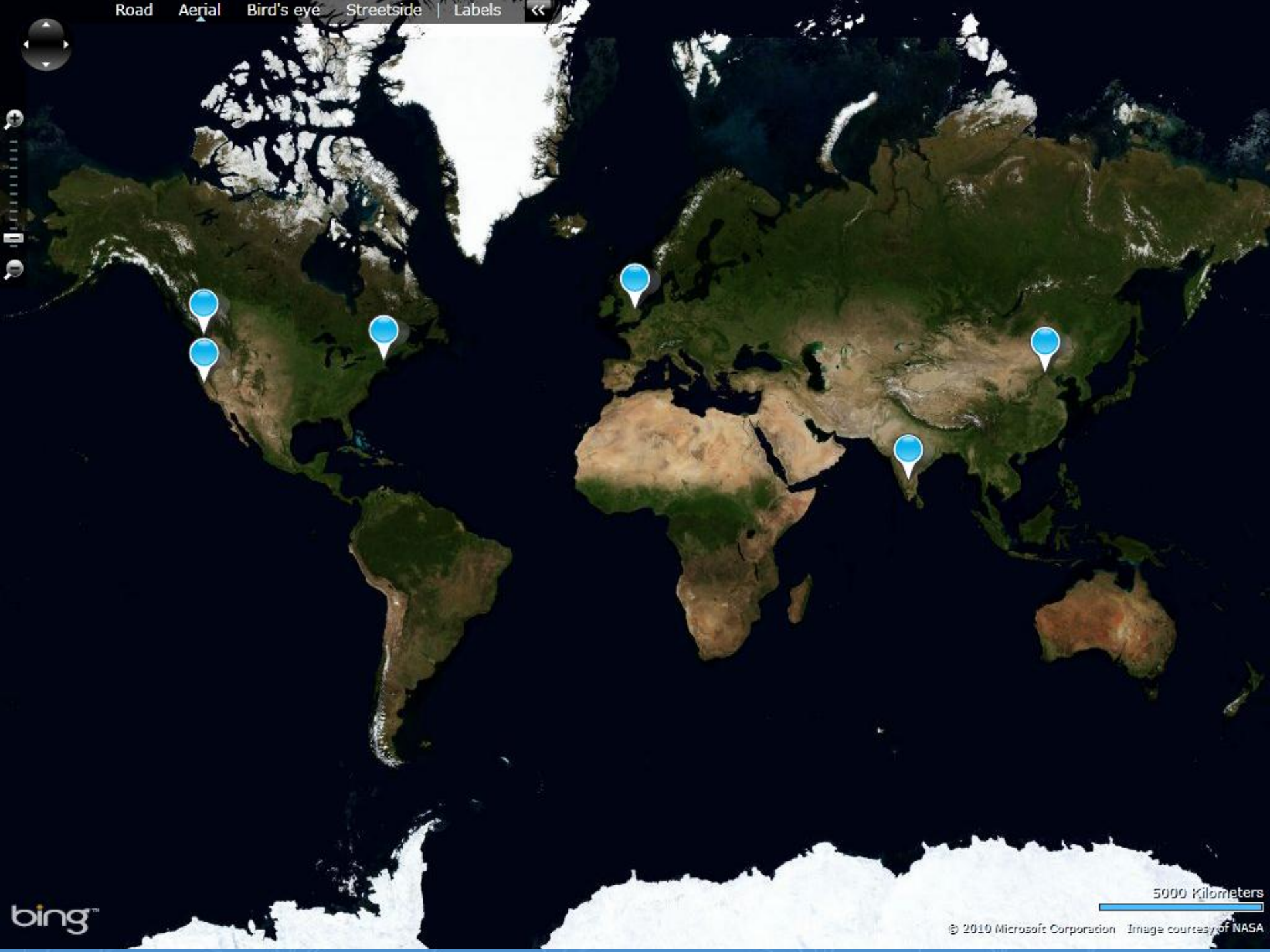
# Microsoft Research Connections
## Our work in the region

Dr Scarlet Schwiderski-Grosche
Scientific Manager

Microsoft Research Connections

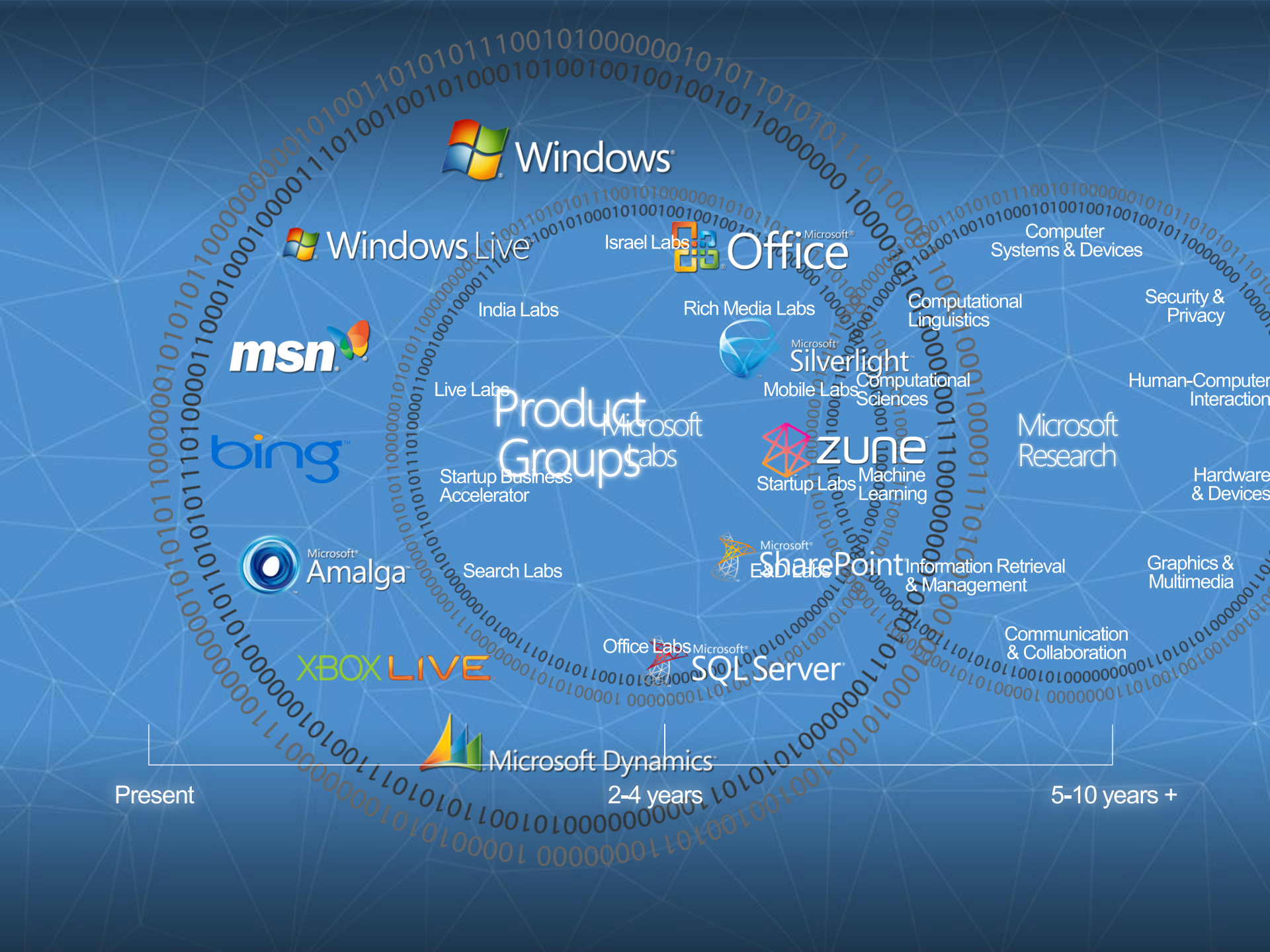
# Objective of this Talk

- To tell you more about Microsoft Research Connections
  - Global
  - EMEA
  - PhD Programme
  - Other engagements


Road
Aerial
Bird's eye
Streetside
Labels
5000 Kilometers
© 2010 Microsoft Corporation
Image courtesy of NASA

# Microsoft Research

- *Expand the state of the art in each of the areas in which we do research*
- *Rapidly transfer innovative technologies into Microsoft products*
- *Ensure that Microsoft products have a future*


Windows
Windows Live
Israel Labs
Office
Computer Systems & Devices
India Labs
Rich Media Labs
Computational Linguistics
Security & Privacy
msn
Silverlight
Live Labs
Mobile Labs
Computational Sciences
Human-Computer Interaction
Product Groups
Microsoft Labs
bing
zune
Microsoft Research
Startup Business Accelerator
Startup Labs
Machine Learning
Hardware & Devices
Amalga
Search Labs
SharePoint
E&D Labs
Information Retrieval & Management
Graphics & Multimedia
Office Labs
SQL Server
Communication & Collaboration
XBOX LIVE
Microsoft Dynamics
Present
2-4 years
5-10 years +

# Microsoft Research Connections

**Work broadly with the academic and research community to speed research, improve education, foster innovation and improve lives around the world.**

Support university research through collaborative partnerships

Accelerate university research and education through technology investments

Inspire the next generation of researchers and scientists

Drive awareness of Microsoft contributions to research

# Engagement and Collaboration Focus

**Core Computer Science**

**Natural User Interface**

**Earth Energy Environment**

**Education and Scholarly Communication**

**Health and Wellbeing**

**Research Accelerators**

- Worldwide Telescope
- Microsoft Biology Foundation
- Chemistry Add-in for Word
- Zentity
- Trident Workflow Workbench

**Global Partnerships**

- Centre for Computational and Systems Biology
- Microsoft Research-INRIA Joint Centre
- Microsoft Research Asia (MSRA) Joint Lab Program
- Microsoft & FAPESP

**People**

- Faculty Fellows
- Graduate Women Scholars
- Jim Gray eScience Award
- Student Internships
- ACM Student Research Competition

# Investment Focus

| Computer Science | Earth, Energy, and Environment | Education & Scholarly Communication | Health & Wellbeing |
|---|---|---|---|
| Programming, Tools, Mobile — *Judith Bishop*<br>Natural User Interfaces — *Kris Tolle* | WW Telescope, Climate Change Earth Sciences — *Dan Fay* | Academic Search, Digital Humanities, Publishing — *Lee Dirks* | MS Biology Foundation & Tools — *Simon Mercer* |

## Regional Outreach/Engagements

EMEA: ***Fabrizio Gagliardi***
LATAM: ***Jaime Puente***
India: ***Vidya Natampally***
Asia: ***Lolan Song***
America/Aus/NZ: ***Harold Javid***

## Engineering

High-quality and high-impact software release and community adoption

***Derick Campbell***

# EMEA

# Innovation Centres

## – Tarek Elabbady

### CMIC

- Bing Arab Countries Market Owners
- Best Arabic Natural Language Processing Technology Stack


### EMIC

- Embedded StreamInsight
- VENUS-C
- SWEPT


### ILDC

- New member of MSR family
- Telecoms, Security, Online services and Entertainment

# Regional Collaborations at Joint Institutes

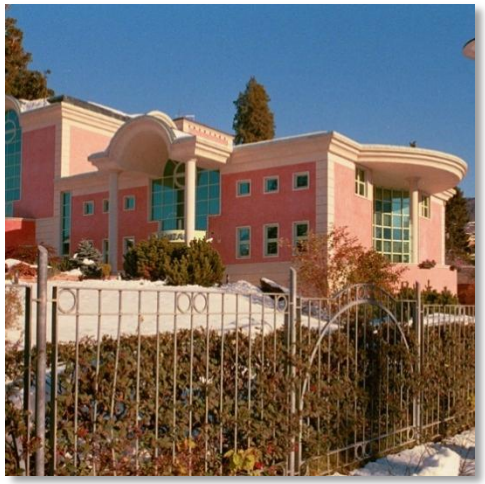
**INRIA, FRANCE**

SOFTWARE SECURITY

SCIENTIFIC INFORMATION

INTERACTION

**WWW.MSR-INRIA.INRIA.FR**



**UNIVERSITY OF TRENTO, ITALY**

COMPUTATIONAL TOOLS FOR SYSTEMS BIOLOGY

**WWW.COSBI.EU**

**BARCELONA SUPERCOMPUTING CENTRE, SPAIN**

MULTI CORE SYSTEMS

ARCHITECTURES AND PROGRAMMING

LANGUAGE RUNTIMES

**WWW.BSCMSRC.EU**

# Microsoft Research Connections EMEA

- Fabrizio Gagliardi (EMEA Director)
- Luisa Marie Küppers (EMEA Business Manager)
  - Noemie Elisa Laverne
- Scarlet Schwiderski-Grosche (EMEA Scientific Manager)
- Kenji Takeda (EMEA Technical Manager)
  - Julia Brading

