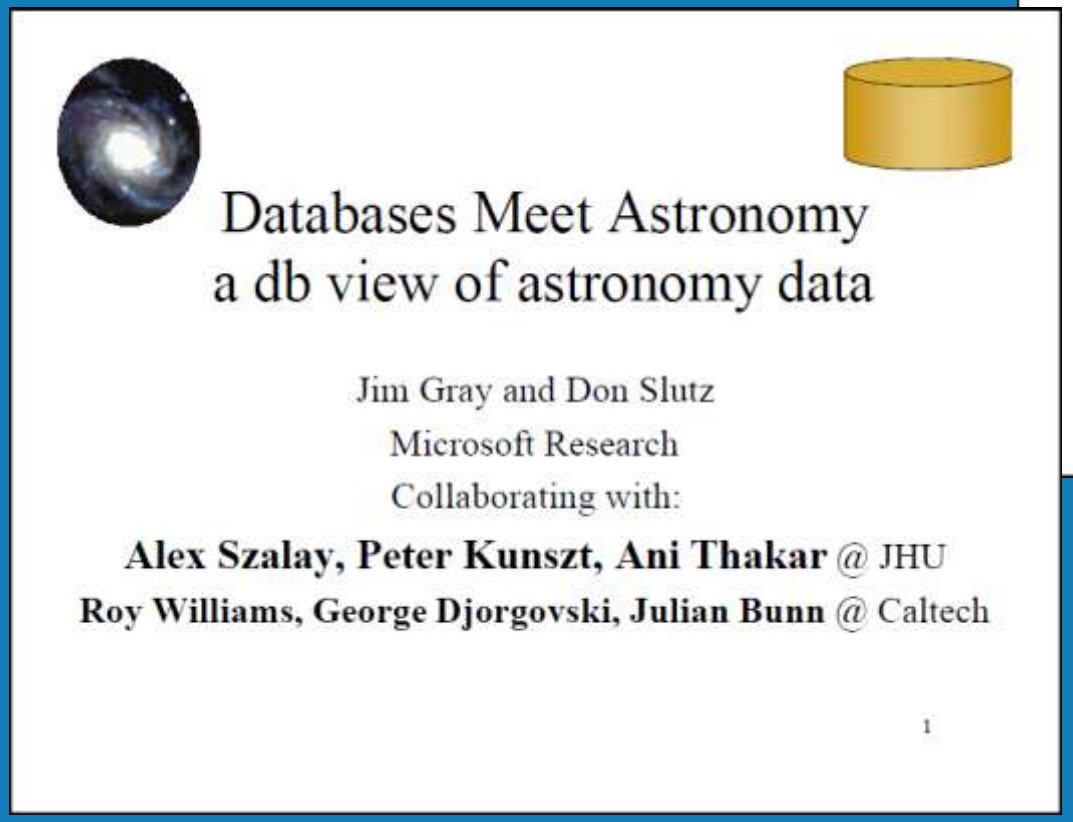
ECR | An Information Model for Interactive Data Visualization

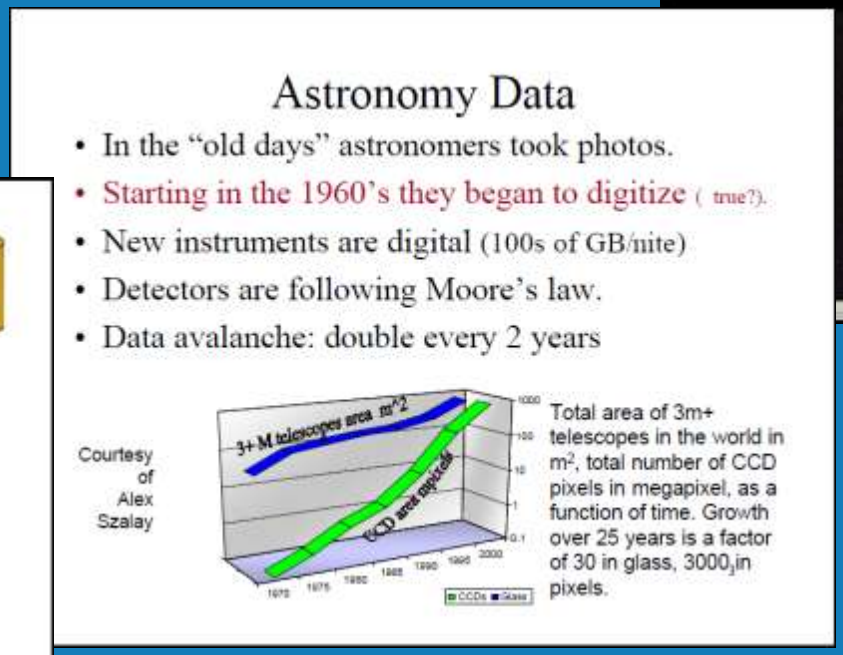
Curtis Wong
Principal Researcher
curtis.wong@microsoft.com

