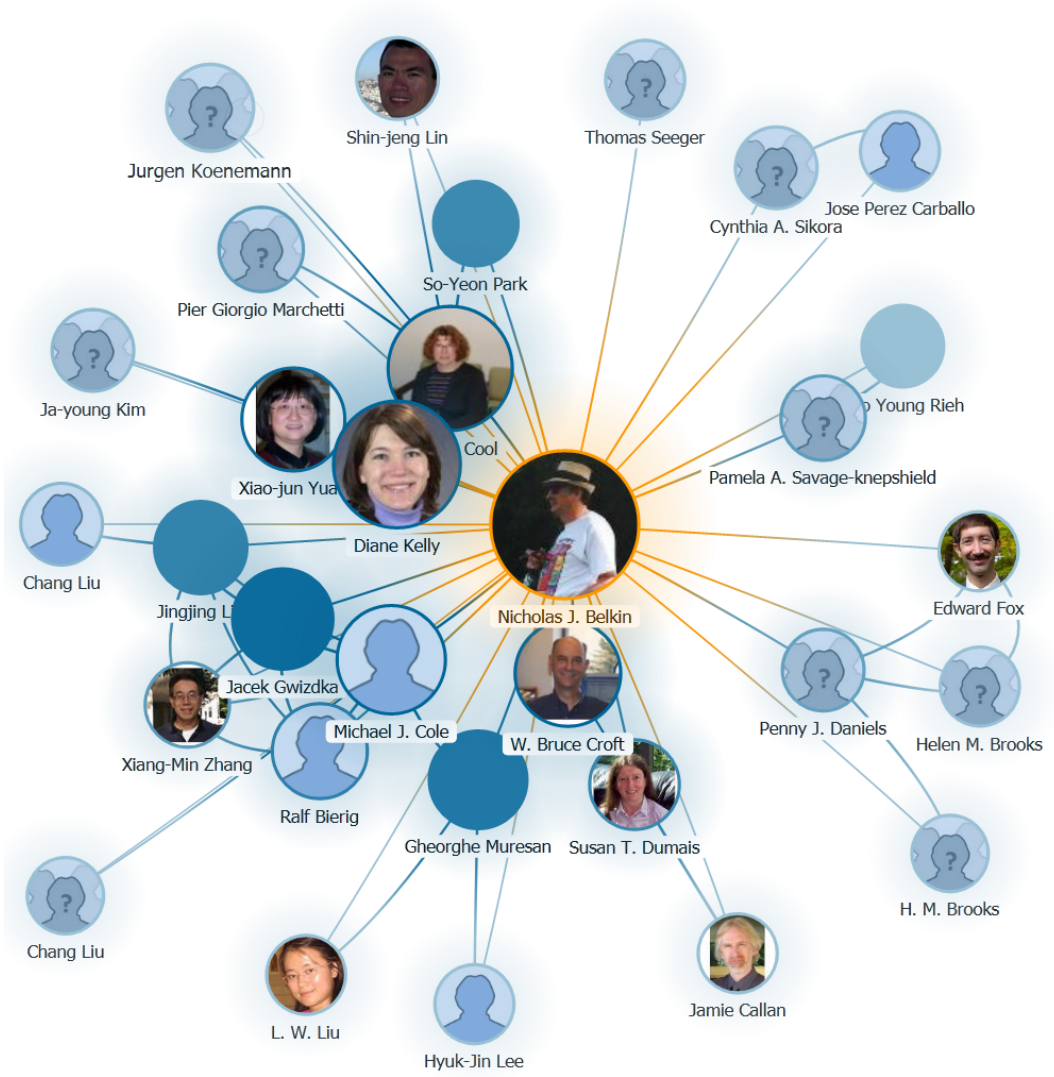
[Combining implicit and explicit topic representations for result diversification]
Jiyin He, Vera Hollink, Arjen de Vries
Conference: Research and Development in Information Retrieval - SIGIR, pp. 851-860, 2012

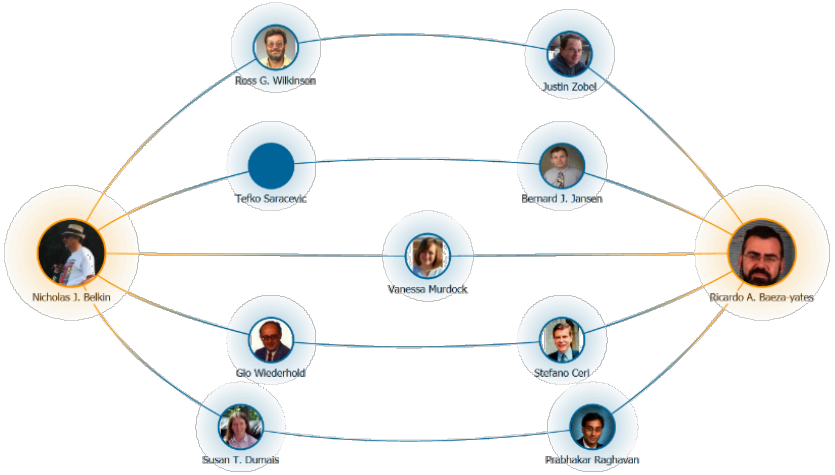
[Personalized social query expansion using social bookmarking systems]
Mohamed Reda Bouadjenek, Hakim Hacid, Mokrane Bouzeghoub, Johann Daigremont
Conference: Research and Development in Information Retrieval - SIGIR, pp. 1113-1114, 2011

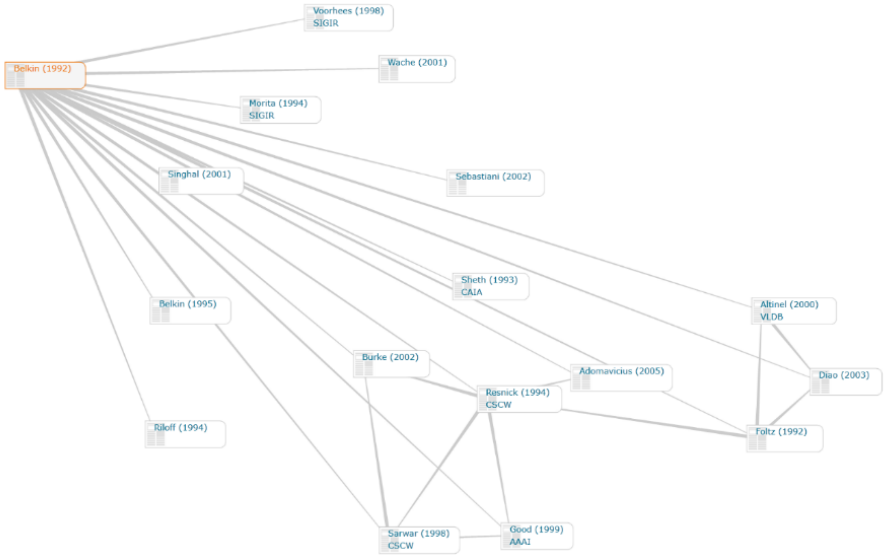
[Evaluating the Synergic Effect of Collaboration in Information Seeking]
Chirag Shah, Roberto Gonzalez-Ibanez
Conference: Research and Development in Information Retrieval - SIGIR, pp. 913-922, 2011

[Jester 2.0 : Evaluation of a New Linear Time Collaborative Filtering Algorithm] (Citations: 10)
Mark Digiovanni, Hiro Narita, Ken Goldberg
Conference: Research and Development in Information Retrieval - SIGIR, 2010

Published in 2012.