Russia:

- Elena Pavlova

# Our Work in the Region

Flexible engagement approach

- Joint institutes
- PhD Scholarships
- Faculty Fellowships
- Conference sponsorships
- Internships
- Events (e.g. workshops)
- Collaborative projects


- Started in 2004
- EMEA academics apply with their research project
- Selected projects start in the following academic year
- Students are co-supervised by an MSRC researcher
- Students often do Internships at MSRC
- Around 25 students a year
- Over 200 PhD students in total (~ 100 active)



Special agreements with

- Max Planck Society in Germany
- UK Research Council (Dorothy Hodgkin Postgraduate Awards and CASE Studentships)
- Irish Research Council IRCSET

Currently exploring

- EU co-funding
- Agreement with University of Edinburgh

# PhD Summer School

- Networking
  - PhD Scholars
  - Students from joint labs (INRIA, BSC, CoSBi)
  - Students from Max Planck
  - Students sent from Innovation Centres
  - Students sent via Cambridge Computer Lab, CosmoComp ITN, Russia MRC
  - MSR researchers
  - Cambridge academics
- 'Transferable skills'
  - Write paper or poster, give talk, become an entrepreneur, apply for funding, career advice
- Research talks
  - Latest 'stuff' from MSR
- MRC projects talks/demo
- Poster sessions
- Social events

# MRC Engagements by Theme

- Computer Science
  - Project Hawaii
  - F#
  - Gadgeteer
- Earth, Energy and Environment
  - WW Telescope
- Natural User Interfaces
  - Kinect research projects
- Education & Scholarly Communication
  - Academic Search
- Health & Wellbeing
  - Microsoft Biology Foundation

# MRC Computer Science Topics

## Software Engineering

- Started Innovation Foundation (SEIF)
- Interaction with RiSE
- Promoting Visual Studio through ICSE events
- Summer School in 2011


## Semantic Computing

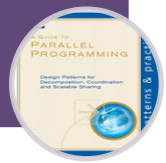
- Interaction with NLP, ISRC, Bing
- Promote cloud-based services
- Developed Web N-gram and language models
- Bing-MSR Speller Challenge
- Developing Knowledge

## Mobile Computing

- Interaction with DB team in MSR-R
- Student projects on phones at 17 universities
- Promoting Azure and WP7
- Prizes through Imagine Cup

## Parallelism and Concurrency

- Book with Ade Miller from the Patterns and Practices group
- Course on Concurrency and Parallelism with Tom Ball in RiSE
- Promoting C# and F#




## Programming Languages

- Interaction with Don Syme on F#
- Commissioned F# courseware from Imperial
- Developing a Try F# system in a browser
- Workshop on "F# in Education"
- Supporting Pex4Fun

## Community interaction

- CS EdWeek
- Work with the NSF, CRA, ACM, IEEE and IFIP
- MSR Software Summit in Paris
- Program for Interns joining the Redmond Lab


CS is research-oriented with strong internal and external impact

# Project Hawaii

Effort to investigate the ability of the **cloud** to enhance end-user experience on **mobile** devices

- To unleash the creative power of students by lowering barriers to writing mobile + cloud apps

# What does Hawaii offer?

## Microsoft Research

Cloud services

- Relay
- Rendezvous
- Speech to Text
- OCR in the cloud
- Compute in Windows Azure
- Storage in Windows Azure

Development Environment (SDKs)

Mobile devices (Windows Phone 7)

## Universities

# Spring 2011 Semester

- Launched Hawaii in **21** universities
- Close to **300** students start using Hawaii (with Windows Phone 7)

STANFORD UNIVERSITY · ILLINOIS · UNIVERSITY OF ARKANSAS · UNIVERSITY OF MINNESOTA Driven to Discover℠ · UCL · UNIVERSITÄT LEIPZIG · PURDUE UNIVERSITY · Duke UNIVERSITY · UC SANTA BARBARA UNIVERSITY OF CALIFORNIA · UNIVERSITY OF MICHIGAN · USC Viterbi School of Engineering

**Spring 2011**

**University College London**


Mobile and Cloud Computing, taught by Brad Karp and Kyle Jamieson

**Duke University**

Wireless Networking and Mobile Computing, taught by Romit Choudhury

**University of Minnesota**

Fundamentals of Advanced Networking, taught by Zhi-Li Zhang

**New York University**

TBA, taught by Lakshminarayanan Subramanian

**Stony Brook University**

TBA, taught by Xin Wang

**Stanford University**

Computer Science Innovation, taught by Jay Borenstein

**University of Arkansas**

Hot Topics in Mobile and Pervasive Computing, taught by Nilanjan Banerjee

**University of Illinois at Urbana-Champaign**

Extending Mobile Computing through Cloud Computing, taught by Yih-Chun Hu

**University of Massachusetts Lowell**

Data Communications, taught by Benyuan Liu

**University of Houston**

Advanced Distributed Computing: Mobile Computing Riding on the Cloud, taught by Rong Zheng

**University of California Santa Barbara**

Mobile Computing, taught by Elizabeth M. Belding

**Temple University**

TBA, taught by Jie Wu

**University of California Santa Barbara**

Network Programming, taught by Ben Y. Zhao

**Indiana University Purdue University Indianapolis**


Advance Mobility and Cloud Computing, co-taught by Arjan Durresi of IUPUI and Raj Jain of WUSTL

**University of Goettingen**

TBA, taught by Xiaoming Fu

**The Ohio State University**

TBA, taught by Dong Xuan

**Purdue University**

Software Development for Mobile Devices I, taught by Kyle D. Lutes

**University of Leipzig, Germany**

TBA, taught by Prof. Dr.-Ing. Christoph Lindemann

**Pontificia Universidade Catolica, Brasil**

Web Engineering, taught by Karin Breitman

**Egypt-Japan University of Science and Technology, Egypt**

Mobile Computing, taught by Moustafa A. Youssef

**University of Washington**



CSE 481M: Home Networking Capstone, co-taught by Ratul Mahajan, David Wetherall and John Zahorjan

mobile computing research center

Microsoft Research Connections

# What are the student building?

http://research.microsoft.com/en-us/um/redmond/projects/hawaii/applications/

## Flagged Down

Flagged Down is a mobile application that aims to allow users to search and hail cabs within their vicinity.


**Singapore Management University, School of Information Systems:** *Alex Chng, Hendry Poh, Nicholas Szetoh, Tan Jun Ming, Bevan.*

## Smart Bike Pedometer

**Smart Bike Pedometer:** GPS & Accelerometer Based, Real-time Information: Burned Calories; Elapsed Time; Distance; Speed

**On the Phone:**Data Acquisition, User Interface.
**In the Cloud:** Record Keeping, Route Sharing, Road Information Datasets, Route Condition Inference Engine.


**University of Southern California:** *Bo-Chun Wang.*

## Parking Assistant

**Crowd-Sourced Parking:** Driven by Social Networking. Users Rank Lots: Parking Availability, Cost, Convenience, System Find Available Spots Close to User.

**On the Phone:** Location Tracking. User Interface
**In the Cloud:** System State. Space Ranking Algorithm


**University of Southern California:** *Gaurav Sanghavi, Nirmit Desai.*

# MRC Engagements by Theme

- Computer Science
  - Project Hawaii
  - F#
  - Gadgeteer
- Earth, Energy and Environment
  - WW Telescope
- Natural User Interfaces
  - Kinect research projects
- Education & Scholarly Communication
  - Academic Search
- Health & Wellbeing
  - Microsoft Biology Foundation

# World Wide Telescope

**Seamless Rich Social Media Virtual Sky**
**Web application for science and education**

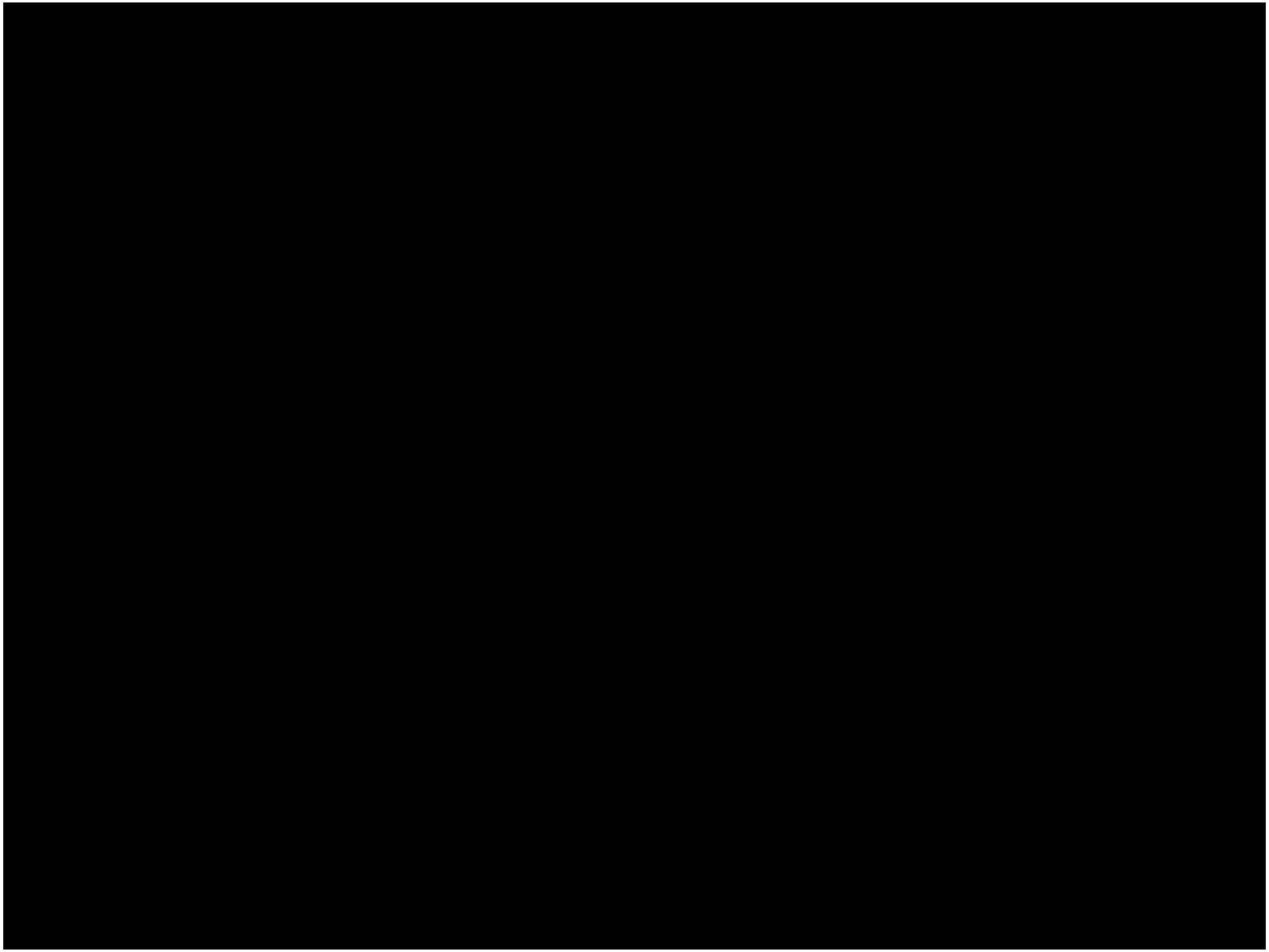
- Science- Seamless integration of multi-wavelength, multiple telescope distributed image/data sets and one click contextual access to distributed web information/data sources
- Education- Easy as Powerpoint, rich social media authoring environment within the sky allowing astronomers, educators and kids to create and share rich narrated guided tours of the universe