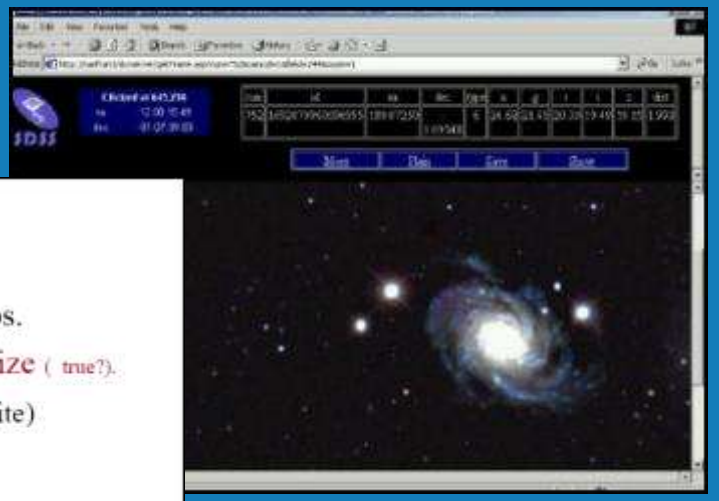
Microsoft Research

A little bit of history…

- 2002 Jim Gray Talk







Call to Action

- If you are a vis-person: we need you (and we know it).

Microsoft Research

Which led to SkyServer DR2…

- March 2004

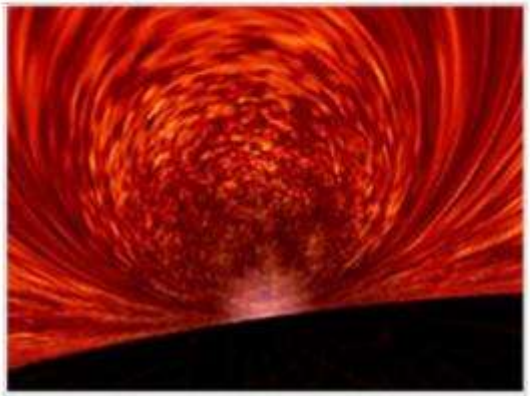
When Alex Szalay Suggested…

- May 2005

Visualization of Astrophysical Data: Bringing Together Science, Art, and Education

May 25 - 27, 2005

Website



Visualization of Astrophysical Data: Bringing Together Science, Art, and Education

The visualization of real astrophysical data sets is a powerful tool for communicating science to the public and for teaching. The confluence of high quality data sets such as SDSS & WMAP, advances in computational techniques, and the continued march of Moore's law has enabled the use of stunningly beautiful and scientifically accurate images, animations, and interactives in a variety of settings (e.g. TV programs, museums, websites, digital planetaria, magazines, & undergraduate classrooms). Two dimensional, three dimensional, and hyper-dimensional (e.g. color coded 3D data) representations convey large amounts of information in a visceral fashion that can inform both experts and the public. As the data and techniques have progressed the boundaries between art and science have begun to blur and move towards research. This workshop will bring together astrophysicists, visualizers, and educators to discuss the current status and to debate the future direction of astronomical visualization as a tool for research, education, and public outreach.

Organizers: Randy Landsberg, Josh Frieman, Andrew Hamilton (CU Boulder), Andrey Kravtsov, Mark SubbaRao, & Alex Szalay (JHU).