Jurgen Koenemann
Shin-jeng Lin
Thomas Seeger
Cynthia A. Sikora
Jose Perez Carballo
So-Yeon Park
Pier Giorgio Marchetti
Ja-young Kim
Xiao-jun Yua
Cool
Diane Kelly
Young Rieh
Pamela A. Savage-knepshield
Chang Liu
Jingjing L
Nicholas J. Belkin
Edward Fox
Jacek Gwizdka
Michael J. Cole
W. Bruce Croft
Penny J. Daniels
Helen M. Brooks
Xiang-Min Zhang
Ralf Bierig
Gheorghe Muresan
Susan T. Dumais
H. M. Brooks
Chang Liu
L. W. Liu
Hyuk-Jin Lee
Jamie Callan


Ross G. Wilkinson
Justin Zobel
Tefko Saracevic
Bernard J. Jansen
Nicholas J. Belkin
Vanessa Murdock
Ricardo A. Baeza-yates
Gio Wiederhold
Stefano Ceri
Susan T. Dumais
Prabhakar Raghavan


Voorhees (1998) SIGIR
Belkin (1992)
Wache (2001)
Morita (1994) SIGIR
Singhal (2001)
Sebastiani (2002)
Sheth (1993) CAIA
Belkin (1995)
Altinel (2000) VLDB
Burke (2002)
Adomavicius (2005)
Diao (2003)
Resnick (1994) CSCW
Riloff (1994)
Foltz (1992)
Sarwar (1998) CSCW
Good (1999) AAAI

Microsoft Academic Graph

- Background
  - Project Libra (2005 – 2008) Wei-Ying Ma, et al.
  - Microsoft Academic Search (2008 – 2012)
  - → Bing, Cortana (2015 –)
- Bing examples
- Get the Data
- WSDM Cup

An Overview of Microsoft Academic Service (MAS) and Applications

Arnab Sinha, Zhihong Shen, Yang Song, Hao Ma, Darrin Eide, Bo-june (Paul) Hsu, Kuansan Wang
Microsoft Research, Redmond, WA 98052, USA
{arsinha, zhihosh, yangsong, haoma, darrine, paulhsu, kuansanw}@microsoft.com

ABSTRACT

In this paper we describe a new release of a Web scale entity graph that serves as the backbone of Microsoft Academic Service (MAS), a major production effort with a broadened scope to the namesake vertical search engine that has been publicly available since 2008 as a research prototype. At the core of MAS is a heterogeneous entity graph comprised of six types of entities that model the scholarly activities: field of study, author, institution, paper, venue, and event. In addition to obtaining these entities from the publisher feeds as in the previous effort, we in this version include data mining results from the Web index and an in-house knowledge base from Bing, a major commercial search engine. As a result of the Bing integration, the new MAS graph sees significant increase in size, with fresh information streaming in automatically following their discoveries by the search engine. In addition, the rich entity relations included in the knowledge base provide additional signals to disambiguate and enrich the entities within and beyond the academic domain. The number of papers indexed by MAS, for instance, has grown from low tens of millions to 83 million while maintaining an above 95% accuracy based on test data sets derived from academic activities at Microsoft Research. Based on the data set, we demonstrate two scenarios in this work: a knowledge driven, highly interactive dialog that seamlessly combines reactive search and proactive suggestion experience, and a proactive heterogeneous entity recommendation.

Keywords

Academic search; Recommender systems; Entity conflation

1. INTRODUCTION

Recent years have witnessed a paradigm shift in how the knowledge on the Web is made available to the users. The trend is highly visible in the evolution of the Web search engine. The traditional Web search outcomes often serve the users' need at best in a "hit-or-miss" fashion [4, 7]. A multi-year initiative in the industry, called Bing Dialog in Microsoft [11] and Knowledge Vault in Google [5], addresses this challenge by using statistical inferences to better organize the Web information and support much richer forms of interaction in recognizing and serving the user needs. In addition to reactively retrieving information and answering questions, the model proactively includes additional dialog acts, such as confirmation, disambiguation, refinement and digression. Coupled with statistical user intent inferences, these acts significantly expedite the process of serving users with the knowledge they need [14]. Our work aims at leveraging this model in addressing the information needs in areas where the sheer amount of information available through a multitude of channels has exceeded the human capacity in processing them. Although most search engines have provided advanced operators for users to compose elaborated queries to better filter out unwanted materials, their arcane syntax has relegated their usages to a negligible rate. A goal of the modern dialog approach to Web search is therefore to utilize advanced techniques to enable the search engines to communicate with users in natural language. Because the dialog inferences inevitably require the system to anticipate or predict the needs of the users, another emerging trend in the search engine evolution is to extend the prediction behaviors into system initiated notifications. The growing prevalence of mobile personal assistants serve as a natural vehicle to deliver proactive notifications, potentially preempting the needs of user initiated search for information [10].

In this paper, we present two applications in the area of academic publications to demonstrate the potentials of the emerging search paradigm. The first application, described in Section 3.1, illustrates a natural language powered interactive search experience. By leveraging the relationships among the entities in the academic domain, the natural language processor is able to harvest the syntactic and semantic cues for parsing and predicting user queries. The second application, described in Section 3.2, demonstrates how a recommendation system can take advantage of the relationships across different types of entities to offer heterogeneous suggestions. Noting that the statistical techniques underlying these two applications are by no means perfect, we further decide to make the data used by the two applications publicly available so that the community can jointly attack the challenging unsolved problems. The data set is an update to the corpus previously released for research purposes [2] and will be described in details in Section 2. The two applications also exemplify a commonly encountered scenario in which the results presented to the users should be properly ranked. The ranking algorithms and the measurements for determining the ranking order remain actively research topics. Given the surge in the count of academic entities and observable limitations of citation count based impact metrics, the problem of defining meaningful impact metrics of academic entities (e.g. papers, authors, conferences) is gaining substantial interest among the researchers [8, 3]. We hope this open corpus can contribute not only to advance information technologies for other innovative applications but also trigger a new horizon of

WWW 2015 Companion, May 18–22, 2015, Florence, Italy.
ACM 978-1-4503-3473-0/15/05.
http://dx.doi.org/10.1145/2740908.2742839.