**ID magazine** *International Design Annual*
*"Best in category; Interactive 2009"*

**TIME magazine**
*"50 Best sites on the Internet 2009"*


**Harvard-Smithsonian:** Alyssa Goodman
**Johns Hopkins University:** Alex Szalay
**Microsoft Research:** Curtis Wong, Jonathan Fay

# NUI – Kinect SDK and WWT



http://www.youtube.com/watch?v=1-tMp4WkQjA

# Worldwide Telescop

***www.worldwidetelescope.org***


## Worldwide Telescope

Project: Seamless Astronomy at Harvard

Windows Client launched at TED'08

Silverlight Client launched at MIX'09

Over 6 Million unique visitors

TED'10 demo by Blaise to show Bing Maps SL integration with WWT SL

## WWT Outreach

WWT at center of China eclipse July '09

Localizations in 5 languages

Community Servers in China & Japan

WWT Coursework developed

Galileo Tour celebrating 400th anniversary launched

WWT Ambassadors program (Harvard & WGBH) NSF funding

## NASA Space Act Agreement

NASA provide content in WWT format for Moon and Mars – launch March '10

SAA allowed MS to have more combined marketing

PDC Azure Demo by DPE – Be A Martian, leveraged GalaxyZoo effort

NASA Explorer Schools to adopt WWT Planetarium

## WWT|Earth

Visualize environmental datasets

Bring gaming experience to environmental data

Have high-end Rich Internet App to complement Bing Maps

Prototype demo'd at AGU'09

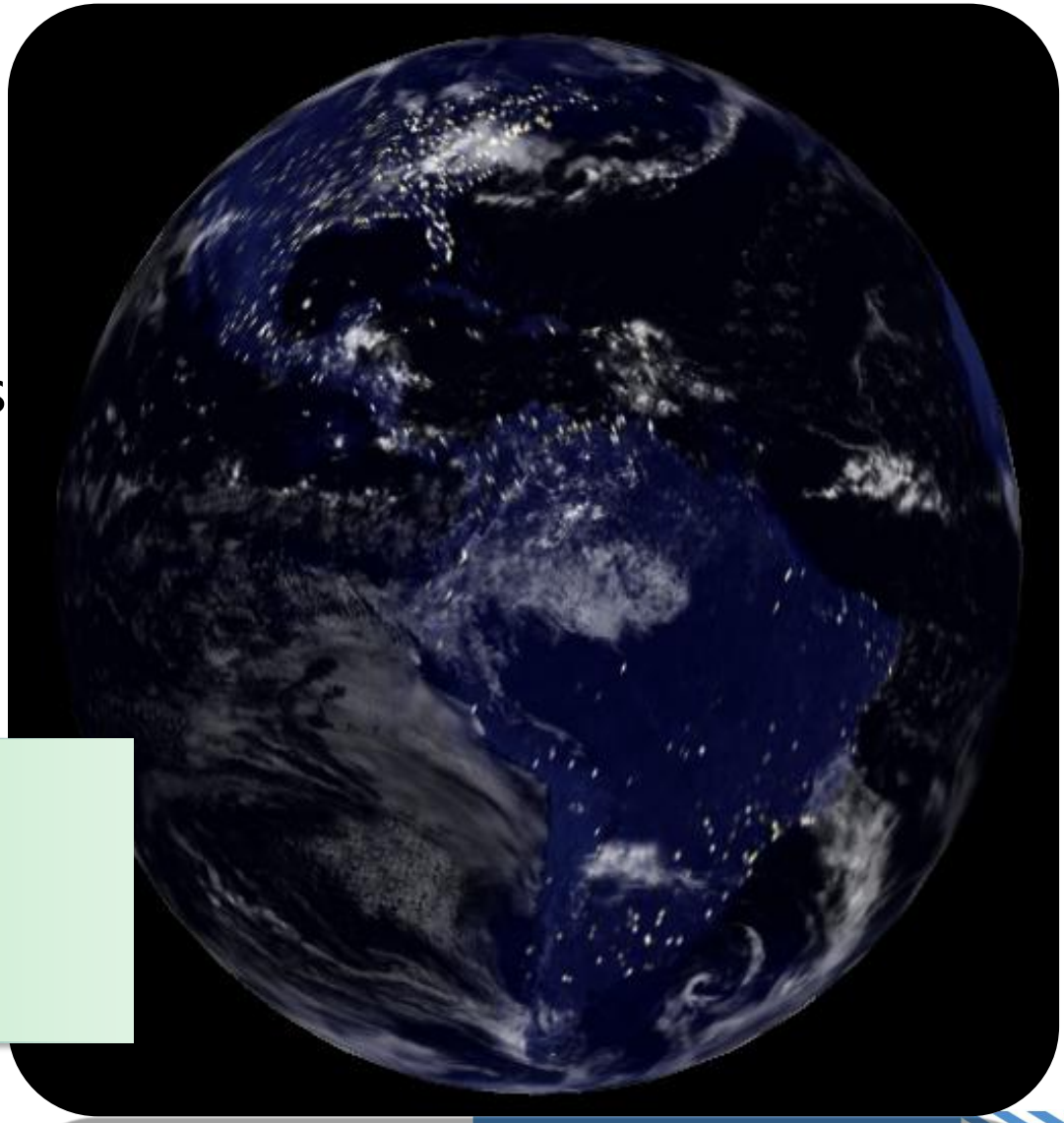
Re-architecture building on Win7, DX11, etc – exposing API

3D Scientific exploration and inquiry tool

# Earth, Energy, and Environment

# Worldwide Telescope | Earth

- A seamless visual environment
- Sky and earth-based visualizations
- Create and share tours of your data
- Run demo



*Introducing an Excel Add-in for geo-spatial data visualizations*

# MRC Engagements by Theme

- Computer Science
  - Project Hawaii
  - F#
  - Gadgeteer
- Earth, Energy and Environment
  - WW Telescope
- Natural User Interfaces
  - Kinect research projects
- Education & Scholarly Communication
  - Academic Search
- Health & Wellbeing
  - Microsoft Biology Foundation

# Natural User Interface

## Kinect Academic Program

- Fund academic projects and competitions around the Kinect SDK
- Leverage MCR Events, Subs and MSR Labs
- Rely on DPE/ADEs for broader reach

## Machine Translation Toolkit

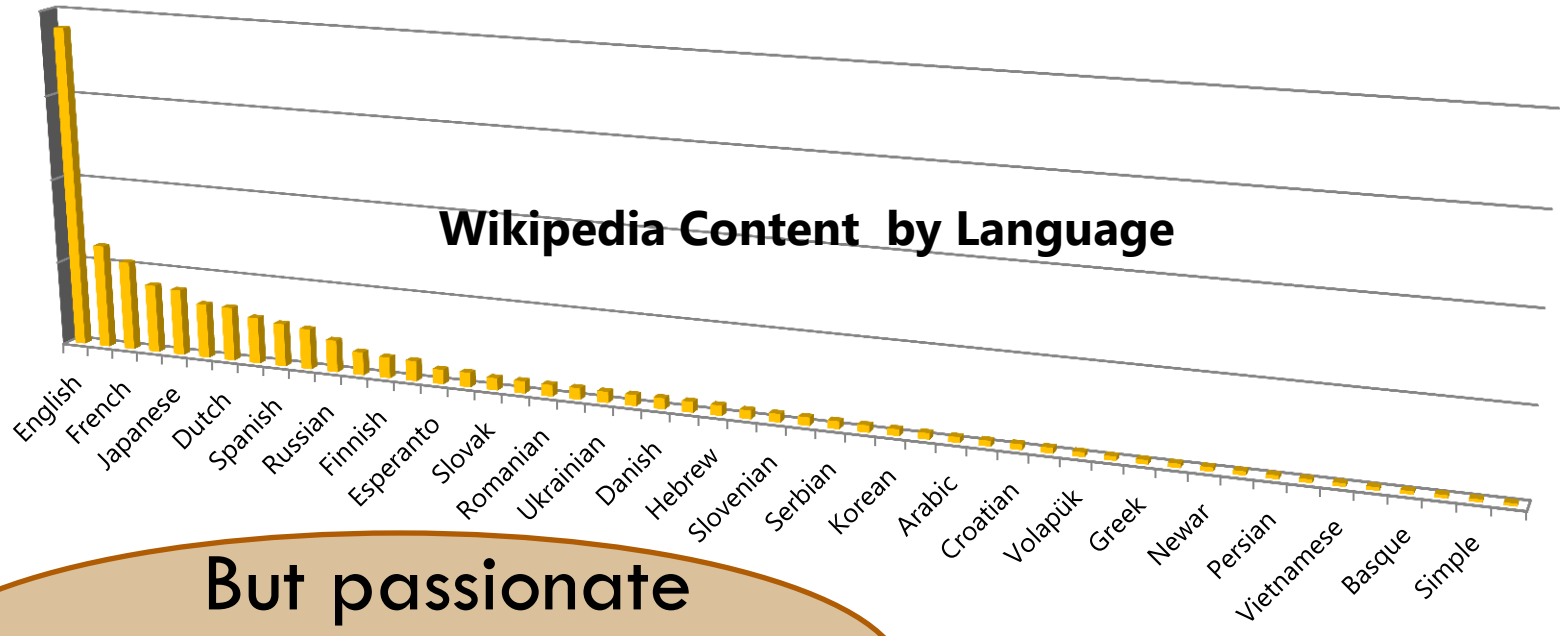
- MRC large "swim lane" project
- Leverage the MSR MT Team for community outreach and development
- Rely on MSR PM for ALL tech transfer work

## Select NUI Engagements

- Focus resources on the success of Kinect and MT
- Focus on Data Visualization
- Rely on regions to PM additional local projects

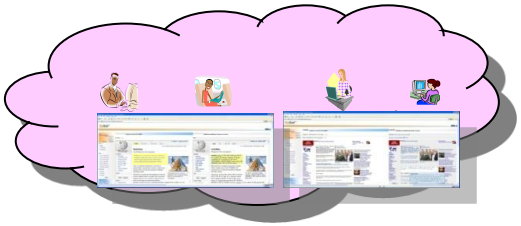
# A Special Data Source: Wikipedia

- Wikipedia is among the most valuable resources
  - LARGE: ~14M Articles
  - META-DATA rich: Mark-ups, Templates, Topics…
  - MULTILINGUAL: Available in 250+ languages
  - SEMANTICALLY LINKED: Articles linked by Semantics
- Wikipedia content is also skewed




But passionate local Wikipedia Communities …

# The WikiBhasha Project: A symbiotic mix of Machine and Human Translation…



Community of users to create content using MT output

WikiBABEL on Wikipedia

WikiBabel BETA
POWERED BY PEOPLE

Article to Target Wikipedia

Machine Translation System

Collaborative Translation Cache

Linguistic Resources

Machine Translation used to create rough content

**Wikipedian's Perspective:** Faster creation of Multilingual content

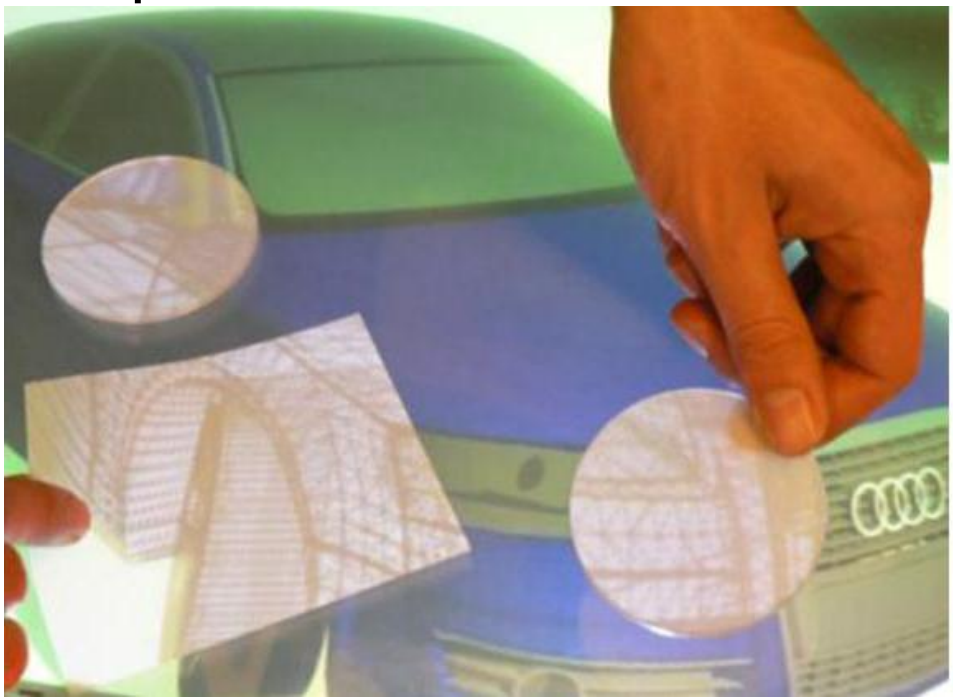
**Technology Perspective:** User corrections captured as data