Read more >>

The Visualization of Astrophysical Data:

Bringing together Science, Art and Education

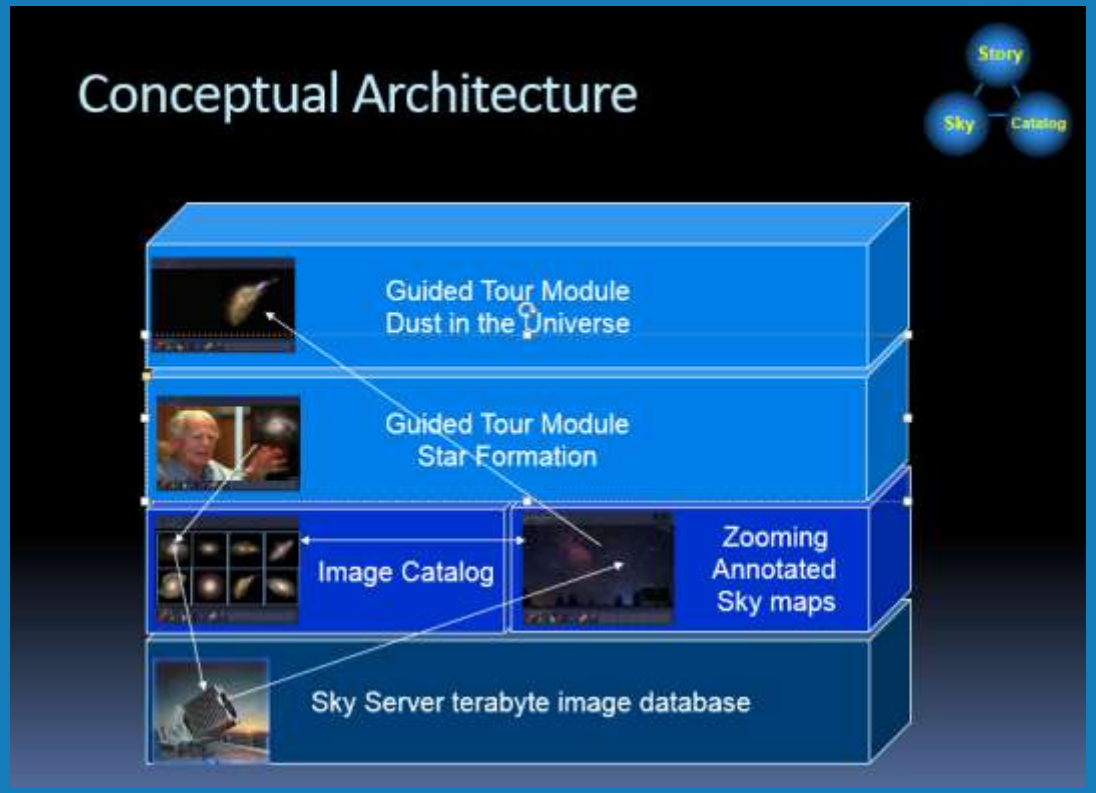
This site will serve as a staging area for workshop material, until the official workshop page is up on the KICP website.

