## The Initiative in Innovative Computing @ Harvard

#### ★ Visiting NSF Today (July 18, 2007)

- Alyssa A. **Goodman**, Director of IIC & Prof. of Astronomy (speaker)
- Felice Frankel, Senior Research Fellow/IIC Visual Computing Group
- Josh **Grindlay**, Prof. of Astronomy, PI of DASCH project
- Michael Halle, Senior Scientist @ IIC/PM of Astronomical Medicine
- Ros **Reid**, Editor, American Scientist & Visitor to IIC
- Rita Tavilla, Dir. of Administration & Operations @ IIC
- Louie Weitzman, Head of Internet Technologies Group @ IIC



## Why me?

COMPLETE Home Page

#### Goodman et al. 2006



#### COMPLETE Postdoc, 2007

#### Referencing Data from the COMPLETE Survey

COMPLETE data are non-proprietary. Please reference Ridge, N.A. et al., "The COMPLETE Survey of Star Forming Regions: Phase 1 Data". 2006. Al. 131. 2921 as

#### Goals for this Hour

Introductions (3 min) What are we doing at IIC? (2 min) Where did IIC come from? (2 min) What is IIC for? ... The "GAP" (5 min) How is IIC organized? (Branches x Projects, 5 min) Example efforts:

Astronomical Medicine (6 min) The Connectome (3 min) Envisioning Science (5 min) iSTEM (2 min) The Space-Time Machine (5 min) Going "Off-the-Desktop" (4 min) Q&A (20 min)



What does the wiring diagram of the human brain look like?

How do stars form in our Galaxy today?

How do humans interact with "big data"?

#### What's "Dark Energy"?

What are the earliest structures in the Universe? How should we model blood flow in the human body from atomic to physiological scales?

How should we use technology in teaching?

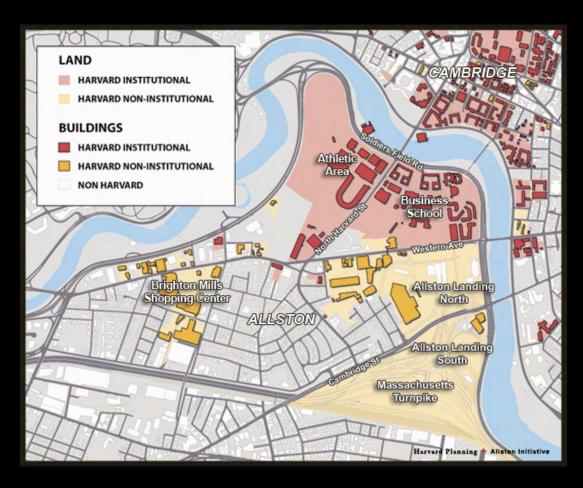
Can we predict how often a genotype will produce disease? How can collaborative technologies accelerate Stem Cell & Neuroinformatics Research?

How should we mine routine medical records, on vast scales, to cure diseases?

## How did IIC begin?

#### Response to Harvard's "expansion" in Science, and into Allston

See IIC <u>Whitepaper</u> (2004) & Task Force on Science & Technology <u>report</u> (2005) at <u>iic.harvard.edu</u> for more







#### **SCIENTIFIC DISCIPLINES**



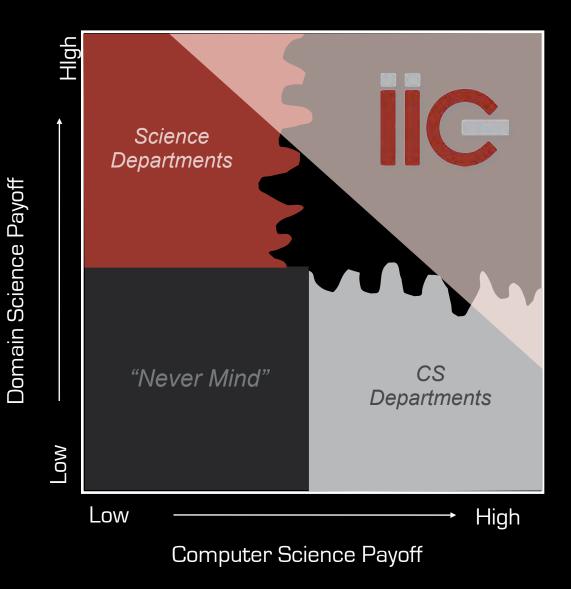
#### **COMPUTER SCIENCE**

Increasingly, core problems in science require computational solution

Typically hire/"home grow" computationalists, but often lack the expertise or funding to go beyond the immediate pressing need Academic researchers often focused on finding elegant solutions to basic computer science challenges

Often see specific, "applied" problems as outside their interests

# IIC's Ideal "Space"



## Original IIC "Branches" ( and Projects Draw upon >1 )

Visualization (Visual Computing)	Distributed Computing (Internet Technologies)	Databases/ Provenance (+HPC/Data)	Analysis & Simulations (+HPC/processing)	Instrumentation (Including "off-the desktop" computing)
Physically meaningful combination of diverse data types.	e-Science aspects of large collaborations. Sharing of data and computational resources and tools in real-time.	Management, and rapid retrieval, of data. "Research reproducibility" where did the data come from? How?	Development of efficient algorithms. Cross-disciplinary comparative tools (e.g. statistical).	Improved data acquisition. Novel hardware approaches (e.g. GPUs, sensors).
V	DC	DB/P	AS	1



Plus...Educational Programs that bring IIC Science to Harvard students, and to the public at large.