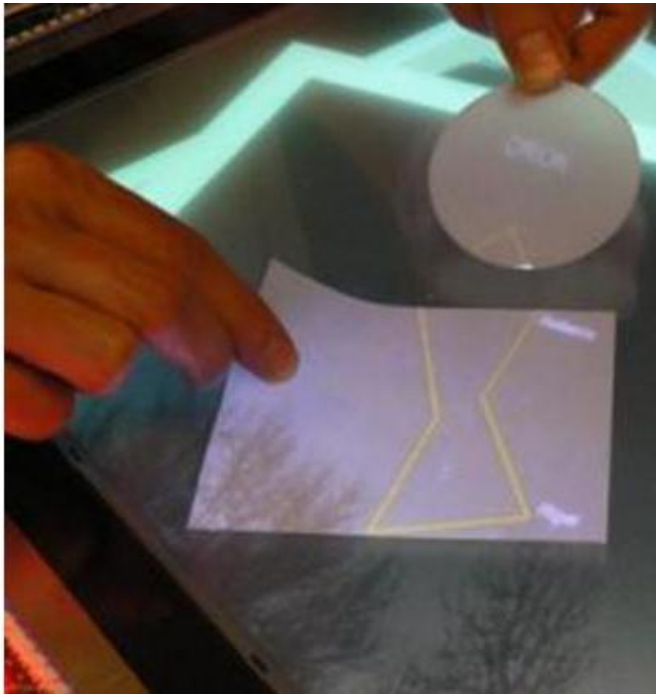
# SenseCam: A Wearable Automatic Digital Camera

- Simple device
  - wide-angle lens
  - automatic capture
  - range of sensors
- Applications
  - remarkably powerful cue for recall
  - security, police
  - activity monitoring
  - sharing experiences
- Large research community
  - over 30 labs using SenseCam WW
  - medical; computer science; social science etc.
  - annual SenseCam Symposium
- Available commercially
  - www.viconrevue.com

# SecondLight: Interaction Beyond the Surface

- New type of rear-projection multi-touch surface computing technology
- Second image displayed through electrically switchable projection screen
- A camera can see through the surface, sensing what happens above, be used for gesture input.





# MRC Engagements by Theme

- Computer Science
  - Project Hawaii
  - F#
  - Gadgeteer
- Earth, Energy and Environment
  - WW Telescope
- Natural User Interfaces
  - Kinect research projects
- Education & Scholarly Communication
  - Academic Search
- Health & Wellbeing
  - Microsoft Biology Foundation

# Education and Scholarly Communication


***The Scholarly Communication Lifecycle***

**Chem4Word**
Semantic chemistry for students and publishers

**Project Tuva**
Enhanced Video Player

**Research Information Centre (RIC)**
Virtual Research Environment (VRE) Toolkit for SharePoint

**Project Garibaldi**
Large Art Display on a Surface (LADS)

**Zentity**
Semantically-enabled repository software

**Project Trident**
Scientific Workflow Workbench

**Academic Search**

# Microsoft Academic Search

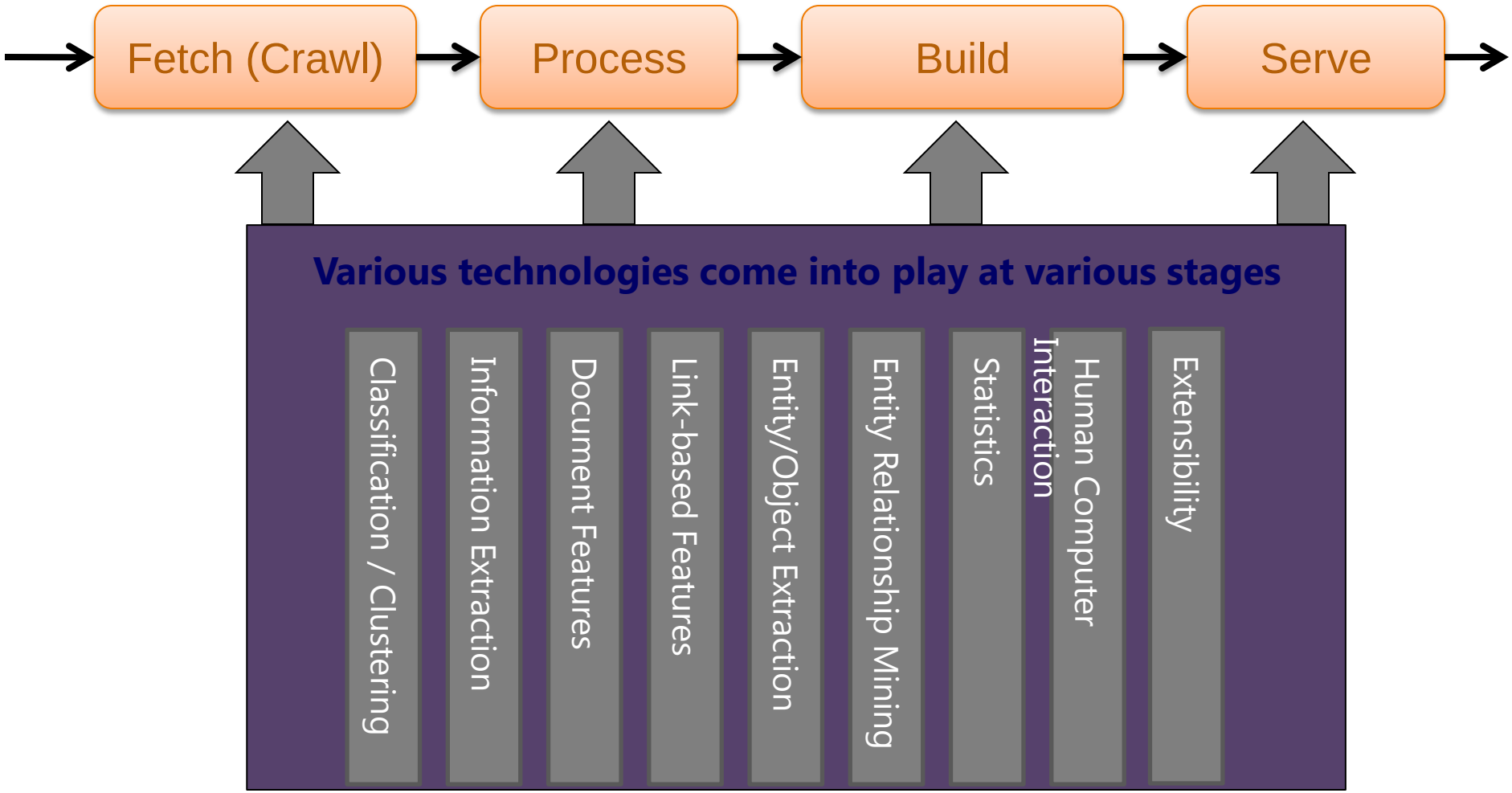
Microsoft Academic Search is a free academic search engine developed by Microsoft Research

- Easily search the top papers, authors, conferences, and journals for a topic
- See details about a specific paper, author, conference, journal or organization
- Quickly explore relationships between authors
- Discover influential papers, authors, conferences, journals and organizations within a domain
- Get the latest call for papers

# From Web Pages to Web Entities

- Entity search and knowledge mining
  - Web-scale entity extraction, integration, and summarization
  - Entity relationship mining
  - Entity ranking
- Academic search as an example
  - Researchers, papers, organizations, conferences, journals
  - Knowledge and insights
  - Visualization & exploration

# Technologies in Academic Search



# Technologies in Academic Search

- Object-Level Ranking: Bringing Order to Web Objects
  Zaiqing Nie, Yuanzhi Zhang, Ji-Rong Wen, and Wei-Ying Ma
  *In the Proceedings of the 14th international World Wide Web conference* (**WWW 2005**), *May 10-14, 2005, in Chiba, Japan.*
- Object-Level Vertical Search
  Zaiqing Nie, Ji-Rong Wen, and Wei-Ying Ma
  *In the Third Biennial Conference on Innovative Data Systems Research* (**CIDR 2007**, *May 10-14, research paper*).
- Extracting Objects from the Web
  Zaiqing Nie, Fei Wu, Ji-Rong Wen, Wei-Ying Ma
  *In the 22nd International Conference on Data Engineering* (**ICDE 2006**, *poster paper* ).
- Simultaneous Record Detection and Attribute Labeling in Web Data Extraction
  Jun Zhu, Zaiqing Nie, Ji-Rong Wen, Bo Zhang, Wei-Ying Ma
  *In the 12th International Conference on Knowledge Discovery and Data Mining (***SIGKDD 2006***, full paper* ).
- 2D Conditional Random Fields for Web Information Extraction Jun Zhu, Zaiqing Nie, Ji-Rong Wen, Bo Zhang, Wei-Ying Ma
  *In the 22nd International Conference on Machine Learning (***ICML 2005***).*
- Web Object Retrieval Zaiqing Nie, Yunxiao Ma, Shuming Shi, Ji-Rong Wen, Wei-Ying Ma
  *In the Proceedings of the 16th international World Wide Web conference (***WWW 2007***).*

# Search by keyword

Author (25002)

Jiawei Han

Mohammed Javeed Zaki

Alex Alves Freitas

Philip S. Yu

Vipin Kumar

Conference (841)

KDD

ICDM

PAKDD

SDM

PKDD

Journal (689)

CORR

Sigkdd Explorations

ESWA

TKDE

DATAMINE

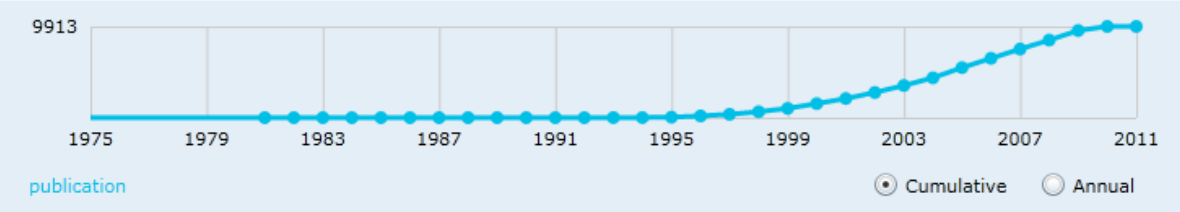
Academic > Keyword > Data Mining

Subscribe

This page shows one keyword best matching your query, you can find other results here.

**Data Mining**

Publications: 15,255 | Citation Count: 103,098




Publication (15255)

Churn Prediction in Telecommunication Using Data Mining Technology

Rahul J. Jadhav, Usharani T. Pawar

Published in 2011.

Clustering and ranking university majors using data mining and AHP algorithms: A case study in Iran

A. Rad, B. Naderi, M. Soltani

Journal: Expert Systems With Applications - ESWA, vol. 38, no. 1, pp. 755-763, 2011

Object Detection with Discriminatively Trained Part-Based Models (Citations: 5) View...