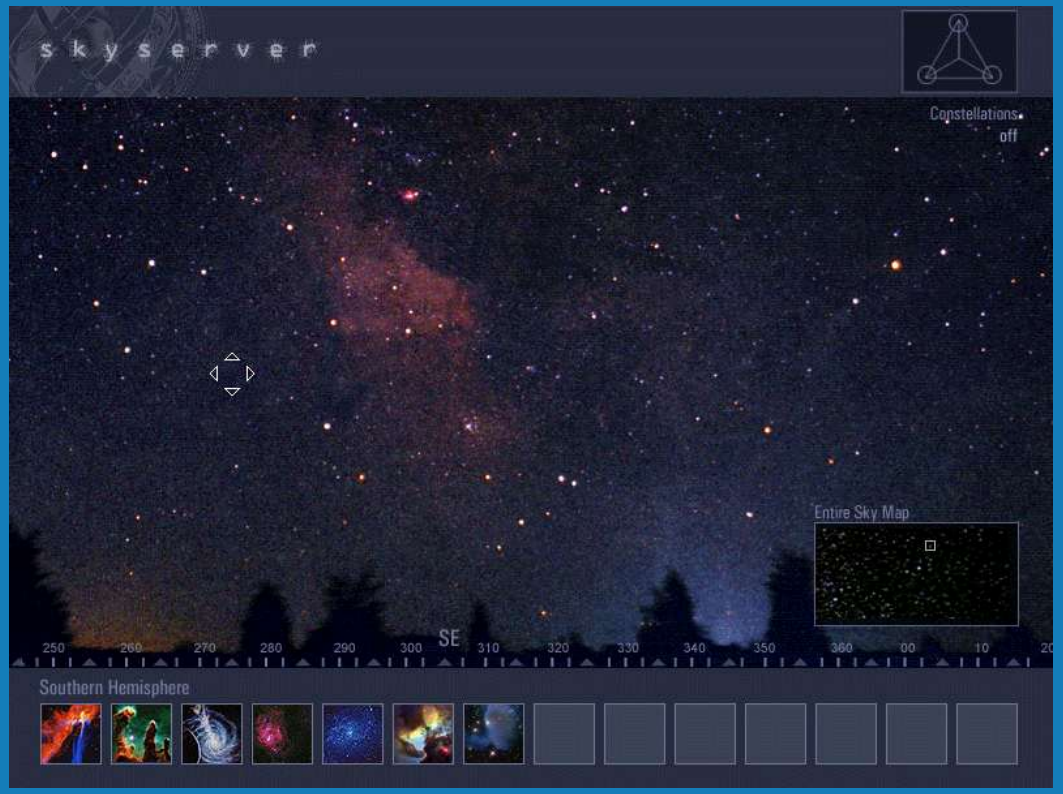
Presentations and Video Clips

	Presenter	Topic	Presentation+Resources
	Randy Landsberg	Welcome and Intro	Talk(PDF)
	Adrienne Gauthier	Instructional Design Strategies for Instructional Technologies Utilizing Data Visualization	Poster(PDF)
	Curtis Wong	Interactive narrative content and context for visualization	Talk(PPT) - Interactive from Da Vinci CDROM (Windows only) Codescope.zip
		Accuracy & Aesthetics: Scientific	