## Building the Best (Startup) Program

Visualization	Distributed Computing	Databases/ Provenance	Analysis & Simulations	Instrumentation
Proje P	ect 1 roject 2		Alisteringerstationers	
		Project 3		
V	DC	DB/P	AS	1



#### Responses to March 2006 IIC Call for Ideas

Atomistic Modeling of Biomolecular Function	V	DC	DB/P	AS	
Multiscale Hemodynamics	V	DC	DB/P	AS	
Gene Pattern + The Virtual Data Center	V	DC	DB/P	AS	
Medical Treatment Outcomes Online		DC	DB/P		1
Enhanced Viz/Analysis Tools for Archaeo/Geo/Seismology	V	DC	DB/P	AS	
Spatial Ontology Mapping (Community-based)	V	DC			
Knowledge Ecology of Science (Peer-to-Peer Collaboration Networks)	V	DC	DB/P		
Framework for Multimodal Studies in Genetics, Biology & the Mind	V	DC	DB/P	AS	1
Connectional Analysis of Synaptic Circuitry in the Mammalian Nervous System (The "Connectome")	V	DC	DB/P	AS	I
LHC/LSST/MWA Consortium for Data-Intensive Science	V	DC	DB/P	AS	1
A Portal for the National Virtual Observatory	V	DC			
Time-Series Research Collaborative	V	DC	DB/P	AS	

## Education is central to IIC's mission

At Harvard:

- Undergraduate & graduate courses focused on "data-intensive science"
- New graduate certificate program, within existing Ph.D. programs
- Research opportunities at undergraduate, graduate, and postdoctoral levels
- Envisioning Science/Picturing to Learn
- **IIC Seminar Series**
- **Faculty** Hiring
- Beyond Harvard:
  - Envisioning Science/Image & Meaning Program
  - New museum, highlighting the kind of science done at the IIC



What does the wiring diagram of the human brain look like?

How do stars form in our Galaxy today?

How do humans interact with "big data"?

#### What's "Dark Energy"?

What are the earliest structures in the Universe?

How should we model blood flow in the human body from atomic to physiological scales?

How should we use technology in teaching?

Can we predict how often a genotype will produce disease? How can collaborative technologies accelerate Stem Cell & Neuroinformatics Research?

How should we mine routine medical records, on vast scales, to cure diseases?

## The AstroMed Story

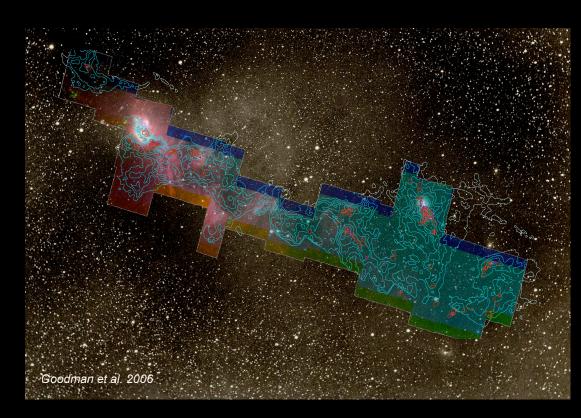


Computer Scientist

Computer Scientist

*"Viz has failed the scientific community…"* 







#### mage size Set 2.4 PLETE = COordinated Molecular Tiew size Set 2.4 Line Exinction Thermal Emission

#### mm peak (Enoch et al. 2006)

sub-mm peak (Hatchell et al. 2005, Kirk et al. 2006)

<sup>13</sup>CO (Ridge et al. 2006)

mid-IR IRAC composite from c2d data (Foster, Laakso, Ridge, et al. in prep.)

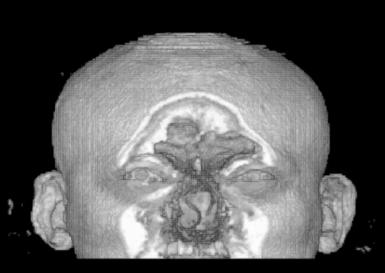
Optical image (Barnard 1927)

n: 163/249 oom: 227% Angle: 0

#### "Slices"

#### "KEITH"

#### "PERSEUS"



lmage slze: 520 x 274 View slze: 1305 x 733 WL: 63 WW: 127	A	thiteenCO_249.tif thiteenCO_249.tif thiteenCO_249.tif thiteenCO_249.tif
		thirteenCU_249.tit 0
		127
		64 L
		•
m: 21/249 700m: 227% Angle: 0	(	11:04:53 AM 7/26/05 Made in OsinW

## Real 3D space



Head "x"