Pedro F. Felzenszwalb, Ross B. Girshick, David A. McAllester, Deva Ramanan

Journal: IEEE Transactions on Pattern Analysis and Machine Intelligence - PAMI, vol. 32, no. 9, pp. 1627-1645, 2010

Privacy Implications of Automated GPS Tracking and Profiling (Citations: 3) View...

Muhammad Usman Iqbal, Samsung Lim

Journal: IEEE Technology and Society Magazine - IEEE TECHNOL SOC MAG, vol. 29, no. 2, pp. 39-46, 2010

# Results in Author/Conference/Journal

## Authors

Jiawei Han
University of Illinois Urbana Champaign
Publications: 604 | Citations: 15299 | G-Index: 115 | H-Index: 57
Interest: Data Mining, Databases, Artificial Intelligence
107 publication(s) from this author

Mohammed Javeed Zaki
Rensselaer Polytechnic Institute
Publications: 214 | Citations: 3762 | G-Index: 58 | H-Index: 32
Interest: Data Mining, Databases, Distributed & Parallel Computing
74 publication(s) from this author

Alex Alves Freitas
University of Kent
Publications: 149 | Citations: 1322 | G-Index: 32 | H-Index: 20
Interest: Artificial Intelligence, Data Mining, Algorithms & Theory
67 publication(s) from this author

Philip S. Yu
University of Illinois Chicago
Publications: 702 | Citations: 9981 | G-Index: 79 | H-Index: 47
Interest: Data Mining, Databases, Distributed & Parallel Computing
60 publication(s) from this author

Vipin Kumar
University of Minnesota
Publications: 483 | Citations: 10021 | G-Index: 89 | H-Index: 46
Interest: Distributed & Parallel Computing, Data Mining, Artificial Intelli
57 publication(s) from this author

Andrew Kusiak
University of Iowa
Publications: 144 | Citations: 1132 | G-Index: 28 | H-Index: 18
Interest: Engineering, Hardware & Architecture, Artificial Intelligence
56 publication(s) from this author

## Conference

KDD - Knowledge Discovery and Data Mining
Publications: 2,089 | Citation Count: 42,059 | Year Range: 1991-2010
473 publication(s) in this conference

ICDM - IEEE International Conference on Data Mining
Publications: 1,720 | Citation Count: 9,269 | Year Range: 2000-2009
236 publication(s) in this conference

PAKDD - Pacific-Asia Conference on Knowledge Discovery a
Publications: 1,269 | Citation Count: 3,247 | Year Range: 1996-2010
125 publication(s) in this conference

SDM - SIAM International Conference on Data Mining
Publications: 716 | Citation Count: 4,811 | Year Range: 2000-2010
115 publication(s) in this conference

PKDD - Principles of Data Mining and Knowledge Discovery
Publications: 1,057 | Citation Count: 5,213 | Year Range: 1996-2010
100 publication(s) in this conference

ICDE - International Conference on Data Engineering
Publications: 3,523 | Citation Count: 43,592 | Year Range: 1984-2010
77 publication(s) in this conference

CIKM - International Conference on Information and Knowled
Publications: 2,273 | Citation Count: 18,484 | Year Range: 1977-2010
75 publication(s) in this conference

SAC - ACM Symposium on Applied Computing
Publications: 4,281 | Citation Count: 11,277 | Year Range: 1990-2011
69 publication(s) in this conference

## Journals

CORR - Computing Research Repository
Publications: 21,068 | Citation Count: 60,876 | Year Range: 1983-2010
168 publication(s) in this journal

Sigkdd Explorations
Publications: 421 | Citation Count: 4,937 | Year Range: 1996-2010
147 publication(s) in this journal

ESWA - Expert Systems With Applications
Publications: 5,348 | Citation Count: 5,745 | Year Range: 1990-2011
135 publication(s) in this journal

TKDE - IEEE Transactions on Knowledge and Data Engineerir
Publications: 2,412 | Citation Count: 32,170 | Year Range: 1987-2010
117 publication(s) in this journal

DATAMINE - Data Mining and Knowledge Discovery
Publications: 436 | Citation Count: 11,118 | Year Range: 1995-2010
102 publication(s) in this journal

BIOINFORMATICS - Bioinformatics/computer Applications in T
Publications: 6,870 | Citation Count: 92,870 | Year Range: 1985-2010
75 publication(s) in this journal

NAR - Nucleic Acids Research
Publications: 30,771 | Citation Count: 217,608 | Year Range: 1974-2010
59 publication(s) in this journal

KAIS - Knowledge and Information Systems
Publications: 620 | Citation Count: 3,036 | Year Range: 1998-2010
43 publication(s) in this journal

# Which Michael Cohen?

Were you looking for these authors:


**Michael Cohen**
University of Aizu


**Michael B. Cohen**
University of Iowa Hospit...


**Michael F. Cohen**
Microsoft


**Michael R. Cohen**
Institute for Safe Medica...


**Michael X. Cohen**
University of Amsterdam

**Michael Cohen**

**Michael A. Cohen**

**Michael H. Cohen**

**Michael J. Cohen**

**Michael P. Cohen**

**Michael S. Cohen**

**Michael Cohen**


**Michael D. Cohen**
University of Michigan


**Michael M. Cohen**
University of California ...


**Michael S. Cohen**
University of California ...

**Michael Cohen**

**M. Michael Cohen**

**Michael H. Cohen**
Harvard University

**Michael I. Cohen**

**Michael J. Cohen**

**Michael P. Cohen**

**Michael V. Cohen**

**Michael A. Cohen**
Boston University

**Michael E. Cohen**
State University of New Y...

**Michael P. Cohen**
U.S. Department of Transp...


**Michael V. Cohen**
University of South Alaba...

**Michael Cohen**

**Michael A. Cohen**

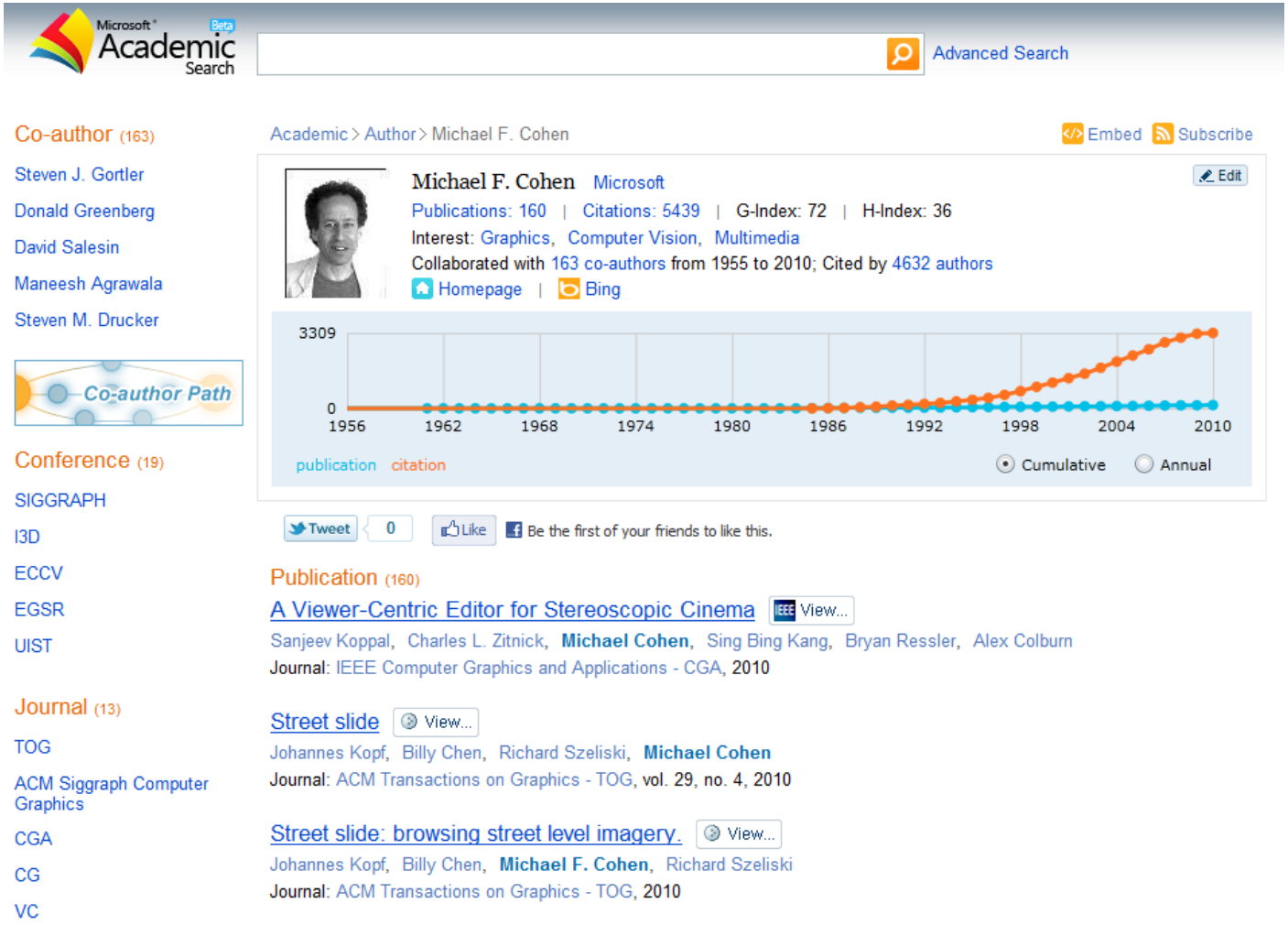
**Michael H. Cohen**

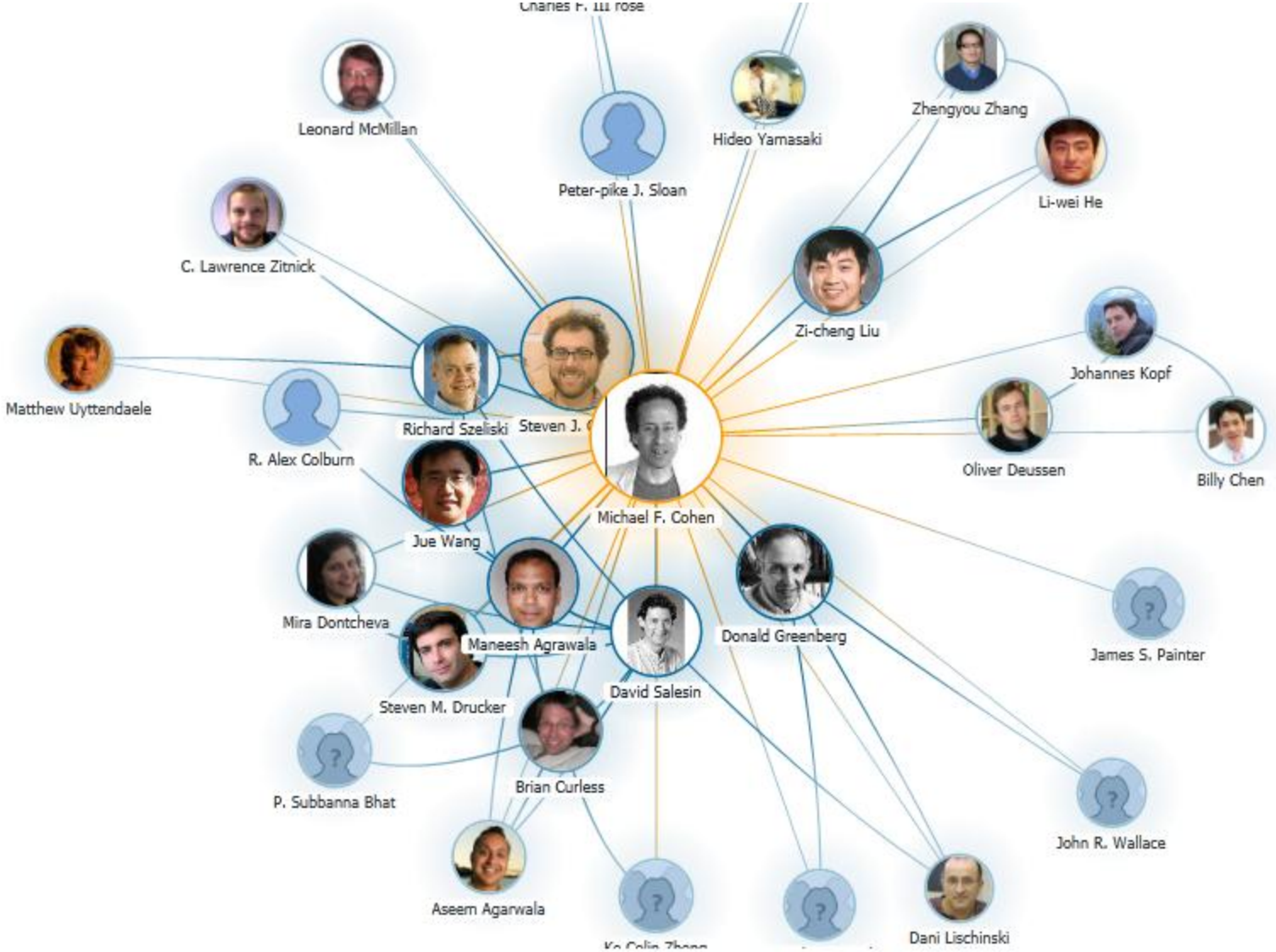
**Michael J. Cohen**

**Michael Lee Cohen**

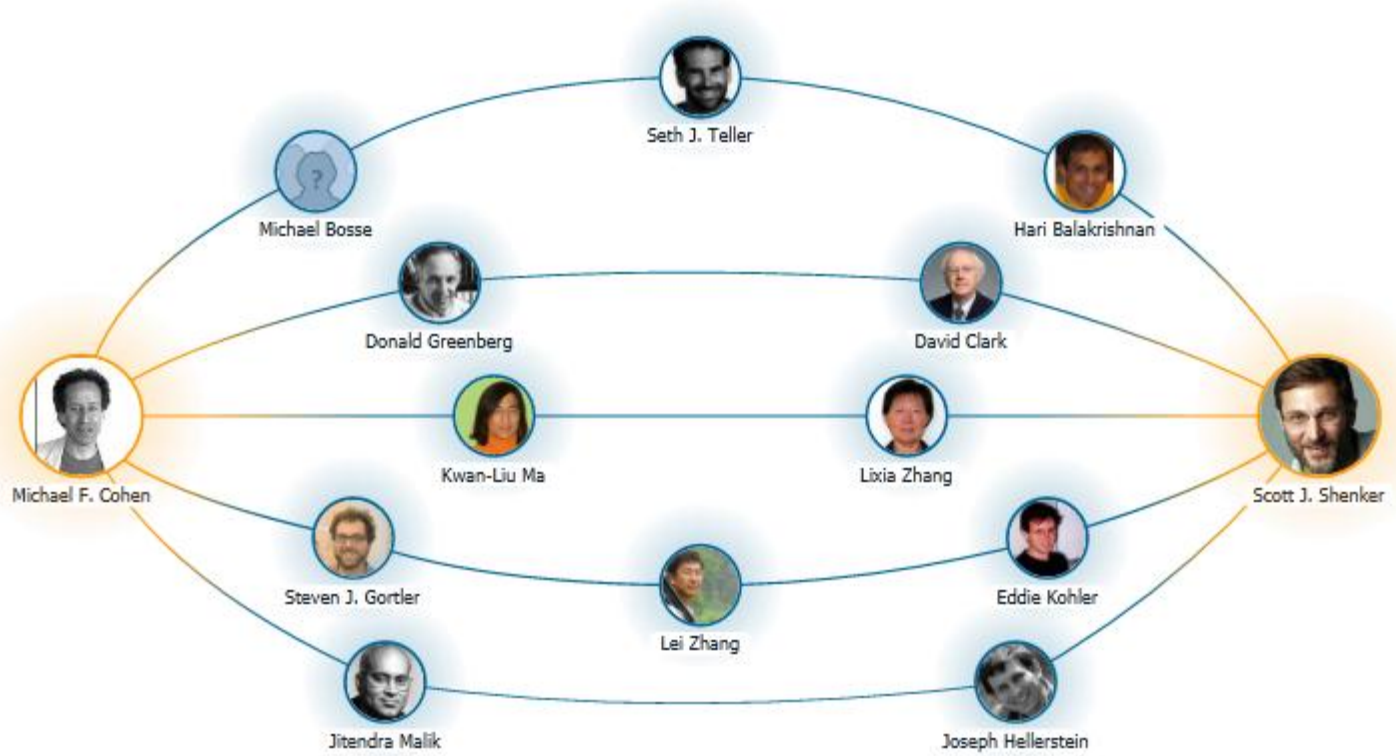
**Michael S Cohen**
University of California ...

# Author Profile




Leonard McMillan
Peter-pike J. Sloan
Hideo Yamasaki
Zhengyou Zhang
Li-wei He
C. Lawrence Zitnick
Zi-cheng Liu
Johannes Kopf
Matthew Uyttendaele
Richard Szeliski
Steven J.
R. Alex Colburn
Oliver Deussen
Billy Chen
Michael F. Cohen
Jue Wang
Mira Dontcheva
Maneesh Agrawala
Donald Greenberg
James S. Painter
David Salesin
Steven M. Drucker
Brian Curless
P. Subbanna Bhat
John R. Wallace
Aseem Agarwala
Dani Lischinski

# Six-degree path



# Organization Profile

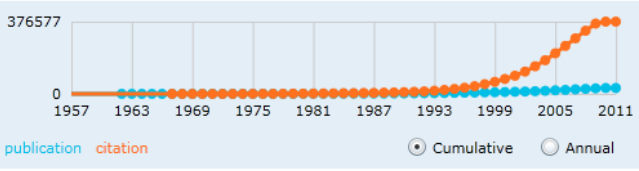
**Microsoft**

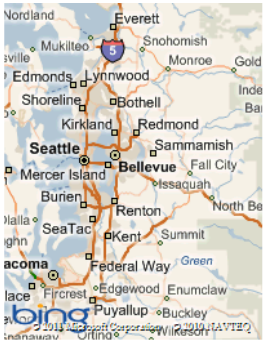
North America

Publications: 41,342 | Citation Count: 679,837 | H-Index: 286

Top Areas: Algorithms & Theory, Networks & Communications, Databases, Engineering, Multimedia

Homepage




Author (1849) All Years

**Rakesh Agrawal**
Microsoft
Publications: 269 | Citations: 20522 | G-Index: 142 | H-Index: 57
Interest: Databases, Data Mining, World Wide Web

**Leslie Lamport**
Microsoft
Publications: 236 | Citations: 16031 | G-Index: 125 | H-Index: 52
Interest: Distributed & Parallel Computing, Programming Languages, Algorithms & Theory

**Charles Antony Richard Hoare (C.A.R. Hoare)**
Microsoft
Publications: 237 | Citations: 12629 | G-Index: 111 | H-Index: 38
Interest: Algorithms & Theory, Software Engineering, Programming Languages

**Anoop Gupta**
Microsoft
Publications: 404 | Citations: 11886 | G-Index: 104 | H-Index: 54
Interest: Hardware & Architecture, Operating Systems, Physics

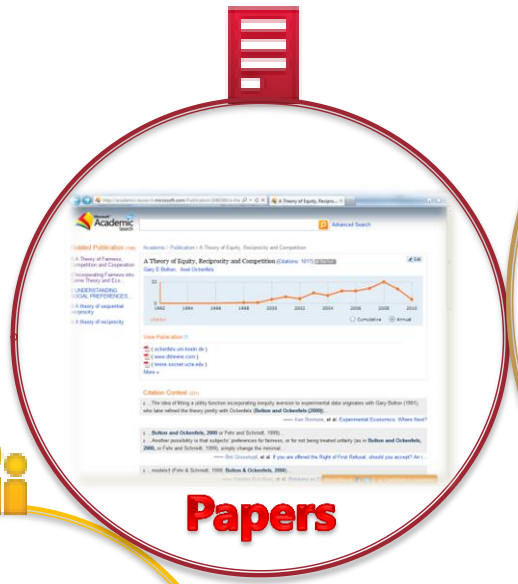
## Computer Science Overall

Filter: Computer Science Overall | Last 5 Years

- Algorithms and Theory
- Artificial Intelligence
- Bioinformatics and Computational
- Computer Education
- Computer Vision
- Data Mining
- Databases
- Distributed and Parallel Computin
- Graphics
- Hardware and Architecture
- Human-Computer Interaction
- Information Retrieval
- Machine Learning & Pattern Reco
- Multimedia
- Natural Language & Speech
- Networks and Communications
- Operating Systems
- Real-Time and Embedded Systen
- Scientific Computing
- Security and Privacy
- Simulation
- Software Engineering & Programn
- World Wide Web
- Computer Science Overall

| Organization (How this rank list is generated) | s | Citations |
|---|---|---|
| Stanford University | | 35130 |
| Microsoft | | 31920 |
| Massachusetts Institute of Technology | | 31836 |
| University of California Berkeley | | 26347 |
| Carnegie Mellon University | | 25241 |
| Harvard University | | 21880 |
| National Institutes of Health | | 19849 |
| University of Illinois Urbana Champaign | | 18383 |
| University of California San Diego | | 15442 |
| University of Wisconsin Madison | | 15040 |
| Cornell University | | 14710 |
| IBM | 4731 | 14630 |
| University of Cambridge | 3485 | 14047 |
| University of California Los Angeles | 3826 | 13985 |
| University of Washington | 3384 | 13858 |
| University of Southern California | 4019 | 13582 |
| University of Michigan | 3735 | 13291 |
| University of Maryland | 3667 | 13012 |
| University of Pennsylvania | 2924 | 12444 |
| University of Toronto | 3140 | 12312 |

# Explore the Relationships







# ChronoZoom

Challenge: The exploration of Big History, with smooth transition from billions of years down to individual nanoseconds.

This is what Walter Alvarez, Professor of Earth and Planetary Science at University of Berkeley set out to do.  And he did it, with the help of Microsoft Research and the Live Labs team.

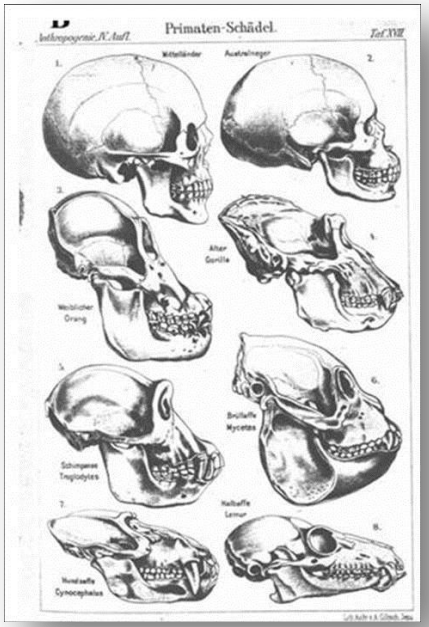

*A new service in development that allows researchers to browse, overlay, and explore interdisciplinary data sources*