Where I pitched the Universe Project…

- May 2005

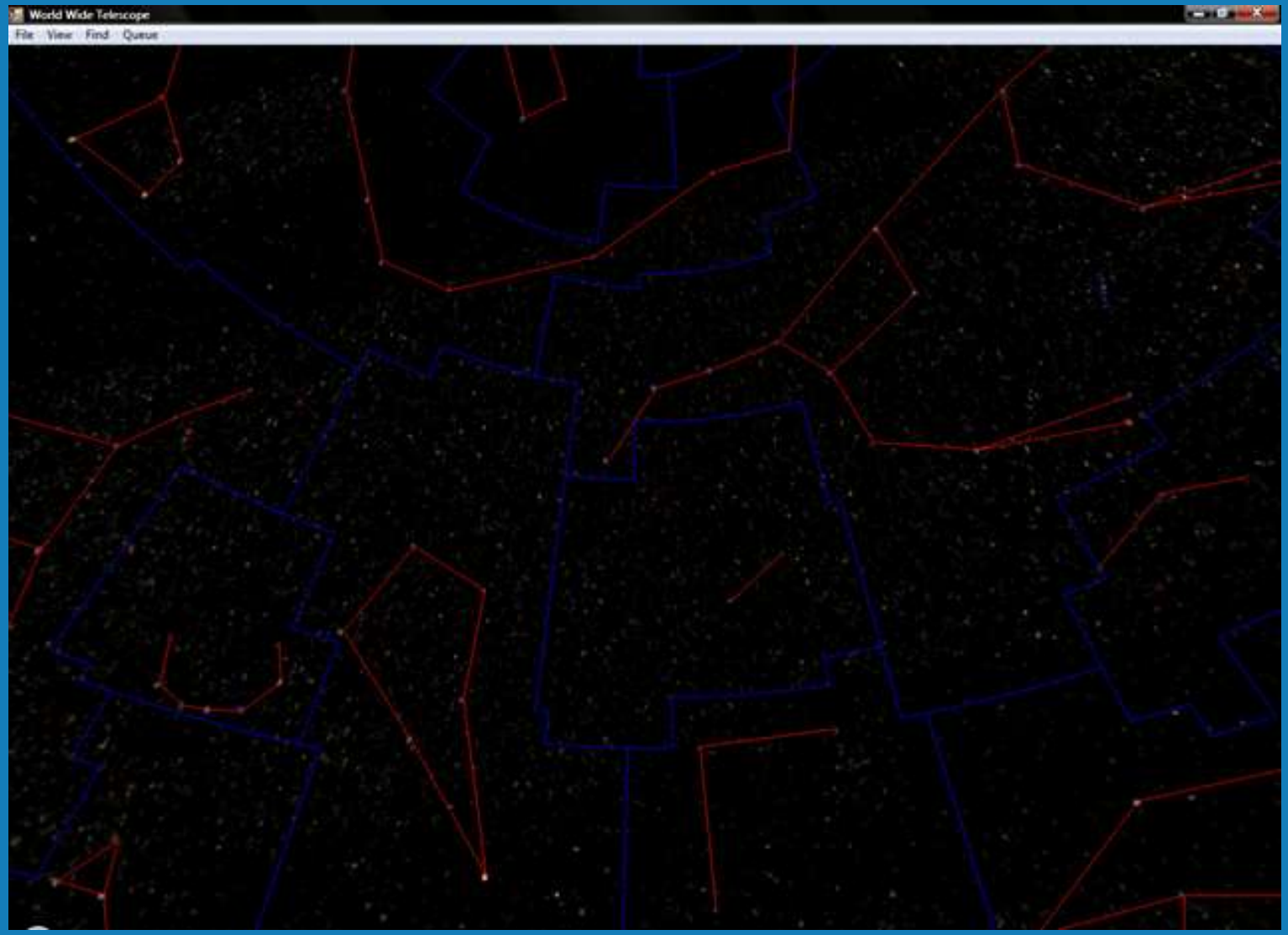




Build a Prototype!

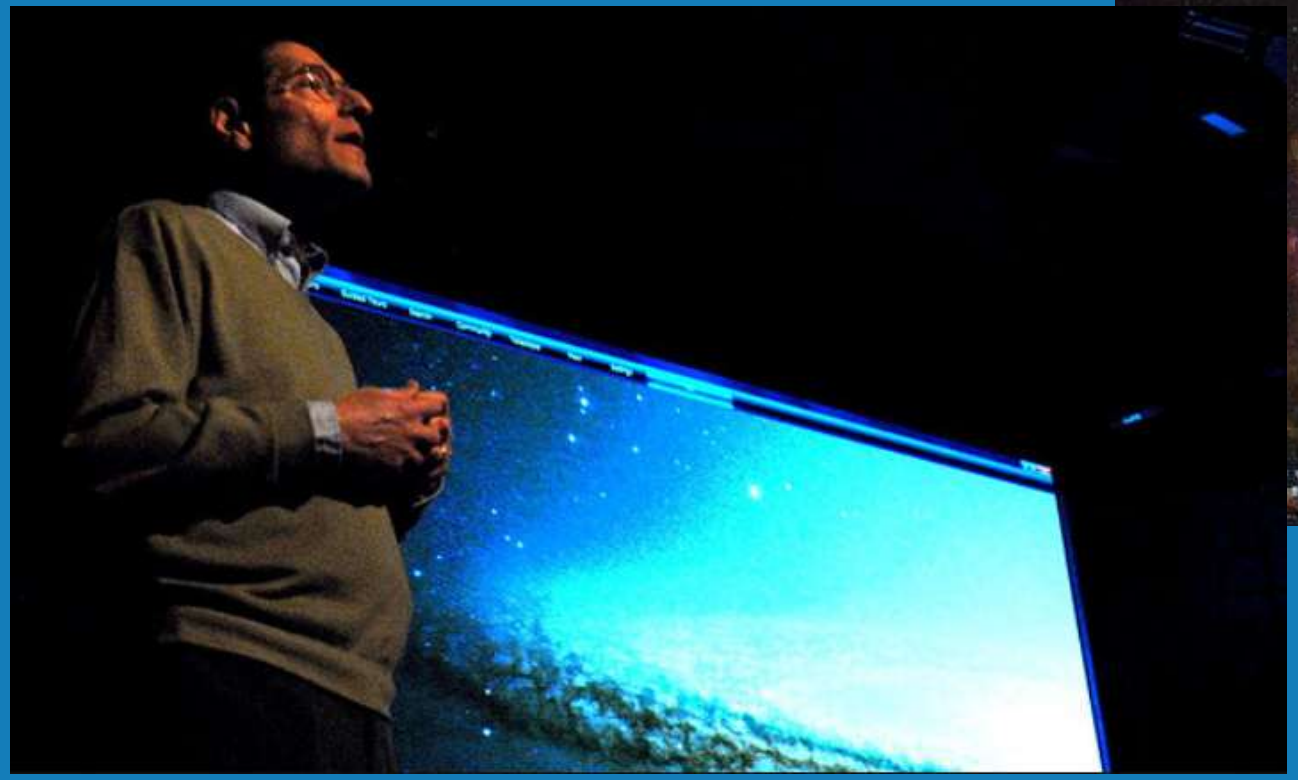
- 2006
- Jonathan Fay's SDSS Tiled image browser