http://aka.ms/academicgraph

Microsoft Academic Graph

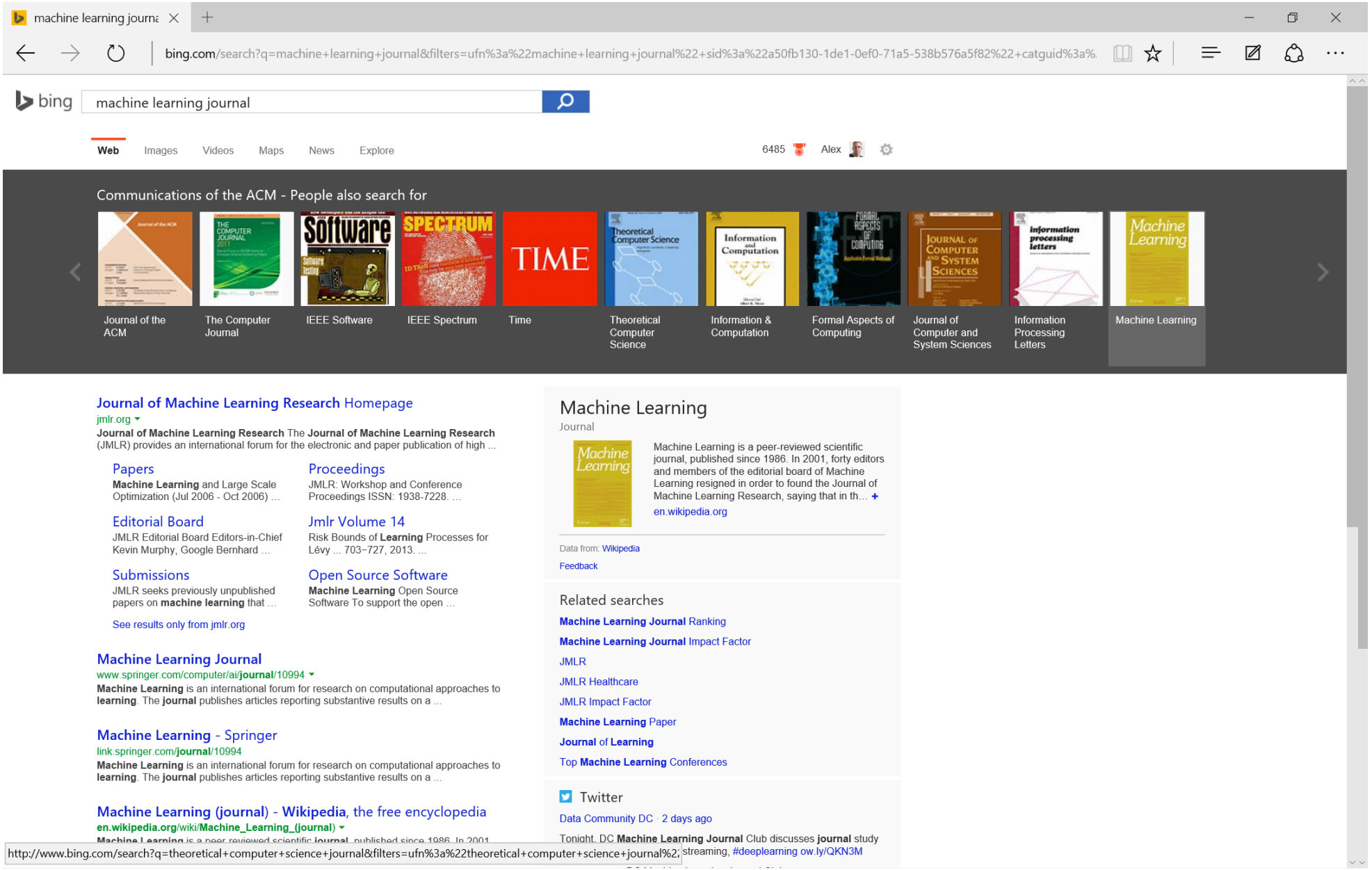
- Background
  - Project Libra (2005 – 2008) Wei-Ying Ma, et al.
  - Microsoft Academic Search (2008 – 2012)
  - → Bing, Cortana (2015 –)

- Bing examples

- The Data

- The Challenge (WSDM Cup)

	Papers	Citations	Authors
Libra	~5M	~20M	~1M
MAS	~50M	~350M	~20M
Bing	~125M	~1B	~40M


machine learning journa
bing.com/search?q=machine+learning+journal&filters=ufn%3a%22machine+learning+journal%22+sid%3a%22a50fb130-1de1-0ef0-71a5-538b576a5f82%22+catguid%3a%
bing
machine learning journal
Web
Images
Videos
Maps
News
Explore
6485
Alex
Communications of the ACM - People also search for
Journal of the ACM
The Computer Journal
IEEE Software
IEEE Spectrum
Time
Theoretical Computer Science
Information & Computation
Formal Aspects of Computing
Journal of Computer and System Sciences
Information Processing Letters
Machine Learning
Journal of Machine Learning Research Homepage
jmlr.org
Journal of Machine Learning Research The Journal of Machine Learning Research (JMLR) provides an international forum for the electronic and paper publication of high ...
Papers
Machine Learning and Large Scale Optimization (Jul 2006 - Oct 2006) ...
Proceedings
JMLR: Workshop and Conference Proceedings ISSN: 1938-7228. ...
Editorial Board
JMLR Editorial Board Editors-in-Chief Kevin Murphy, Google Bernhard ...
Jmlr Volume 14
Risk Bounds of Learning Processes for Lévy ... 703–727, 2013. ...
Submissions
JMLR seeks previously unpublished papers on machine learning that ...
Open Source Software
Machine Learning Open Source Software To support the open ...
See results only from jmlr.org
Machine Learning Journal
www.springer.com/computer/ai/journal/10994
Machine Learning is an international forum for research on computational approaches to learning. The journal publishes articles reporting substantive results on a ...
Machine Learning - Springer
link.springer.com/journal/10994
Machine Learning is an international forum for research on computational approaches to learning. The journal publishes articles reporting substantive results on a ...
Machine Learning (journal) - Wikipedia, the free encyclopedia
en.wikipedia.org/wiki/Machine_Learning_(journal)
Machine Learning
Journal
Machine Learning is a peer-reviewed scientific journal, published since 1986. In 2001, forty editors and members of the editorial board of Machine Learning resigned in order to found the Journal of Machine Learning Research, saying that in th... +
en.wikipedia.org
Data from: Wikipedia
Feedback
Related searches
Machine Learning Journal Ranking
Machine Learning Journal Impact Factor
JMLR
JMLR Healthcare
JMLR Impact Factor
Machine Learning Paper
Journal of Learning
Top Machine Learning Conferences
Twitter
Data Community DC · 2 days ago
Tonight, DC Machine Learning Journal Club discusses journal study
streaming, #deeplearning ow.ly/QKN3M
http://www.bing.com/search?q=theoretical+computer+science+journal&filters=ufn%3a%22theoretical+computer+science+journal%22

michael i. jordan - Bing

bing.com/search?q=michael+i.+jordan&filters=ufn%3a%22michael+i.+jordan%22+sid%3a%2298f2e97b-a8e5-c5af-3fed-7a1951d184ed%22+catguid%3a%228a82e2c7-39c4

bing michael i. jordan

Web Images Videos Maps News Explore 6485 Alex

Andrew Ng - People also search for

Daphne Koller | Michael I. Jordan, Academic Advisor | Geoffrey Hinton | Ray Kurzweil | Sebastian Thrun | Raj Reddy | Stuart J. Russell | Dan Boneh | Jeffrey Ullman | Peter Norvig | Mehran Sahami | Donald Knuth

Connection to Andrew Ng

Michael I. Jordan and Andrew Ng both wrote Latent dirichlet allocation.

Michael I. Jordan's Home Page
www.cs.berkeley.edu/~**jordan**
Michael I. Jordan Pehong Chen Distinguished Professor Department of EECS Department of Statistics University of California, Berkeley. Emails: EECS Address:
Publications · Yuchen Zhang · Martin Wainwright · Andre Wibisono · Ashia Wilson

Michael Jordan - Wikipedia, the free encyclopedia
en.wikipedia.org/wiki/**Michael_Jordan**
Michael Jeffrey **Jordan** (born February 17, 1963), also known by his initials, MJ, is an American former professional basketball player. He is also an entrepreneur, and ...
Early years · High school · College · Professional career · Olympic career

Images of **michael i. jordan**
bing.com/images

See more images of **michael i. jordan**

Michael I. Jordan - Wikipedia, the free encyclopedia
en.wikipedia.org/wiki/**Michael_I._Jordan**
Michael Irwin **Jordan** (born 1956) is an American scientist, Professor at the University of

Michael I. Jordan

Scientist

Michael Irwin Jordan is an American scientist, Professor at the University of California, Berkeley and leading researcher in machine learning and artificial intelligence. Jordan was born in Ponchatoula, Louisiana, to a working-class fa... +

Wikipedia

Born: Feb 25, 1956 (age 59) · Louisiana, United States, with Territories

Education: Arizona State University · University of California, San Diego

Academic advisor: David Rumelhart

Written works: Probabilistic Grammars and Hierarchical Dirichlet Processes · Hierarchical Dirichlet Processes · Latent dirichlet allocation

Timeline

1988: He was a professor at MIT from 1988-1998.