3D rendering: <u>GE Healthcare</u>

#### "Position-Position-Velocity" Space

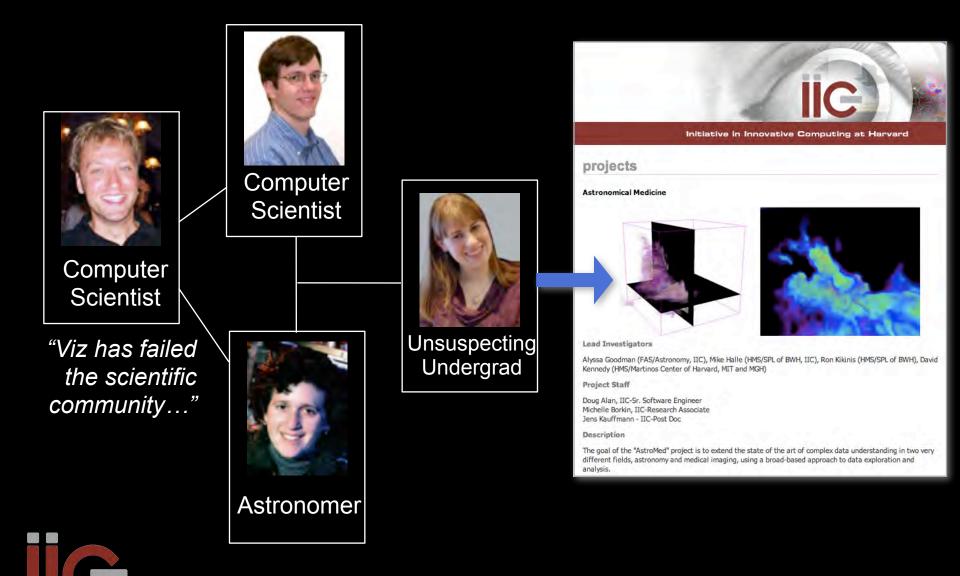


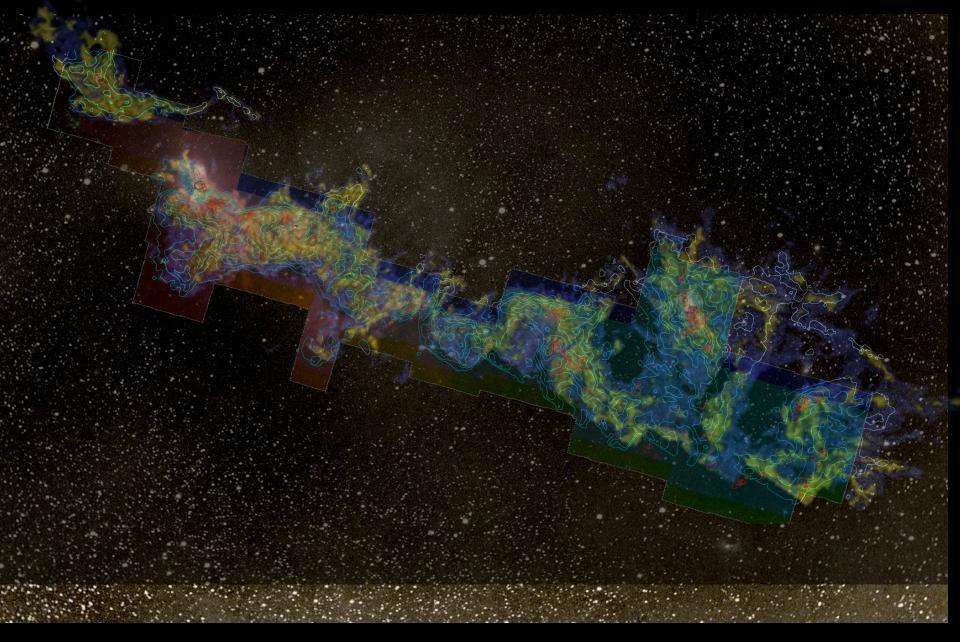
Sky "x" (Right Ascension)



3D rendering: AstroMed / N. Holliman (U. Durham), using VolView (ITK-based)

## The AstroMed Story







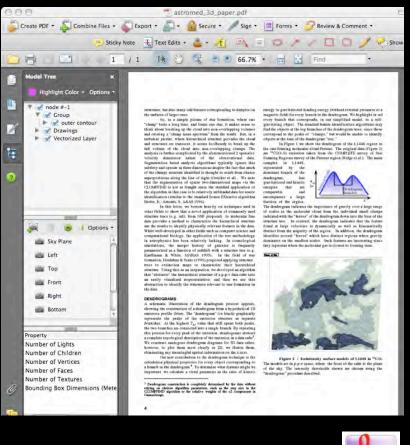


#### **Generalizing & Sharing**

# Open-Source code released, and explained, as it is developed.



# Changing the future of scientific publishing.





#### The Connectome:

#### Wiring Diagram for a Complete Brain Circuit

(Connectional Analysis of Synaptic Circuitry in the Mammalian Nervous System)

3D images from electron-microsope images of serial sections (slices)

- Large volumes studies: up to 500 mm cubes
- High resolution: ~5nm x-y; 50 nm in z ( $10^5 \times 10^5 \times 10^4 = 10^{14} \text{ voxels}$ )
- Large datasets: 10-100 TB

Potentially intractable computationally w/o a hierarchical approach

Start with the large, dominant pathways:

<u>The biggest wires and the biggest excitatory connections.</u>

Use this as scaffolding to then solve other pathways: inhibition, lateral connections, feedback.

"PetaScale Computing" most certainly...



40,000x40,000 pixels 1.6 GB 120x120 mm (3 nm/ pixel)

Here shown 40x undersampled or 1/1600<sup>th</sup> of the data in a single section. The goal, 10,000 such sections making up a 3D volume

ım

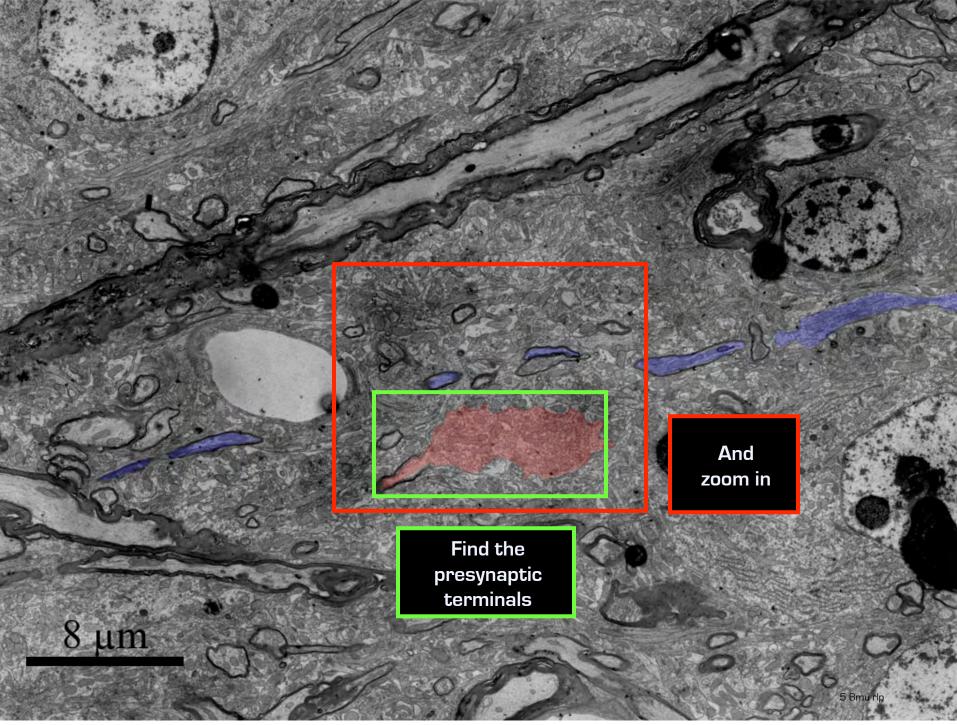
Reconstruct the big axons first: the "wires" (Davi Bock and Clay Reid)