- Heads down for 18 months…



Launch it at the TED Conference

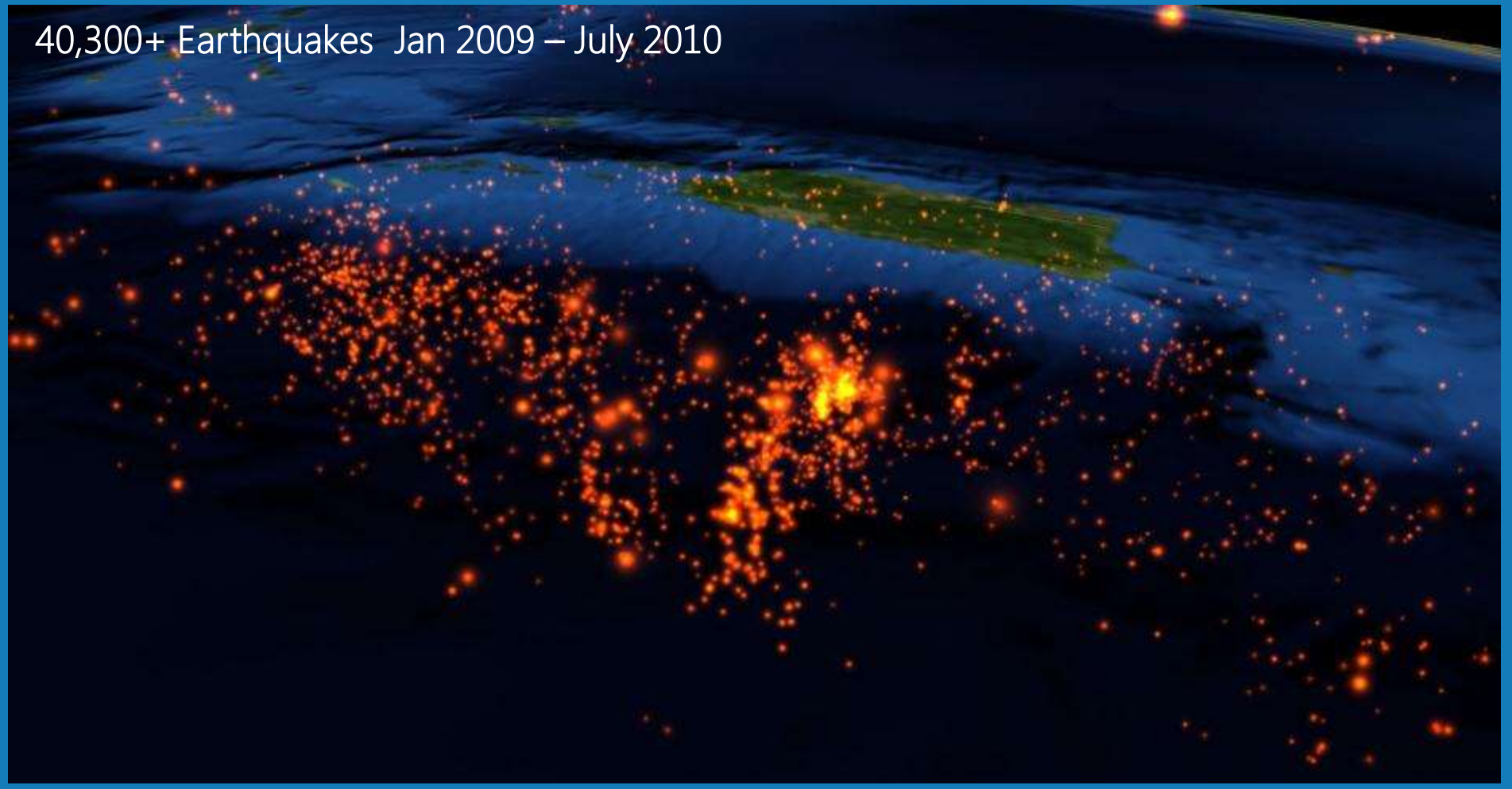
- 2008



Roy Gould speaking at TED

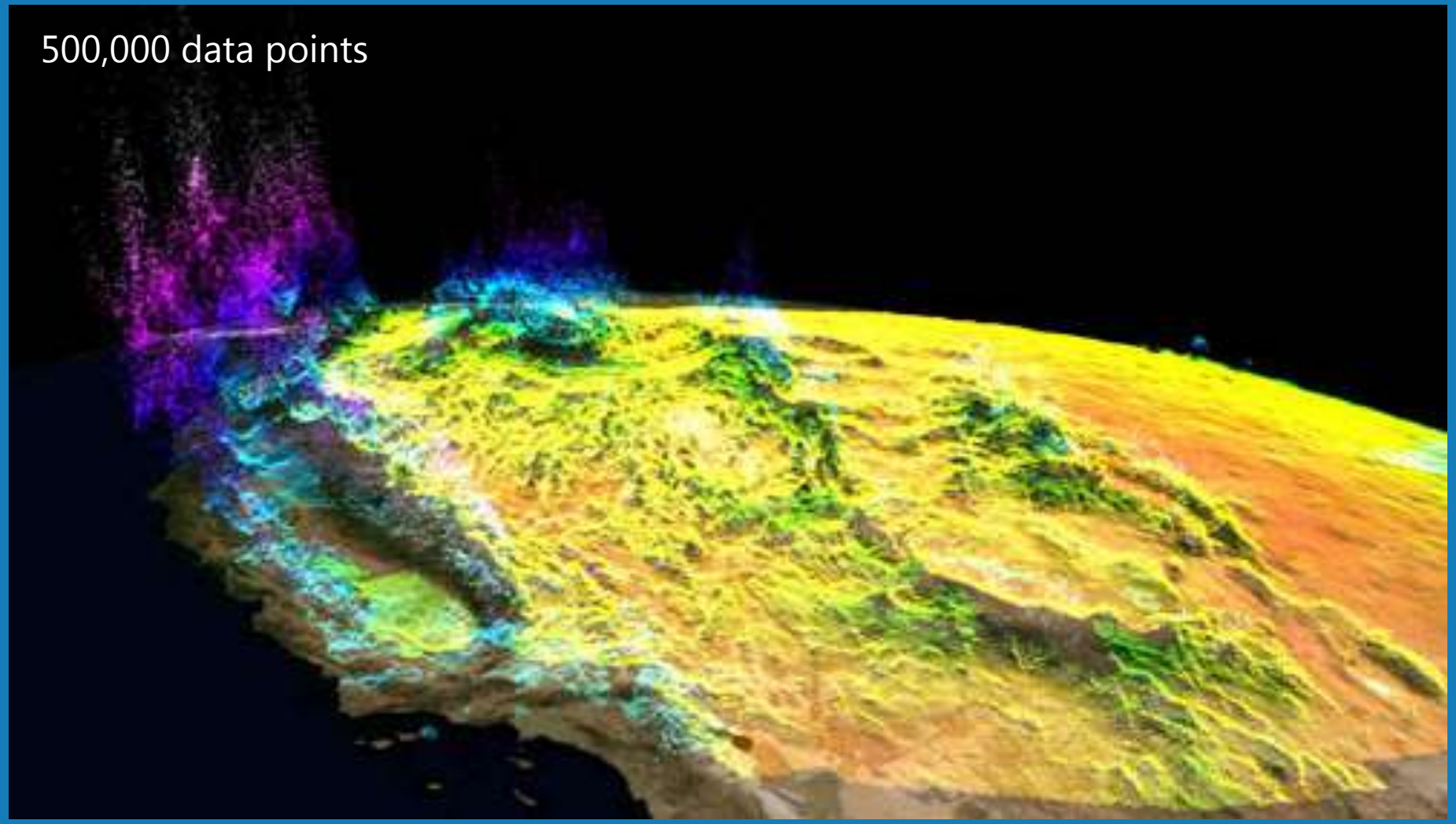
Visualizing Earthquake data

- 2010



Visualize 30 year precipitation

- 2011



Visualize sensor data over time with WWT

- 2011