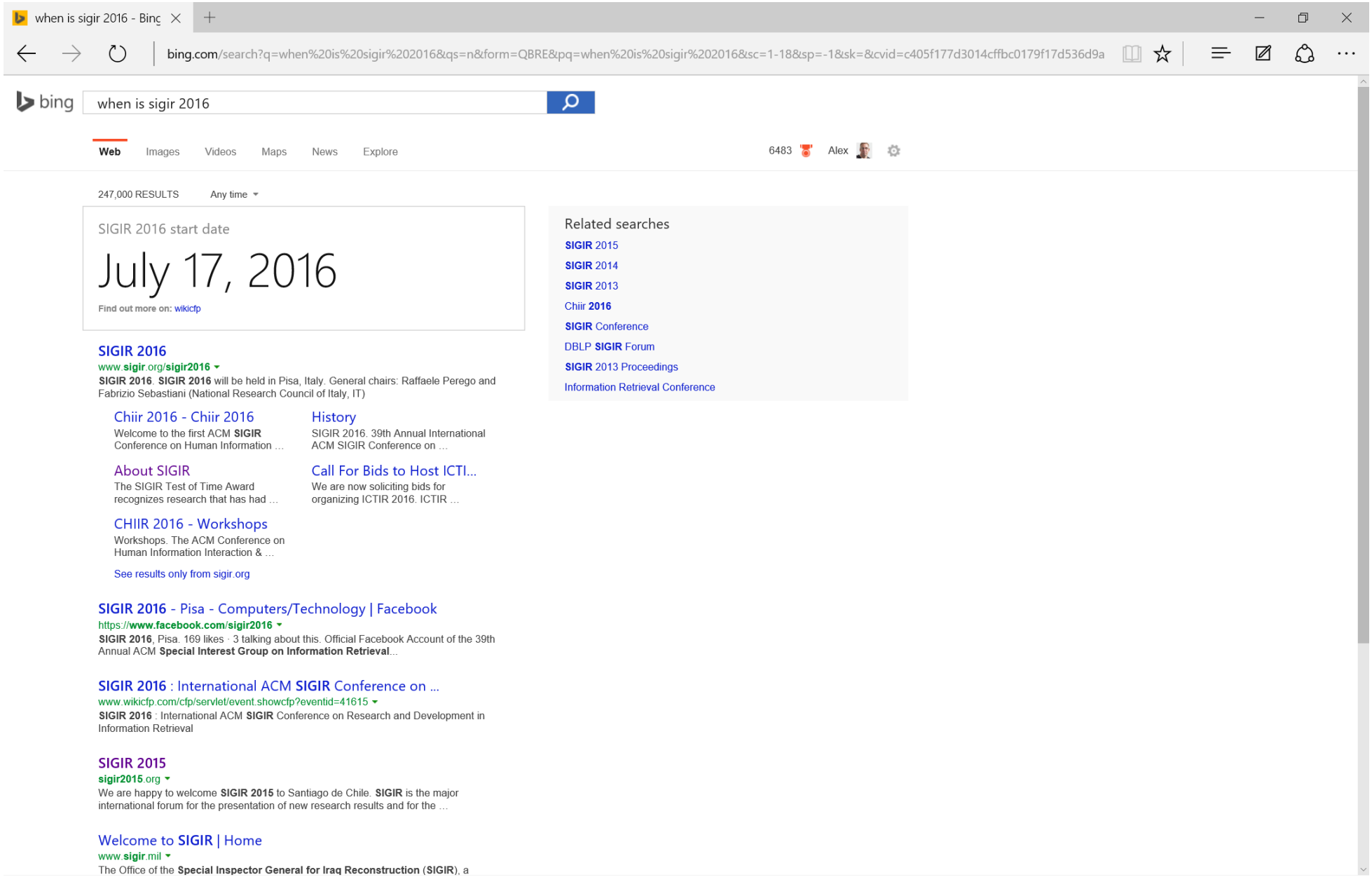
1991: Adaptive mixtures of local experts written by Michael I. Jordan was first published in 1991.

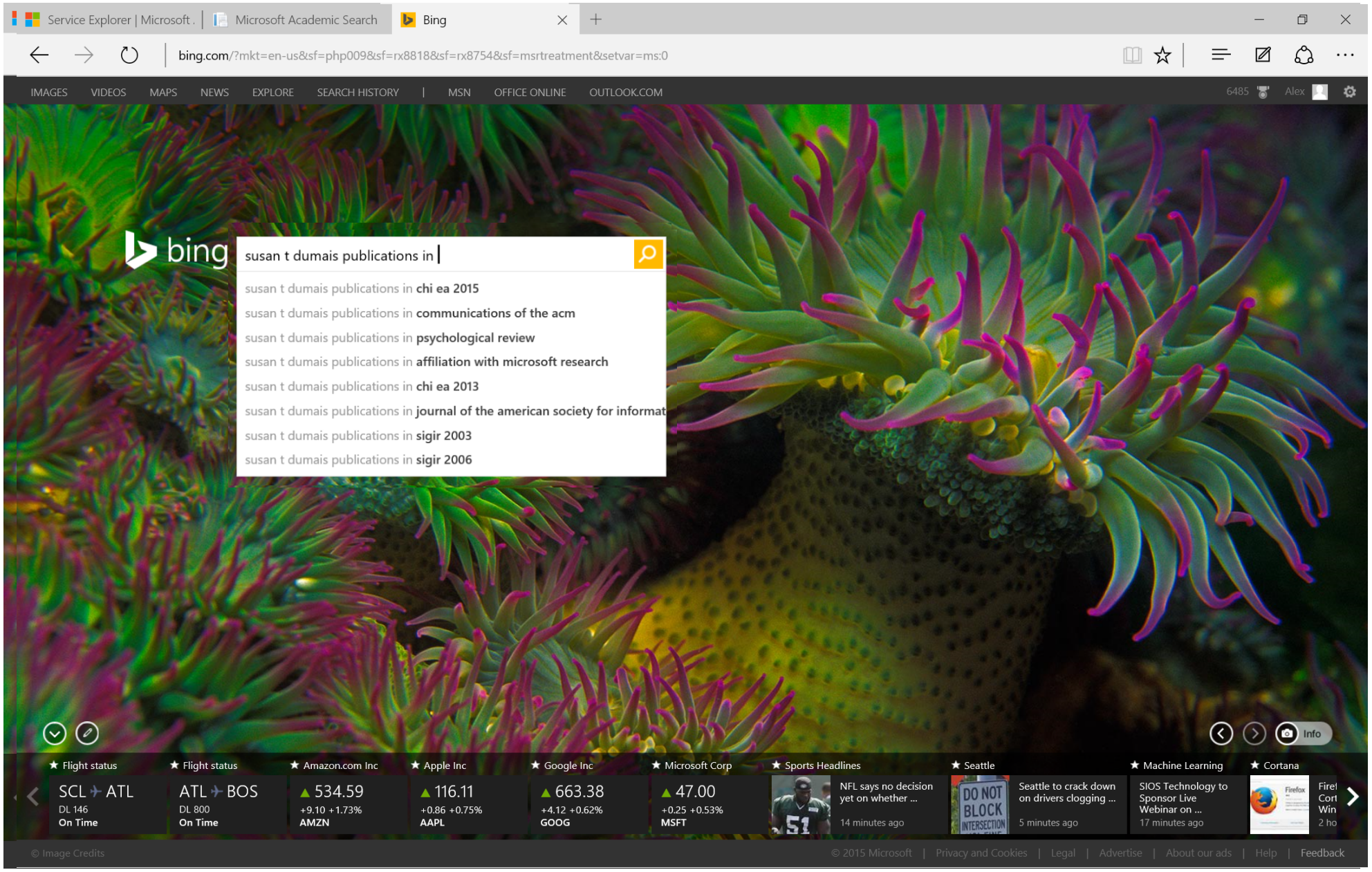
2003: Latent dirichlet allocation written by Michael I. Jordan was first published on March 01, 2003.