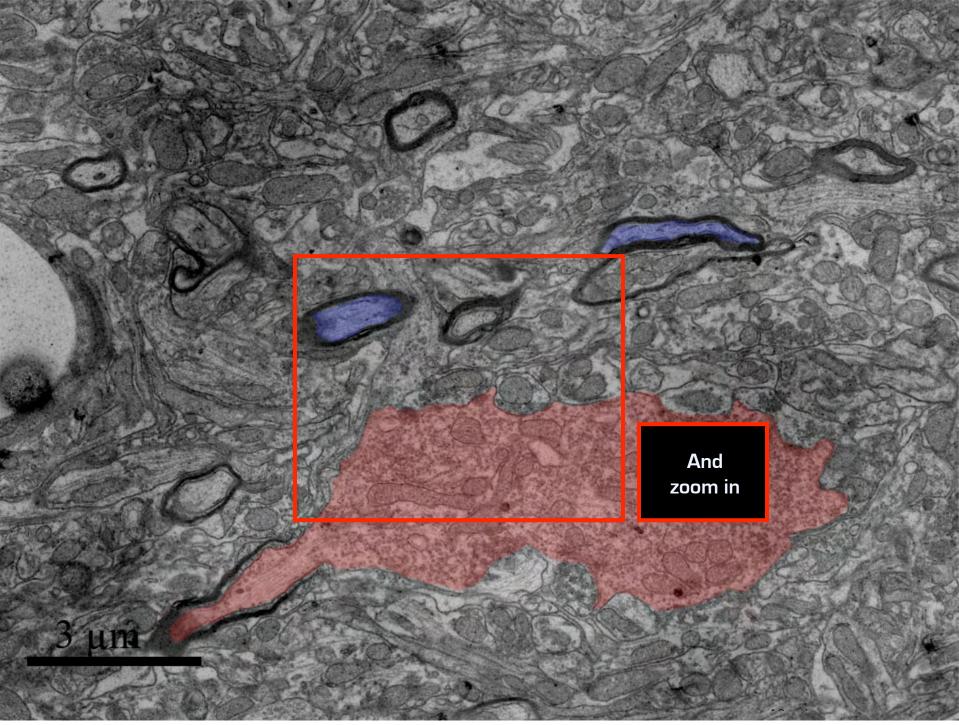
Then find the big synapses: the connections