# Interactive Opportunities with ChronoZoom

## Exploring Life


Significant Events

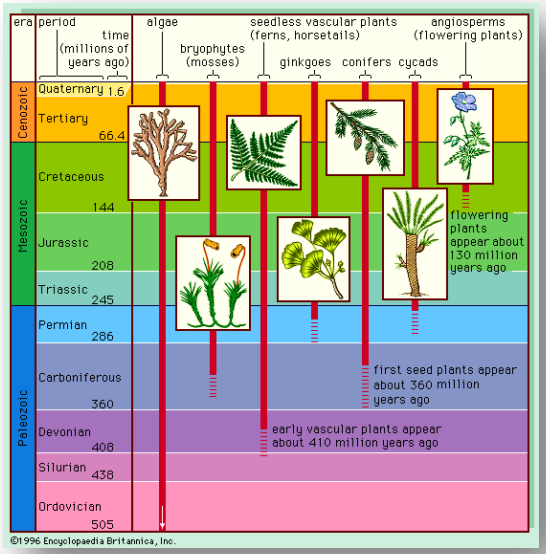
## Scientific Works



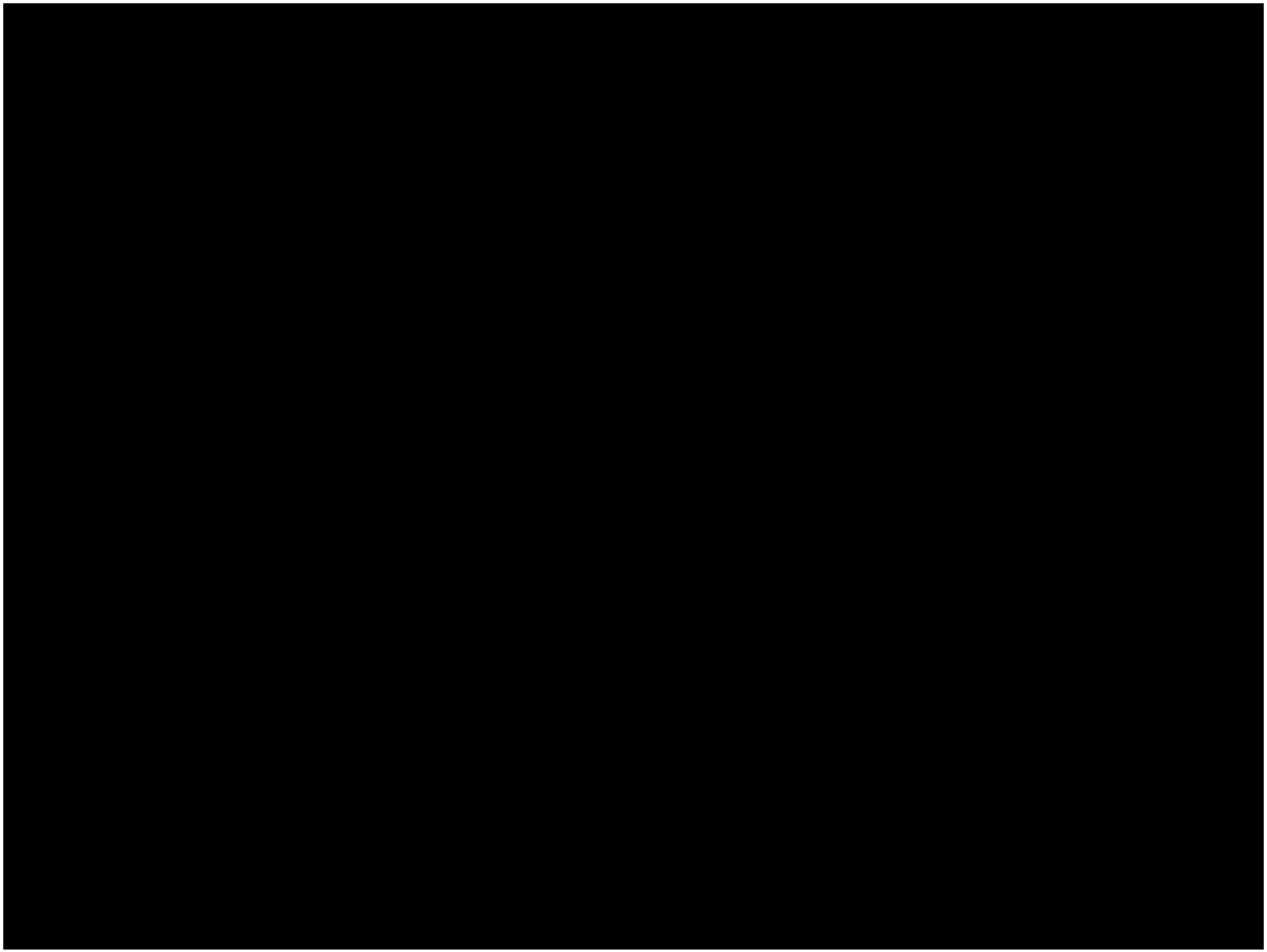

National History

## Cultural Heritage

## Classification and Evolution



# Zoom Technology for Big History



http://www.youtube.com/watch?v=V2q-U02HLb0

# MRC Engagements by Theme

- Computer Science
  - Project Hawaii
  - F#
  - Gadgeteer
- Earth, Energy and Environment
  - WW Telescope
- Natural User Interfaces
  - Kinect research projects
- Education & Scholarly Communication
  - Academic Search
- Health & Wellbeing
  - Microsoft Biology Foundation

# Understanding the genetics of human disease

Collaboration with the
Wellcome Trust Sanger Institute

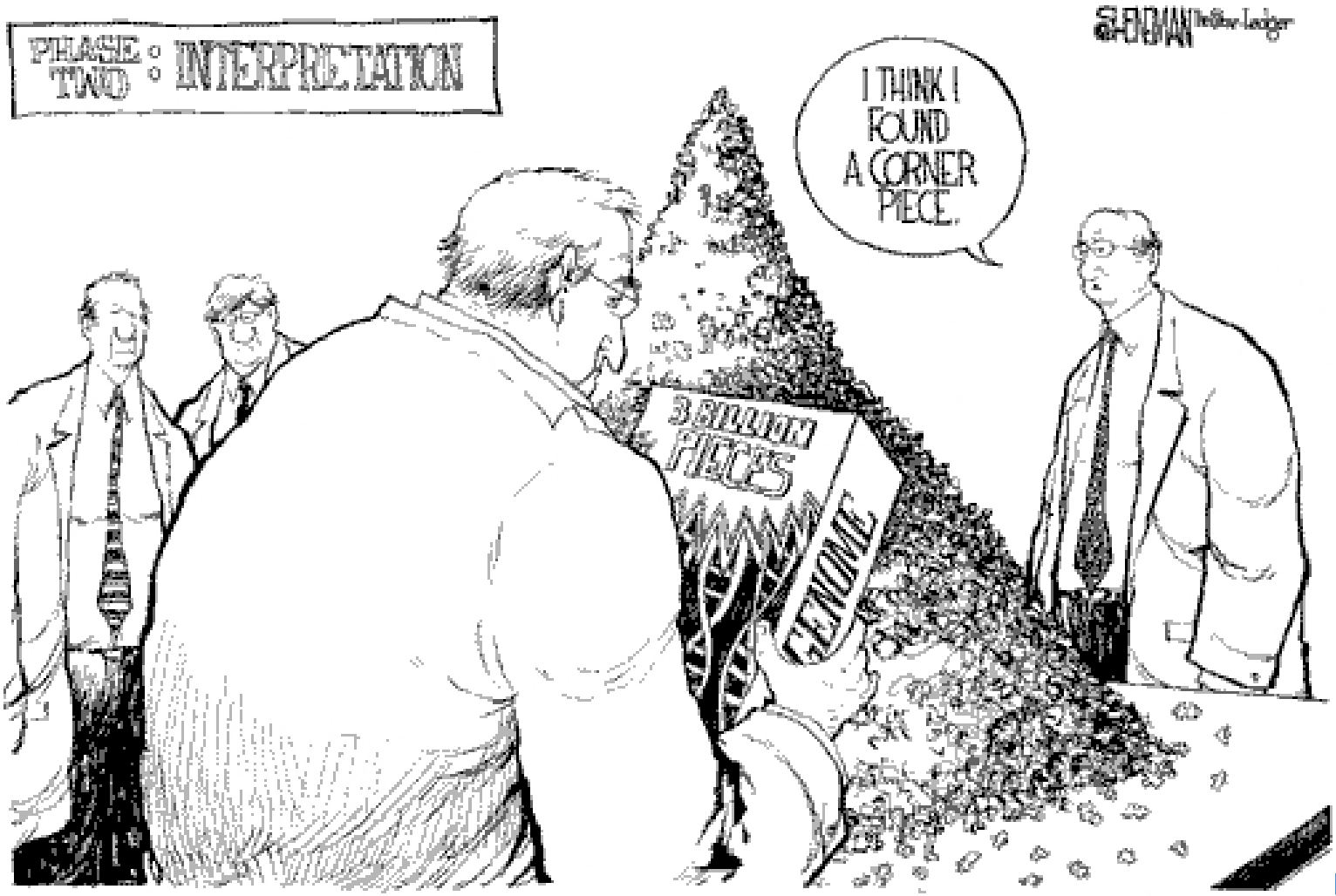
Microsoft Research Cambridge & wellcome trust sanger institute

John Winn,
Wolfgang Lehrach

Richard Durbin,
Manolis Dermitsakis

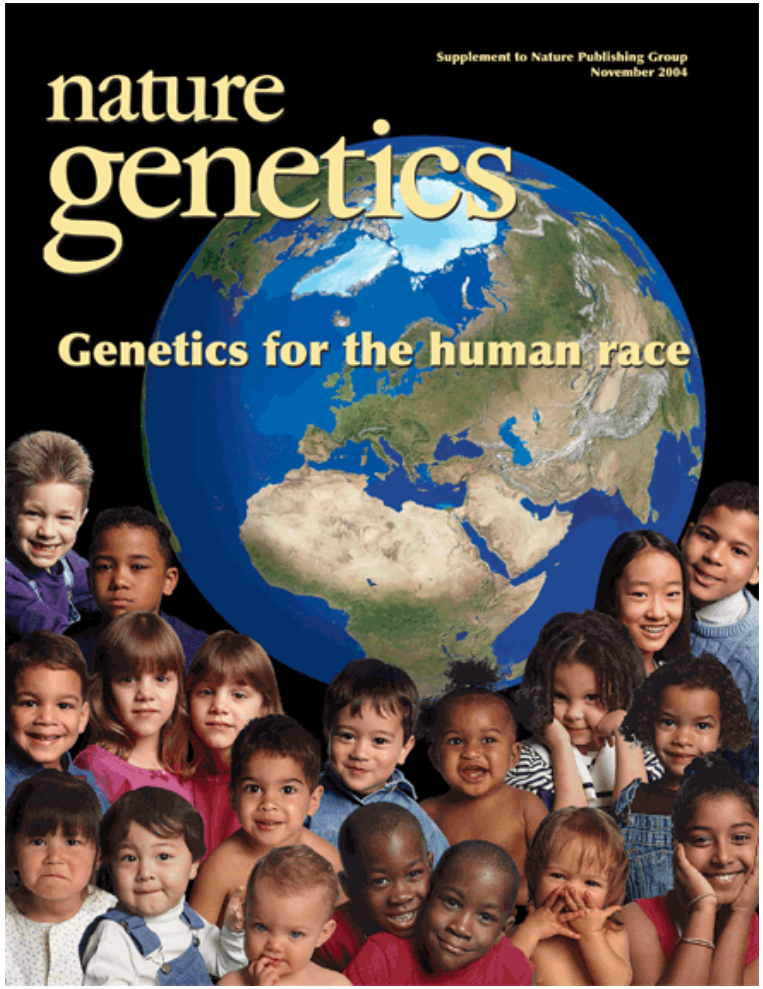
# The Challenge: Gene Function

**Phase 1: Human Genome project**



# Human Genetic Variation

- Accounted for by changes in just 0.1% of the genome (about 30 million bases!)
- Variations may have:
  - Harmless effect *e.g.* change in hair colour
  - No effect
  - Harmful effects
    - Higher risk to certain diseases
    - Genetic disorders
    - Drug toxicity/effectiveness



# Human Diseases and Genes

- **Variations in genetic makeup** defines our susceptibility to diseases: *coronary heart disease, diabetes, arthritis, Crohn's disease, hypertension, bipolar disorder, asthma…*
- Finding genetic causes for cancer, diabetes, heart disease, obesity etc. is very challenging as it depends on **variations in multiple genes**, almost all with **weak effects.**

# Human Diseases and Genes

- Collaborative project to use machine learning tools and large scale data to understand the effect of these variations
  - *Tools*
    - Infer.NET for rapid model development, modification and testing.
    - Parallelisation of tools to allow processing of large scale data sets
  - *Data*
    - A large catalogue of variations between and within populations
    - High throughput measurements of expression of entire genome
    - Disease-labelled genomic data

# What is MBF?

- Microsoft Biology Foundation (MBF) is a bioinformatics toolkit
  - built on top of the .NET Framework 4.0
  - open source under Apache license
  - foundation upon which other tools can be built
- Provides various components useful for biological analysis
  - parsers to read and write common bioinformatics formats
  - support for DNA, RNA and protein sequences
  - algorithm framework for analysis and transformation
  - web connector framework for web-service interaction