See more

Data from: Wikipedia · Microsoft · Berkeley · Freebase

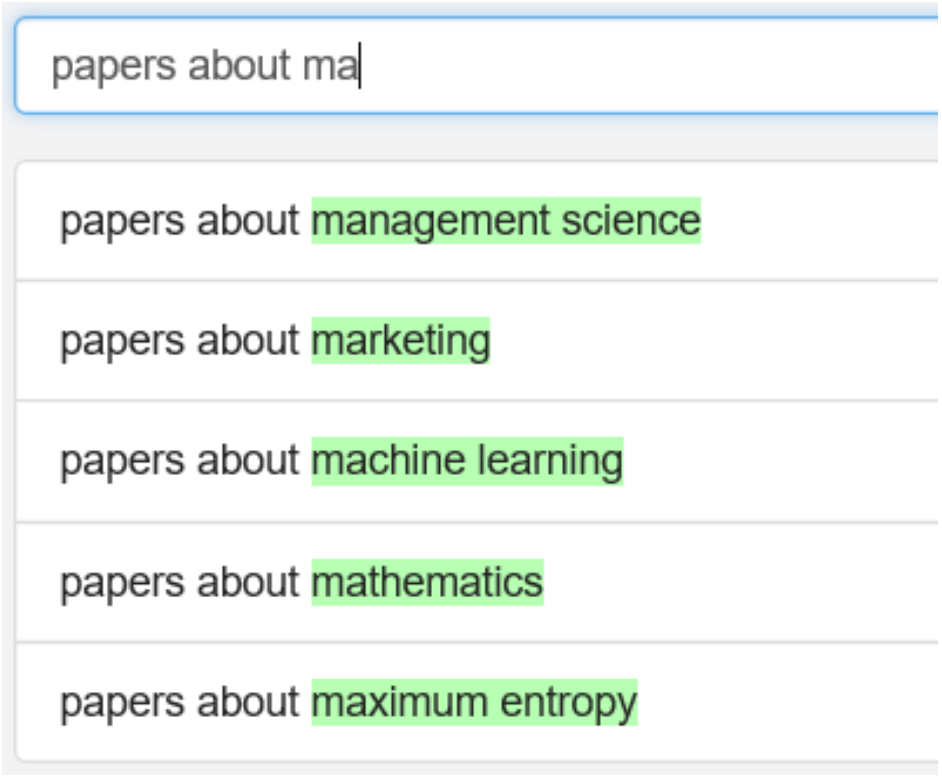
Feedback


when is sigir 2016 - Binç
bing.com/search?q=when%20is%20sigir%202016&qs=n&form=QBRE&pq=when%20is%20sigir%202016&sc=1-18&sp=-1&sk=&cvid=c405f177d3014cffbc0179f17d536d9a
bing
when is sigir 2016
Web
Images
Videos
Maps
News
Explore
6483
Alex
247,000 RESULTS
Any time
SIGIR 2016 start date
July 17, 2016
Find out more on: wikicfp
SIGIR 2016
www.sigir.org/sigir2016
SIGIR 2016. SIGIR 2016 will be held in Pisa, Italy. General chairs: Raffaele Perego and Fabrizio Sebastiani (National Research Council of Italy, IT)
Chiir 2016 - Chiir 2016
Welcome to the first ACM SIGIR Conference on Human Information ...
History
SIGIR 2016. 39th Annual International ACM SIGIR Conference on ...
About SIGIR
The SIGIR Test of Time Award recognizes research that has had ...
Call For Bids to Host ICTI...
We are now soliciting bids for organizing ICTIR 2016. ICTIR ...
CHIIR 2016 - Workshops
Workshops. The ACM Conference on Human Information Interaction & ...
See results only from sigir.org
SIGIR 2016 - Pisa - Computers/Technology | Facebook
https://www.facebook.com/sigir2016
SIGIR 2016, Pisa. 169 likes · 3 talking about this. Official Facebook Account of the 39th Annual ACM Special Interest Group on Information Retrieval...
SIGIR 2016 : International ACM SIGIR Conference on ...
www.wikicfp.com/cfp/servlet/event.showcfp?eventid=41615
SIGIR 2016 : International ACM SIGIR Conference on Research and Development in Information Retrieval
SIGIR 2015
sigir2015.org
We are happy to welcome SIGIR 2015 to Santiago de Chile. SIGIR is the major international forum for the presentation of new research results and for the ...
Welcome to SIGIR | Home
www.sigir.mil
The Office of the Special Inspector General for Iraq Reconstruction (SIGIR), a
Related searches
SIGIR 2015
SIGIR 2014
SIGIR 2013
Chiir 2016
SIGIR Conference
DBLP SIGIR Forum
SIGIR 2013 Proceedings
Information Retrieval Conference