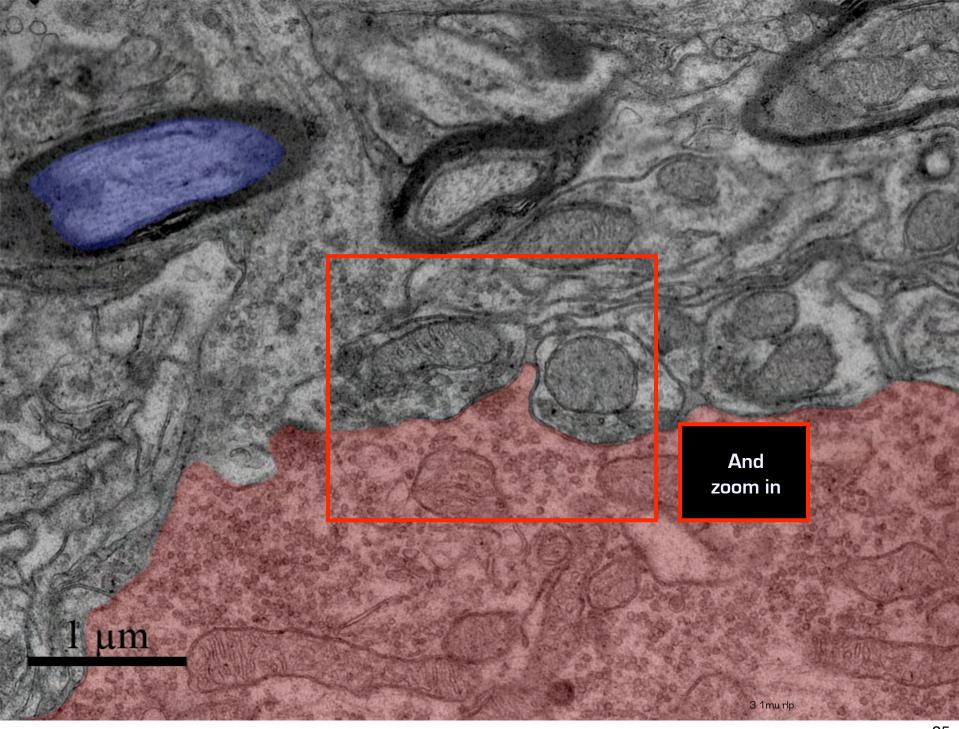
6 15mu EM big view

Then

zoom in



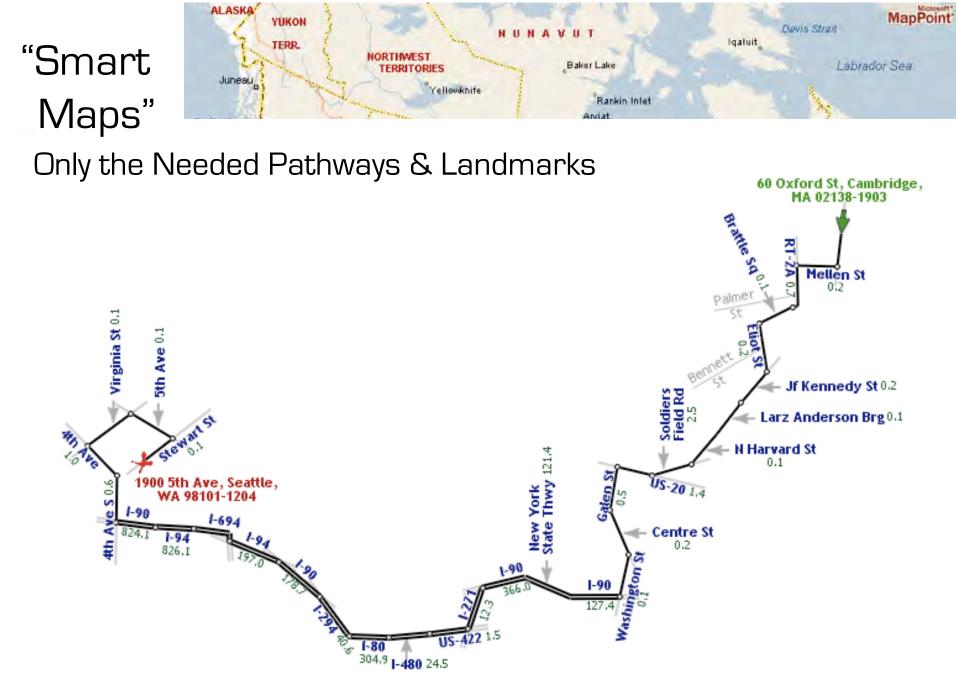




Zoomed in ~40x, or 1/1600<sup>th</sup> of the total field of view. Displayed at native resolution.

> Until you identify a synapse (one connection out of many thousand in the circuit)





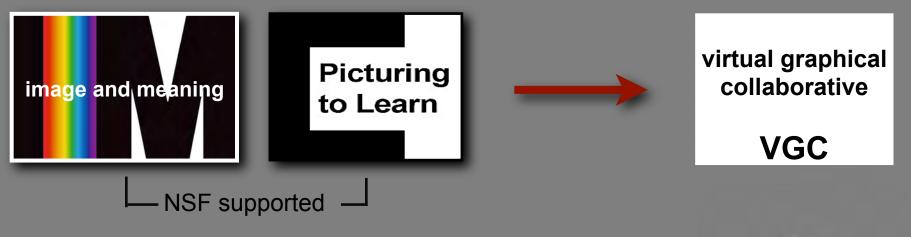
LineDrive (Microsoft & Maneesh Agrawala)

## Scalability

10 <sup>6</sup> pixels 10 <sup>7</sup> voxels 10 <sup>8</sup> voxels	this projector an MRI of your brain, at 0.5 mm resolution Perseus COMPLETE data cube
10 <sup>14</sup> voxels	the <b>Connectome</b> , 0.5 mm <sup>3</sup> of brain tissue
10 <sup>16</sup> pixels	Google Earth Imagery at 1 foot resolution
10 <sup>19</sup> voxels	Google <mark>Earth 3D</mark> , ±1000 feet of elevation, 1ft. res.
10 <mark>22</mark> voxels	the Connectome, full human brain

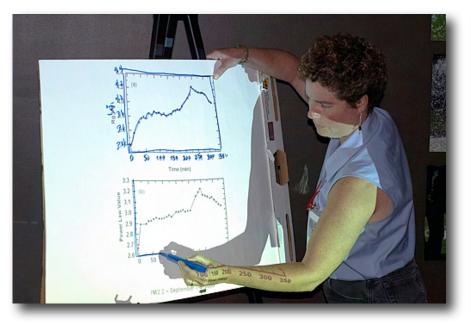


# envisioning science



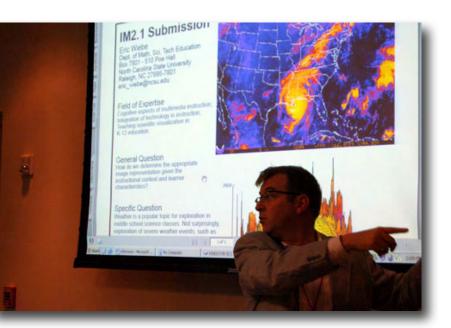


visual computing









# face-to-face -----> virtual



visual computing

# virtual graphical collaborative project

synchronous and/or asynchronous, ENGAGED, small, interdisciplinary, groups

- researchers (via Image and Meaning program)
- students (via Picturing to Learn program)

to create and archive unique "libraries" showing: the process = the evolution of conversations commonalities = tagging beyond the nouns of disciplines



visual computing

# Μ

Title: The Milky Way Initiator: Ross B Initiated: May 24th, 2007

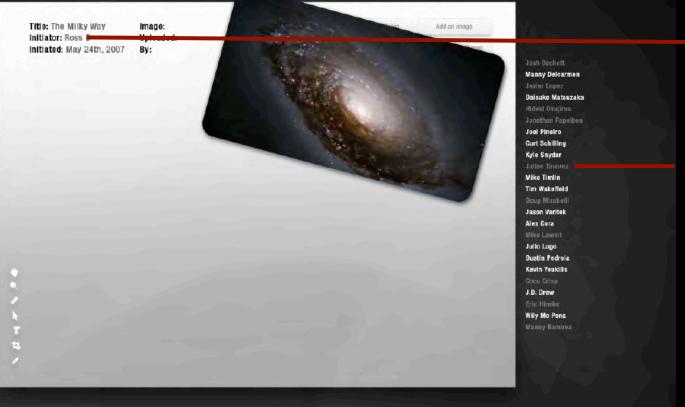
\*\*\*\*\*

By:

Image: Add an image Uploaded:

Manny Delcarmen Daisuke Matsuzaka Joel Pineiro Curt Schilling Kyle Snyder Mike Timlin Tim Wakefield Jason Varitek Alex Cora Julio Lugo **Dustin Pedroia** Kevin Youkills J.D. Drew Wily Mo Pena