# What is MBF intended to do?

- Primarily focused on genomics
  - reusable data structures to represent sequences + symbols
  - I/O framework to load/save sequences
  - algorithm framework to process loaded sequences
- Provides an alternative to other biology frameworks
  - similar concepts to BioJava or BioPerl
  - takes advantage of Microsoft developer tools and .NET
  - will evolve as Microsoft and other contributors add features
- Designed to manipulate large data sets
  - efficient storage of data internally
  - utilizes lazy loading techniques for creation of sequences
  - scalable algorithms that take advantage of multiple cores

# MBF Design Goals

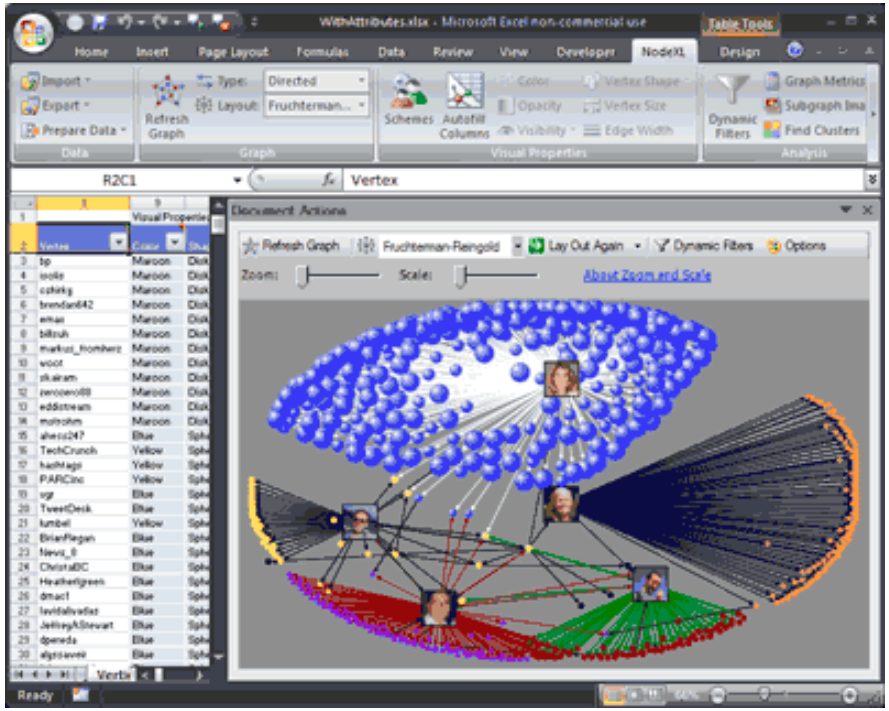
- Extensibility was a primary goal
  - core concepts mapped as interfaces and ABCs
  - can easily provide alternative implementations or add any missing features you need
- Language Neutral
  - built on top of .NET – use any supported language (C++, F#, …)
  - supports dynamic languages such as IronPython
- Designed and implemented using best practices
  - commented source code provided so nothing is a black box
  - algorithms all cite publications
- Interoperability
  - code can be run on several mainstream platforms

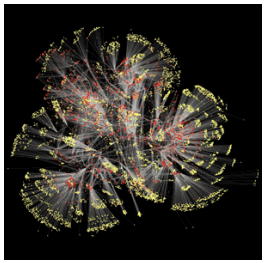
# Thank you

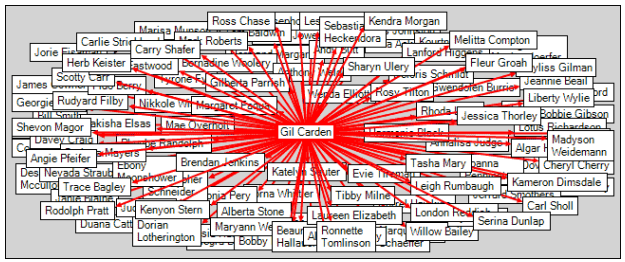
Microsoft®
Research Connections

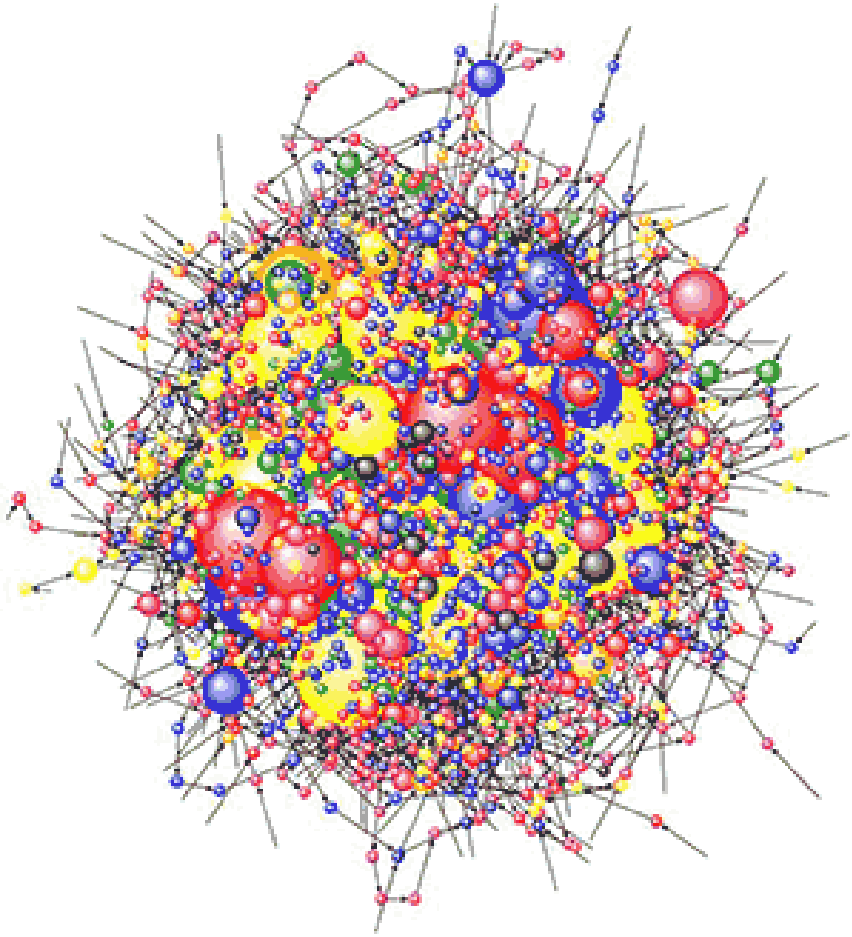
# NodeXL

## Network graph visualization







Binary and source code:
http://nodexl.codeplex.com

20110605-NodeXL-Twitter-#OR11 - Microsoft Excel

Table Tools

Home | Insert | Page Layout | Formulas | Data | Review | View | Developer | NodeXL | Design

Import | Export | Prepare Data — Data

Refresh Graph | Type: Directed | Layout: Harel-Koren ... | Automate — Graph

Autofill Columns | Color | Opacity | Visibility | Vertex Shape | Vertex Size | Edge Width — Visual Properties

Dynamic Filters | Graph Metrics | Subgraph Images | Groups — Analysis

Import | Export | Use Current for New — Options

Workbook Columns | Graph Elements — Show/Hide

Home Page | Sample Workbook | Discussions | Check for Updates | Registration | About Microsoft NodeXL — Help

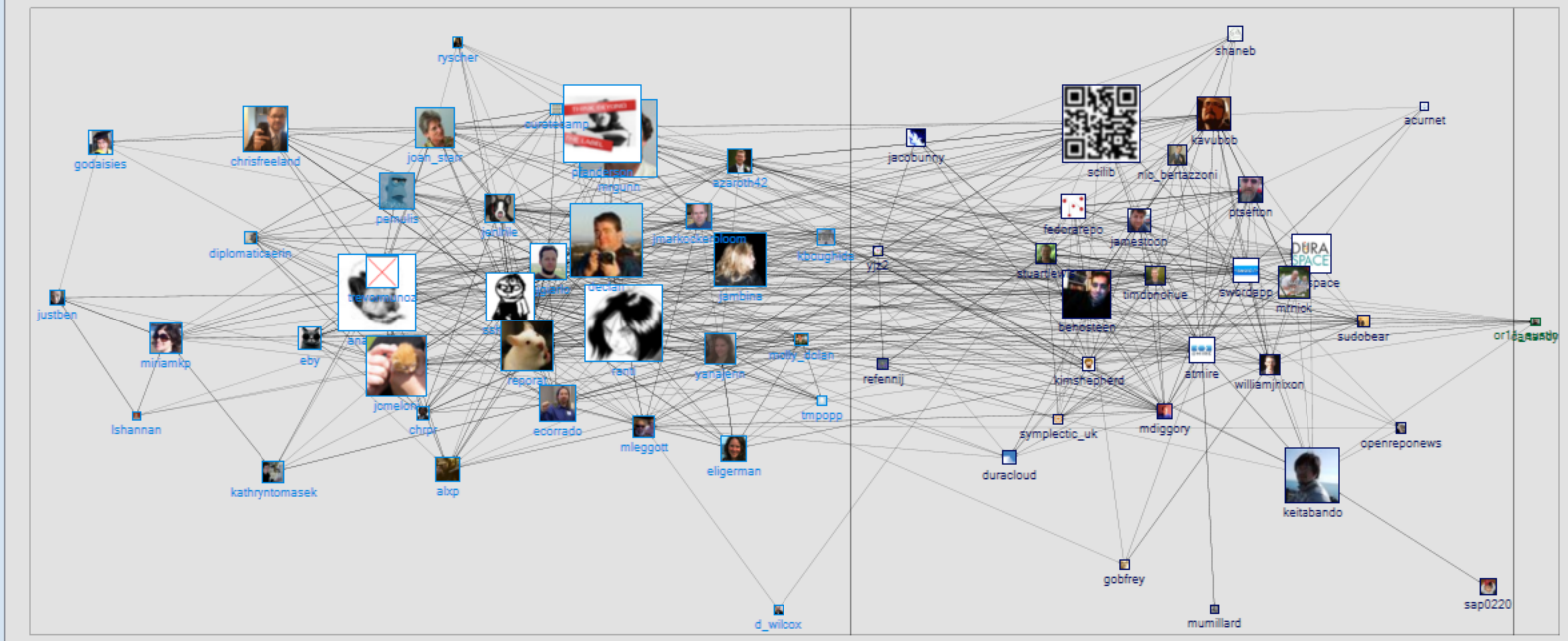
A2 | Vertex

| | A | B | C | D | E | F | G | H | I | AC | AD | AE | AF | AG | AH | AI | AJ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | | | Graph Metrics | | | | | | | | | | | | | | |
| 2 | Vertex | Subgraph | In-Degree | Out-Degree | Betweenness Centrality | Closeness Centrality | Eigenvector Centrality | PageRank | Clustering Coefficient | Followed | Followers | Tweets | Favorites | Description | Web | Time Zone | Time Zone UT Offset (Secon |
| 3 | atmire | | | | | | | | | | | | | | | | |
| 4 | sshreeves | | | | | | | | | | | | | | | | |
| 5 | reporat | | | | | | | | | | | | | | | | |
| 6 | anarchivist | | | | | | | | | | | | | | | | |
| 7 | stuartlewis | | | | | | | | | | | | | | | | |
| 8 | kimshepherd | | | | | | | | | | | | | | | | |
| 9 | fedorarepo | | | | | | | | | | | | | | | | |
| 10 | swordapp | | | | | | | | | | | | | | | | |
| 11 | keitabando | | | | | | | | | | | | | | | | |
| 12 | duraspace | | | | | | | | | | | | | | | | |
| 13 | jcmeloni | | | | | | | | | | | | | | | | |

Document Actions

Refresh Graph | Harel-Koren Fast Mul | Lay Out Again | Dynamic Filters | Graph Options

Zoom: | Scale:




Edges | Vertices | Groups | Group Vertices | Overall Metrics

Ready | Count: 45 | 100%