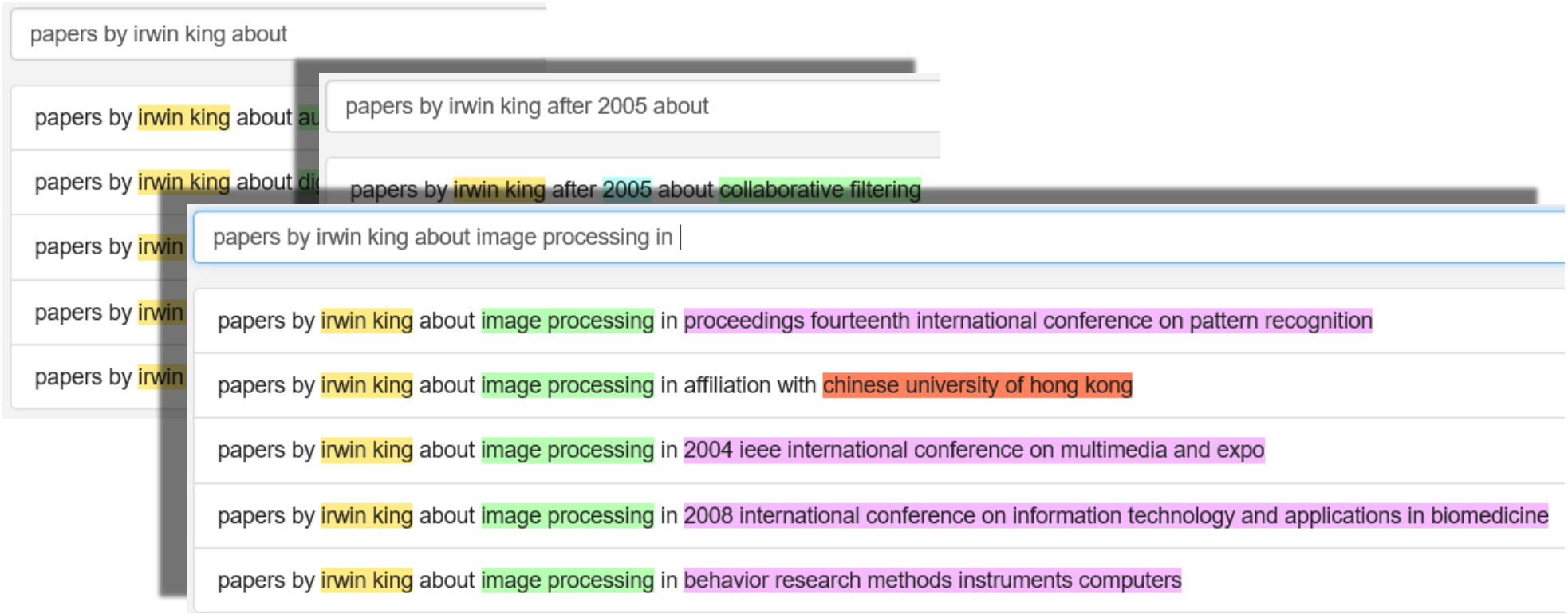

Service Explorer | Microsoft .
Microsoft Academic Search
Bing
bing.com/?mkt=en-us&sf=php009&sf=rx8818&sf=rx8754&sf=msrtreatment&setvar=ms:0
IMAGES VIDEOS MAPS NEWS EXPLORE SEARCH HISTORY | MSN OFFICE ONLINE OUTLOOK.COM
6485 Alex
bing
susan t dumais publications in
susan t dumais publications in chi ea 2015
susan t dumais publications in communications of the acm
susan t dumais publications in psychological review
susan t dumais publications in affiliation with microsoft research
susan t dumais publications in chi ea 2013
susan t dumais publications in journal of the american society for informat
susan t dumais publications in sigir 2003
susan t dumais publications in sigir 2006
Info
★ Flight status
SCL ✈ ATL
DL 146
On Time
★ Flight status
ATL ✈ BOS
DL 800
On Time
★ Amazon.com Inc
▲ 534.59
+9.10 +1.73%
AMZN
★ Apple Inc
▲ 116.11
+0.86 +0.75%
AAPL
★ Google Inc
▲ 663.38
+4.12 +0.62%
GOOG
★ Microsoft Corp
▲ 47.00
+0.25 +0.53%
MSFT
★ Sports Headlines
NFL says no decision yet on whether ...
14 minutes ago
★ Seattle
DO NOT BLOCK INTERSECTION
Seattle to crack down on drivers clogging ...
5 minutes ago
★ Machine Learning
SIOS Technology to Sponsor Live Webinar on ...
17 minutes ago
★ Cortana
2 ho
© Image Credits
© 2015 Microsoft | Privacy and Cookies | Legal | Advertise | About our ads | Help | Feedback

NL Query Completion/Recommendation

- How to complete never foreseen academic queries?
- How to rank completion suggestions?
- How to avoid making completions leading no search results?



Finding reviewers/co-PI made easier



Understanding a colleague made easier

Microsoft Academic Graph

- Background
  - Project Libra (2005 – 2008) Wei-Ying Ma, et al.
  - Microsoft Academic Search (2008 – 2012)
  - → Bing, Cortana (2015 –)

- Bing examples

- The Data

- The Challenge (WSDM Cup)

	Papers	Citations	Authors
Libra	~5M	~20M	~1M
MAS	~50M	~350M	~20M
Bing	~125M	~1B	~40M

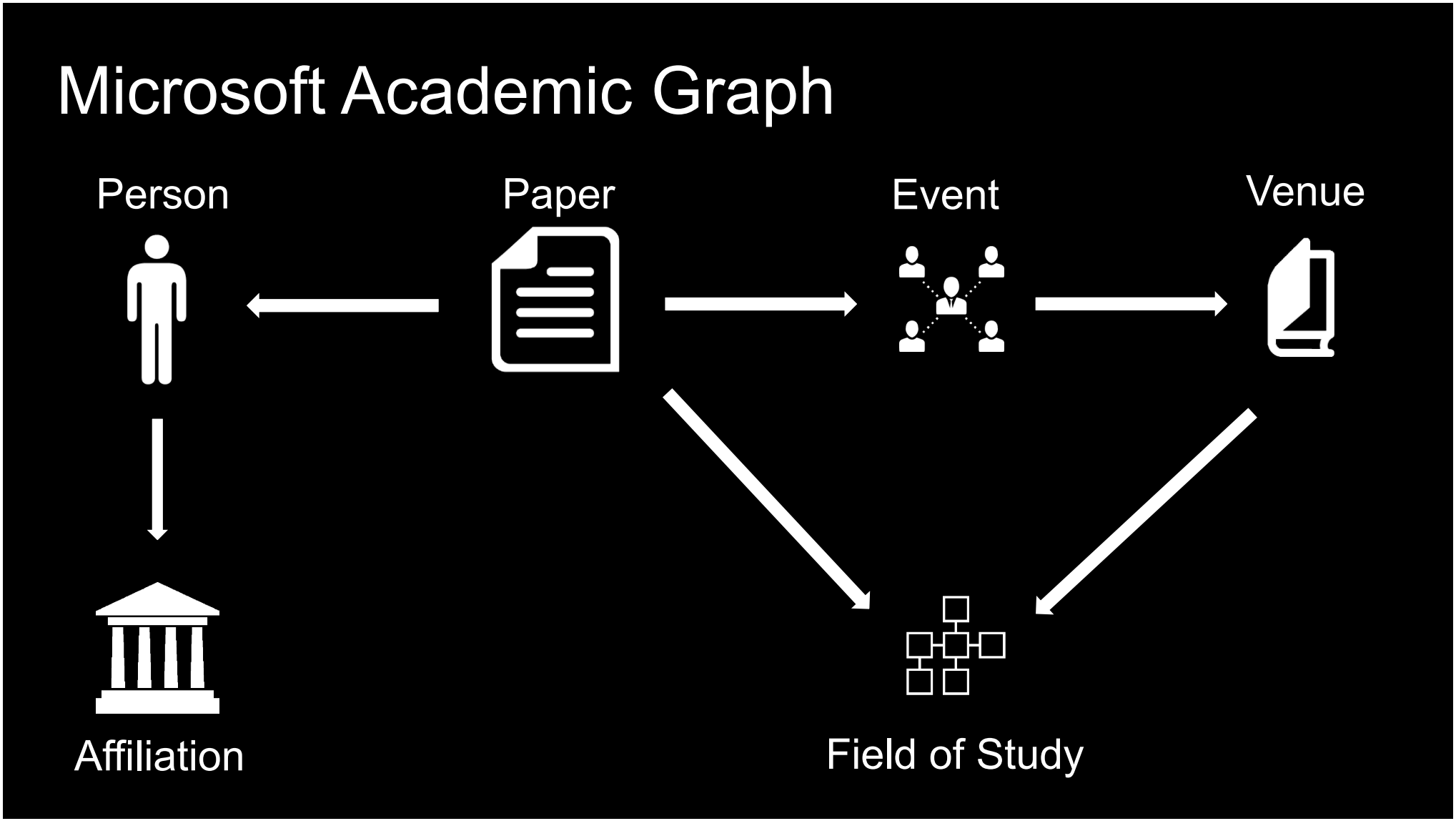
	Papers	Citations	Authors
Microsoft Academic	~40M	~250M	~20M
Microsoft Academic Graph	~100M	~700M	~25M