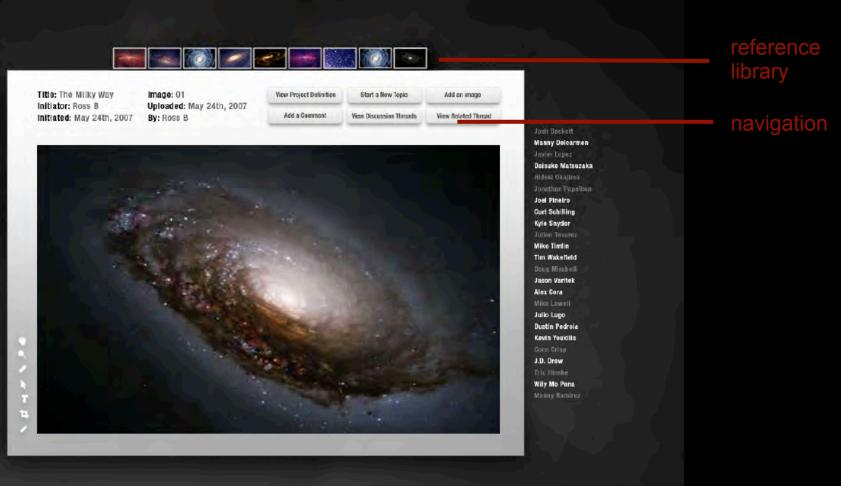
# M



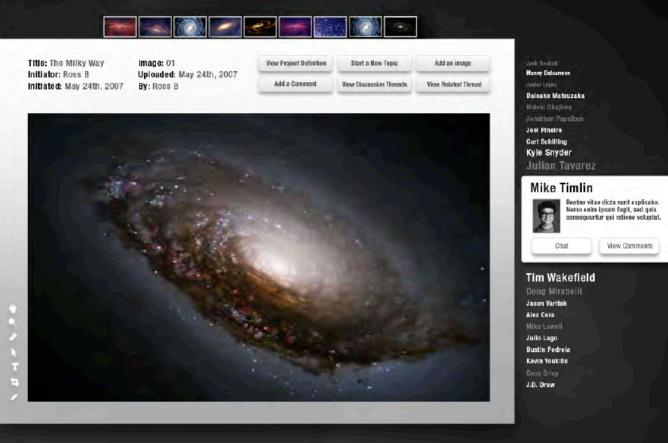
#### "ross" uploads fly-in image

workspace collaborators, active and others











\*

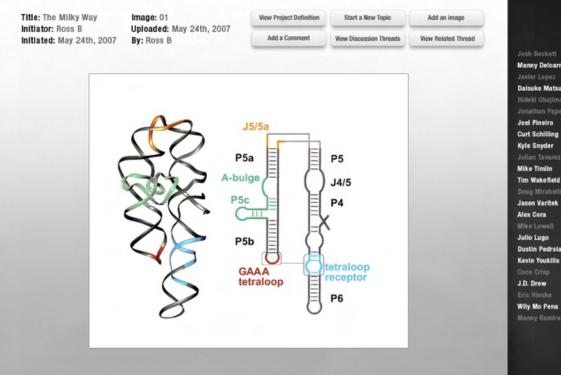
k

T

4

1





Manny Delcarmen Dalsuke Matsuzaka **Joel Pineiro Curt Schilling** Kyle Snyder Mike Timlin **Tim Wakefield** Jason Varitek Julio Lugo Dustin Pedrola Kevin Youkilis



۷

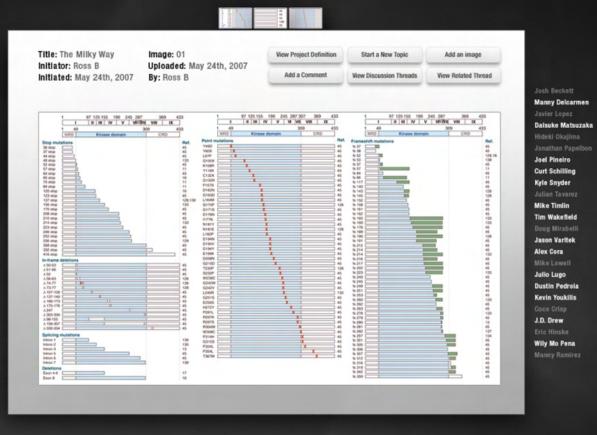
•

k

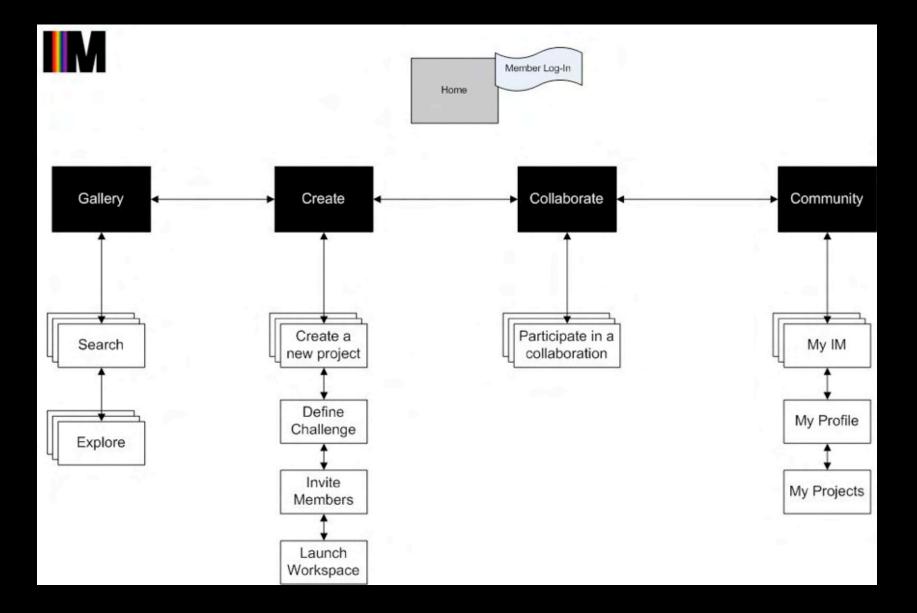
Т

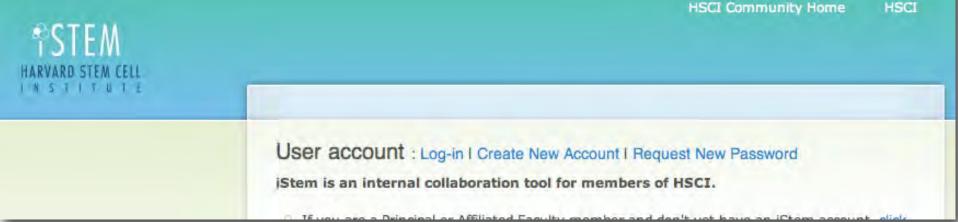
ŧ,

1



38





iSTEM: A "Collaboratory" for the Stem Cell Research Community at Harvard (An IIC-HSCI Joint Project)