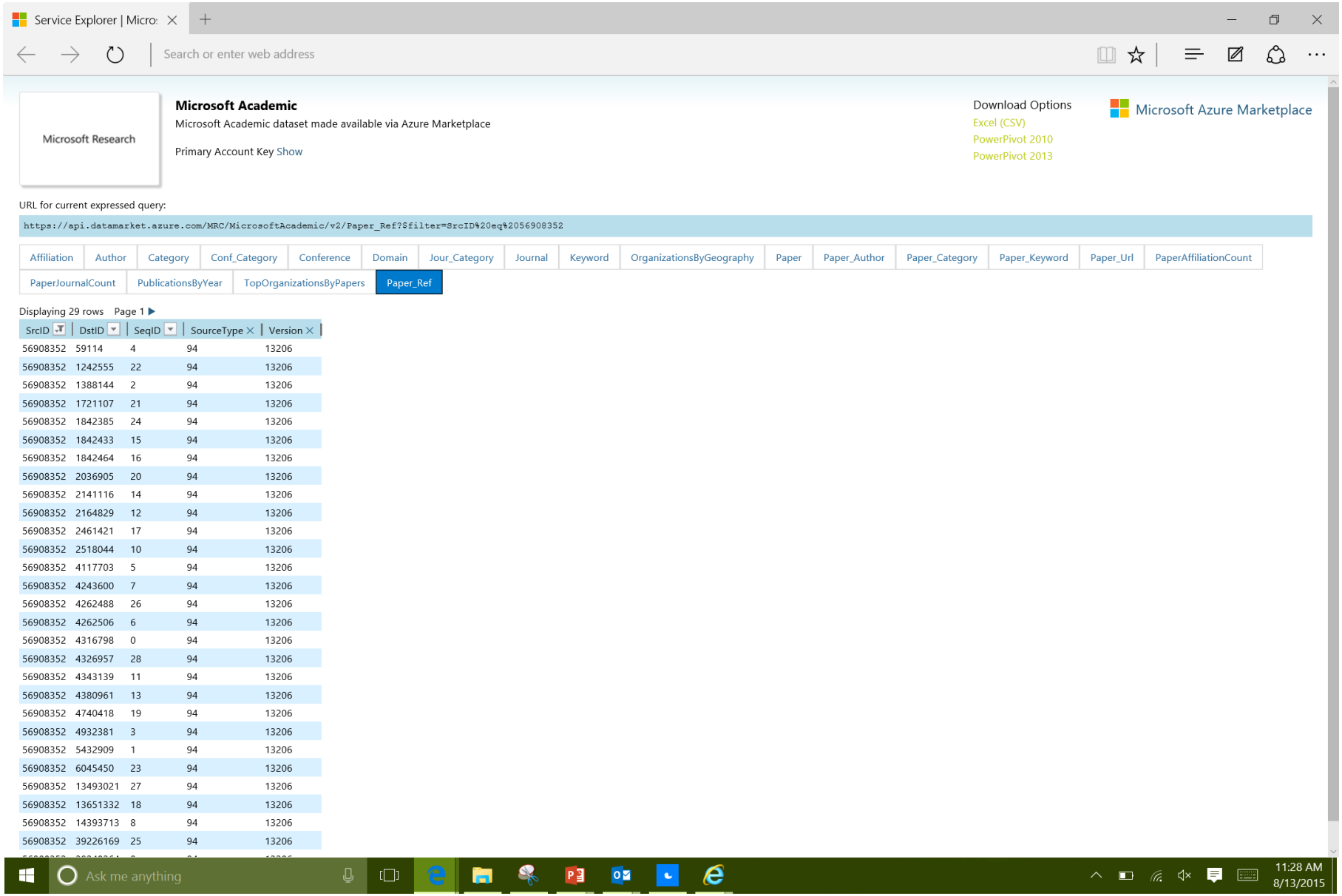
Data Releases

- Microsoft Academic (2012) - https://datamarket.azure.com/dataset/mrc/microsoftacademic
  - Azure Datamarket
  - REST API

- Microsoft Academic Graph (2015) - http://aka.ms/academicgraph
  - Increased coverage
  - Azure blob storage
  - No API (yet!)
  - Improved topic classification
  - More current (→ early 2015)

	Papers	Citations	Authors
Microsoft Academic	~40M	~250M	~20M
Microsoft Academic Graph	~100M	~700M	~25M


Microsoft Academic Graph
Person
Paper
Event
Venue
Affiliation
Field of Study


Service Explorer | Micro:
Search or enter web address
Microsoft Research
Microsoft Academic
Microsoft Academic dataset made available via Azure Marketplace
Primary Account Key Show
Download Options
Excel (CSV)
PowerPivot 2010
PowerPivot 2013
Microsoft Azure Marketplace
URL for current expressed query:
https://api.datamarket.azure.com/MRC/MicrosoftAcademic/v2/Paper_Ref?$filter=SrcID%20eq%2056908352
Affiliation
Author
Category
Conf_Category
Conference
Domain
Jour_Category
Journal
Keyword
OrganizationsByGeography
Paper
Paper_Author
Paper_Category
Paper_Keyword
Paper_Url
PaperAffiliationCount
PaperJournalCount
PublicationsByYear
TopOrganizationsByPapers
Paper_Ref
Displaying 29 rows
Page 1
SrcID
DstID
SeqID
SourceType
Version
56908352 59114 4 94 13206
56908352 1242555 22 94 13206
56908352 1388144 2 94 13206
56908352 1721107 21 94 13206
56908352 1842385 24 94 13206
56908352 1842433 15 94 13206
56908352 1842464 16 94 13206
56908352 2036905 20 94 13206
56908352 2141116 14 94 13206
56908352 2164829 12 94 13206
56908352 2461421 17 94 13206
56908352 2518044 10 94 13206
56908352 4117703 5 94 13206
56908352 4243600 7 94 13206
56908352 4262488 26 94 13206
56908352 4262506 6 94 13206
56908352 4316798 0 94 13206
56908352 4326957 28 94 13206
56908352 4343139 11 94 13206
56908352 4380961 13 94 13206
56908352 4740418 19 94 13206
56908352 4932381 3 94 13206
56908352 5432909 1 94 13206
56908352 6045450 23 94 13206
56908352 13493021 27 94 13206
56908352 13651332 18 94 13206
56908352 14393713 8 94 13206
56908352 39226169 25 94 13206
Ask me anything
11:28 AM
8/13/2015

Microsoft Research

Search Microsoft Research

Our research | Connections | Careers | About us

All | Downloads | Events | Groups | News | People | Projects | Publications | Videos

Microsoft Academic Graph

The Microsoft Academic Graph is a heterogeneous graph containing scientific publication records, citation relationships between those publications, as well as authors, institutions, journals and conference "venues" and fields of study. This data is available as a set of zipped text files stored in Microsoft Azure blob storage and available via HTTP. The file size is ~37GB.

We encourage researchers working with this data to apply for an Azure for Research Award. Details are available on the Azure for Research award submission process at http://azure4research.com. Please include the hashtag #academicgraph in your submission title for easier tracking.

To download the data you need to first agree to the terms of use.

☐ I agree to abide by the terms of use for the Microsoft Academic Graph. Get the data!

Please note: in order to emphasize one important technical challenge that is common in web-scale data collection and aggregation, the data released here have undergone only rudimentary processing, for example in areas of author and paper conflation/deduplication. This noisy yet realistic dataset can provide additional avenues for research in the big data arena.

We kindly request that any published research that makes use of this data cites our data paper listed below.

Publications

- Arnab Sinha, Zhihong Shen, Yang Song, Hao Ma, Darrin Eide, and Kuansan Wang, An Overview of Microsoft Academic Service (MAS) and Applications, WWW – World Wide Web Consortium (W3C), 18 May 2015.