### Goals

-communicate digitally in multiple ways (text, speech, graphics, video) -share, annotate, archive and retrieve digital content -share and search joint resources (reagents, protocols publications and

computational tools)

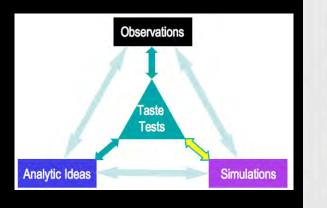
### **Features**

-web-based "social technology" framework

-based on reusable open-source components

-intended to be generalizable at Harvard, and on an international scale

"Taste Tests": Simulations, Observations, and Delicious Insight



Star Formation Taste Tests > Overview ŵ 1-C A A + 🚱 https://iic.grouphub.com/projects/700257/pr 😋 ^ 📿-Google Calendar Wikis Etc. \* Directories \* Delicious \* RSS (1423) \* ADS Best BevondADS \* Image Search \* Weather \* Dashboard | Choose a project Settings | Permissions | My account | Log-out Star Formation Taste Tests CFA Messages To-Do Milestones Writeboards Files Search Overview Chat Time People lic **Project overview** Add message | Add to-do list | Add milestone | Add file

Welcome to the Tasting Room



This is the collaborative space for those who do simulations of star forming regions, and those who observe them. It was inspired, in the Fall of 2006, by the NSF proposal entitled "Star Formation Taste Tests," by A. Goodman & E. Rosolowsky. Today, it is used to host conversations about and short descriptions of simulatons, along with links to longer descriptions (e.g. Journal articles & web sites). In the future, we are planning to connect more enhanced descriptions of those simulations directly to online code bases and sample outputs (likely with help from our friends at NCSA). So, stay tuned.

### What's fresh?

Today

No activity today

### Yesterday

General Remarks and Suggestions by Jürgen 5. in Full Group Communications, 23 Nov

### **Before Yesterday**

Go to "https://iic.grouphub.com/P4080344"

CoodmanNSF1106Basecamp.pdf
(PDF, 3151K) uploaded by Alyssa G., 17 Nov

Subscribe to project's RSS feed

Want to be notified any time someone posts a message, comment or file, or adds or completes a to-do item or milestone in this project? Just subscribe to your project RSS feed RSS (What is RSS?)

People on this project

### IIC

Alyssa Goodman

Douglas Alan Last login about 22 hours ago

Michelle Borkin Last login 2 days ago

Jens Kauffmann Last login 18 days ago

Tim Clark Last login 105 days ago

### American Museum of Natural History

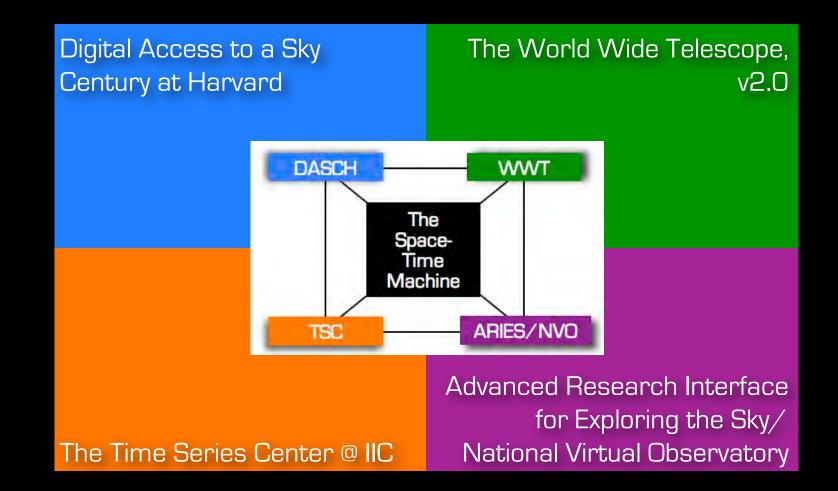
Mordecai-Mark Mac Low Last login 28 days ago

**Cardiff University** 

Anthony Whitworth No login yet

CfA

## The Space-Time Machine



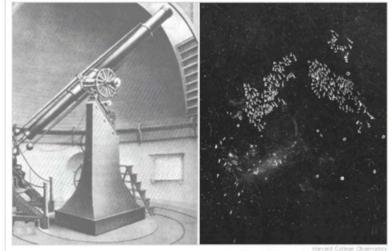


The Space-Time Machine: DASCH



Ehe Ne Get Hom		ork Eimes Very		Science					Science All NYT Search					
WORLD	U.S.	N.Y. / REGIO	BUSINESS	TECHNOLOGY	SCIENCE	HEALTH	SPORTS	OPINION	ARTS	STYLE	TRAVEL	JOBS	REAL ESTATE	AUTOS
					ENVIR	RONMENT	SPACE & C	OSMOS						

### A Trip Back in Time and Space



The Great Refractor, left, which captured the first picture of a star in 1850, and an image of the Large Magellanic Cloud taken in 1900. More Photos >

By GEORGE JOHNSON

In the summer of 1889, when this was still an analog world, a young A PRINT astronomer named Solon I. Bailey carefully packed two crates of glass REPRINTS photographic plates taken at his outpost in the Peruvian Andes for CA SAVE shipment to Harvard College Observatory. Carried down the + SHARE mountain on muleback and across a suspension bridge to the village of Chosica, the fragile load was put on a train bound for Lima and the long voyage to Boston Harbor.





Harvard's Cosmos



For nearly 18 months the data stream

continued - more than 2,500 plates from what Mr. Bailey had quaintly named Mount Harvard - followed in the coming years by tens of thousands more from a second Peruvian station in Arequipa. Over the decades more streams came from Chile, South Africa and New Zealand, joining the growing piles produced by telescopes in Massachusetts.

The accumulated result weighs heavily on its keepers on Observatory Hill, just up Garden Street from Harvard Square: more than half a million images constituting humanity's only record of a century's worth of sky.

Besides being 25 percent of the world's total of

#### Next Article in Science (2 of 9) »







How, and where, did Bill propose to Hillary?

#### Also in Travel:

- Which 17th-century watering hole was George H.W. Bush spotted in?
- Mhich four presidents were born in Boston?

#### MOST POPULAR

E E-MAIL

ARTICLE TOOL

Ioshua

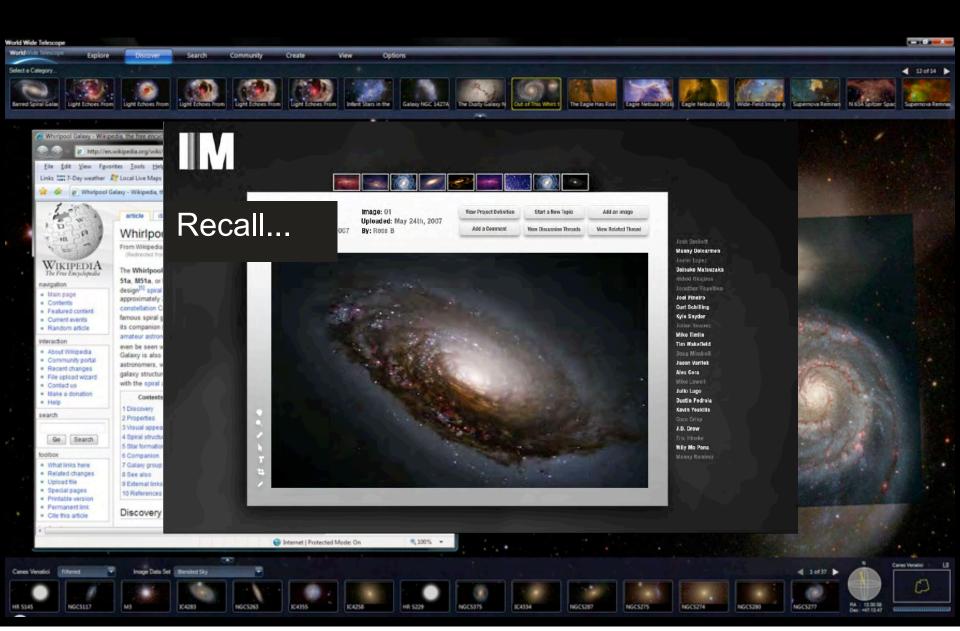
- E-MAILED BLOGGED SEARCHED
- 1. Increasing Rate of Foreclosures Upsets Atlanta
- 2. A Hipper Crowd of Shushers
- 3. New York Wineries Face Tastings Gone Wild
- 4. The Gregarious Brain
- 5. The Park Slope Parent Trap
- 6. Editorial: The Road Home
- 7. Portugal's Hidden 'Dream Places'
- 8. 🛜 Paul Krugman: Health Care Terror
- 9. Op-Ed Contributors: Five Ways to End Harry Potter
- 10. Op-Ed Contributor: Mr. Cheney's Minority Report

Go to Complete List »



Watch David Poque put the iPhone to the test

## The Space-Time Machine: WWT 2.0



# The Space-Time Machine: A Public-Private Partnership

Digital Access to a Sky Century at Harvard: Private Donor + NSF The World Wide Telescope, v2.0: Microsoft Research + Private Foundation (+ NSF?)



The Time Series Center @ IIC: NSF + Microsoft Technical Computing Initiative Advanced Research Interface for Exploring the Sky/National Virtual Observatory [NSF?!]





# Going "Off-the-Desktop"

ĪC

Going Off-the-Desktop in Research & Education: The Scientists' Discovery Room





## The Scientists' Discovery Room: Version 0.01



Initiative in Innovative Computing (IIC)

### A P 🖄 🖉 C A A + Ohttp://lic.harvard.edu/index.html

🖸 🗖 🔍 🖌 Google

### iic.harvard.edu

home about lic projects education people seminars



### Initiative in Innovative Computing at Harvard

employment links of interest reaching the iic

#### iic login

enter email here



Astronomical Medicine

### Astronomical Medicine Movie

The goal of the Astronomical Medicine project is to make the specific modifications to existing medical imaging programs that enable their use in astronomical research. In particular, two software packages (3D Slicer and OsiriX are being modified and used on data from the "COMPLETE" Survey or Star-Forming Regions, in order to characterize the physical nature of gas that forms new stars in our galaxy.



The Initiative in Innovative Computing (IIC) is an interdisciplinary research and development center at Harvard, dedicated to using innovative computing tools to accelerate discovery across all of the scientific disciplines. The IIC's researchers work in close collaboration with scientists and engineers in other fields, fostering a two-way collaborative flow of ideas and inventions between basic science and computer science, academia and industry, professional staff and faculty, teachers and students. The IIC trains the next generation of creative and computationally capable scientists, and communicates with the public at large the value of computing in science, and about the science it enables.

### strategic role of the iic

We are taking the leadership position in many sciences today, by creating a strong interdisciplinary collaboration between experts in the scientific and computational disciplines, supported by an advanced computing infrastructure and skilled personnel. Unprecedented

### 07.10.07

New York Times Article A Trip Back in Time and Space

### 06.12.07

NY Times-Science-Felice Frankel Article She Calls It 'Phenomena.' Everyone Else Calls It Art.

#### 03.21.07

Initiative in Innovative Computing Inaugural Symposium

. IIC Insumuel Commercium





What does the wiring diagram of the human brain look like?

How do stars form in our Galaxy today?

How do humans interact with "big data"?

### What's "Dark Energy"?

What are the earliest structures in the Universe?

How should we model blood flow in the human body from atomic to physiological scales?

How should we use technology in teaching?

Can we predict how often a genotype will produce disease? How can collaborative technologies accelerate Stem Cell & Neuroinformatics Research?

How should we mine routine medical records, on vast scales, to cure diseases?