http://aka.ms/academicgraph

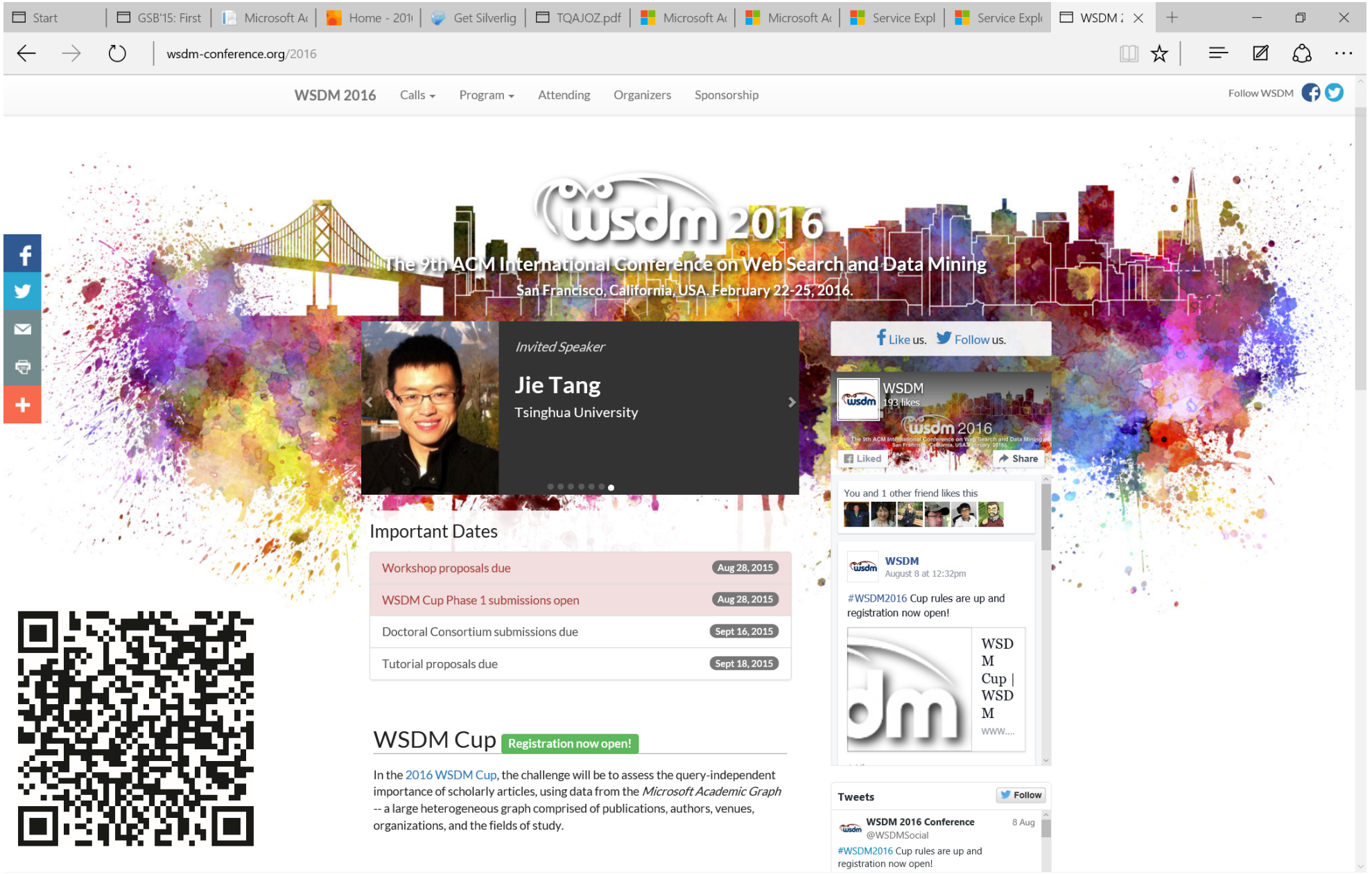
Microsoft Academic Graph

- Background
  - Project Libra (2005 – 2008) Wei-Ying Ma, et al.
  - Microsoft Academic Search (2008 – 2012)
  - → Bing, Cortana (2015 –)
- Bing examples
- The Data
- The Challenge (WSDM Cup)

	Papers	Citations	Authors
Libra	~5M	~20M	~1M
MAS	~50M	~350M	~20M
Bing	~125M	~1B	~40M

	Papers	Citations	Authors
Microsoft Academic	~40M	~250M	~20M
Microsoft Academic Graph	~100M	~700M	~25M

SIGIR2015
SANTIAGO - CHILE

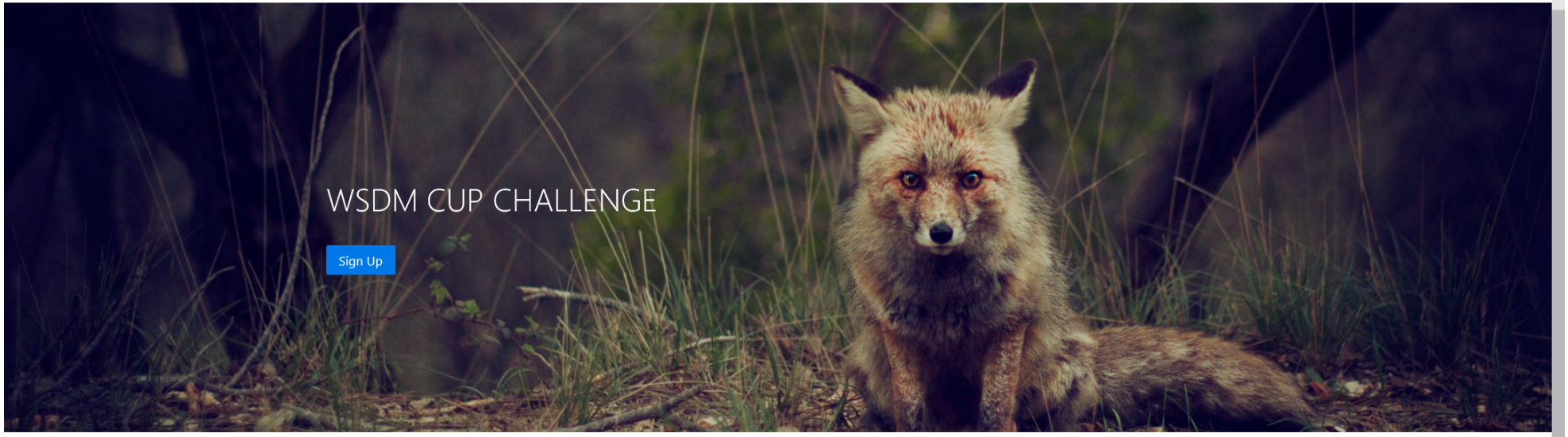

Start
GSB'15: First
Microsoft Ac
Home - 201
Get Silverlig
TQAJOZ.pdf
Microsoft Ac
Microsoft Ac
Service Expl
Service Explo
WSDM 2
wsdm-conference.org/2016
WSDM 2016
Calls
Program
Attending
Organizers
Sponsorship
Follow WSDM
wsdm 2016
The 9th ACM International Conference on Web Search and Data Mining
San Francisco, California, USA. February 22-25, 2016.
Invited Speaker
Jie Tang
Tsinghua University
Like us.
Follow us.
WSDM
193 likes
wsdm 2016
Liked
Share
You and 1 other friend likes this
WSDM
August 8 at 12:32pm
#WSDM2016 Cup rules are up and registration now open!
WSDM Cup | WSDM
WWW....
Important Dates
Workshop proposals due
Aug 28, 2015
WSDM Cup Phase 1 submissions open
Aug 28, 2015
Doctoral Consortium submissions due
Sept 16, 2015
Tutorial proposals due
Sept 18, 2015
WSDM Cup
Registration now open!
In the 2016 WSDM Cup, the challenge will be to assess the query-independent importance of scholarly articles, using data from the Microsoft Academic Graph -- a large heterogeneous graph comprised of publications, authors, venues, organizations, and the fields of study.
Tweets
Follow
WSDM 2016 Conference
@WSDMSocial
8 Aug
#WSDM2016 Cup rules are up and registration now open!

Start | GSB'15: First In | Microsoft Aca | Home - 2 | TQAJOZ.pdf | Microsoft Aca | Microsoft Aca | Service Explor | Service Explor | WSDM 2016. S

wsdmcupchallenge.azurewebsites.net

Microsoft Research

Home Rules Leaderboard WSDM 2016 Log in



The Graph

The Microsoft Academic Graph is a heterogeneous graph containing scientific publication records, citation relationships between publications, as well as authors, institutions, journal and conference "venues," and fields of study.

The Data

This data is available as a set of zipped text files stored in Microsoft Azure blob storage and available via HTTP. The file size (zipped) is ~37GB. The data provided for the challenge may be access and downloaded here.

The Challenge

The goal of the Ranker Challenge is to assess the query-independent importance of scholarly articles, using data from the Microsoft Academic Graph--a large heterogeneous graph comprised of publications, authors, venues, organizations, and the fields of study.

Azure For Research

We encourage researchers working with this data to apply for an Azure for Research Award. Details are available on the Azure for Research award submission process at http://azure4research.com. Please include the hashtag #academicgraph in your submission title for easier tracking.

Microsoft Research

Questions?

Alex Wade

Microsoft Research

Microsoft

@alexwade

SIGIR2015
SANTIAGO - CHILE