Exploring linked visualizations in WWT

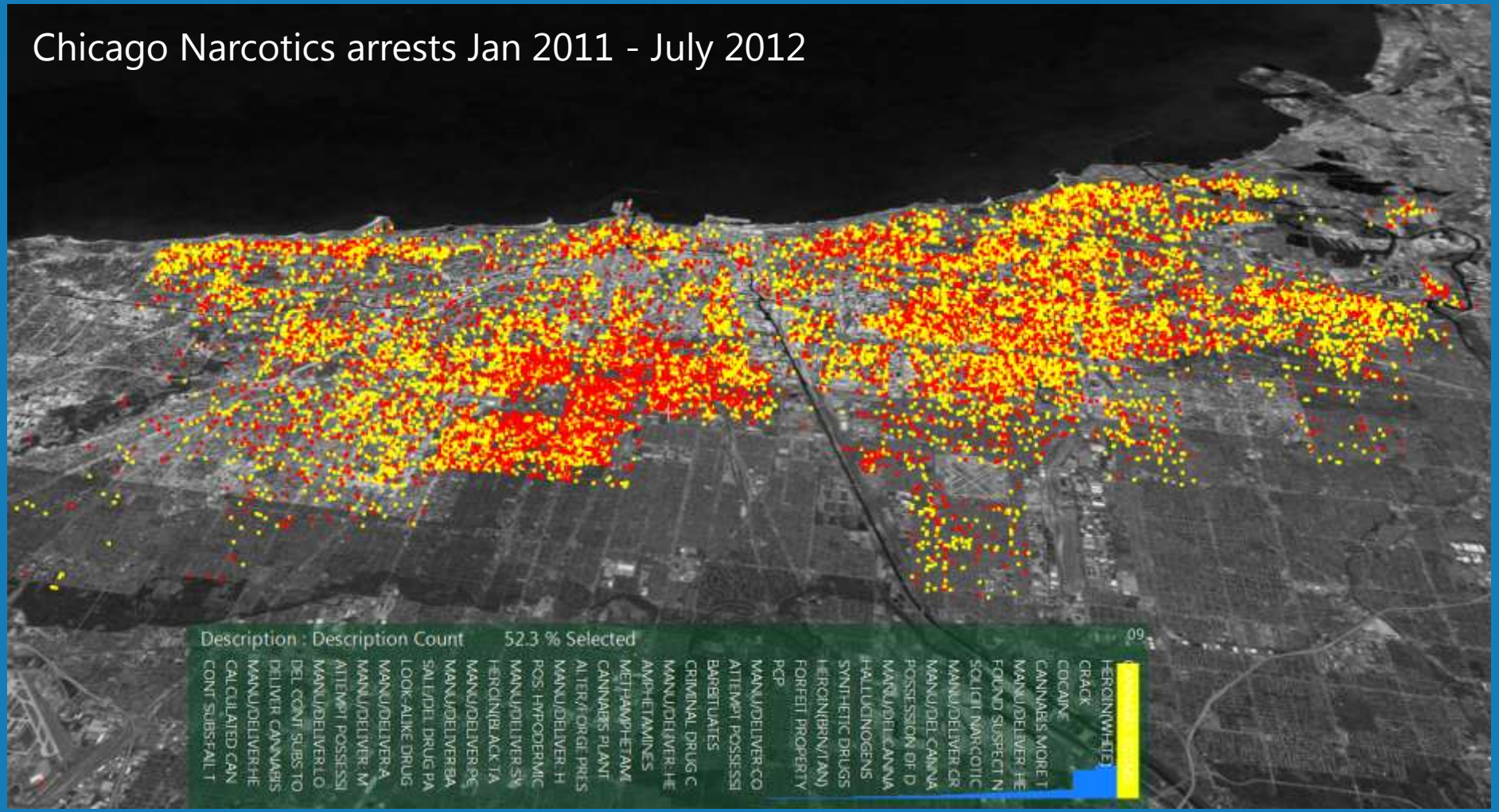
- 2011

	A	B	C	D	E	F	G	H	I	J
1	Case #	Date	Block	IUCR	Description	Location Description	Beat	Ward	Latitude	Longitude
2	HT655009	12/31/2011 23:10	063XX S EVANS AVE	1811	CANNABIS 30GMS OR LESS	STREET	312	20	41.7792	-87.6073
3	HT654959	12/31/2011 22:50	001XX S LOCKWOOD AVE	1812	CANNABIS MORE THAN 30GMS	VEHICLE NON-COMMERCIAL	1522	29	41.87895	-87.7578
4	HT654969	12/31/2011 22:41	040XX W MONROE ST	2024	HEROIN(WHITE)	RESIDENCE PORCH/HALLWAY	1115	28	41.87985	-87.7269
5	HT654966	12/31/2011 22:30	078XX S VINCENNES AVE	1811	CANNABIS 30GMS OR LESS	STREET	621	17	41.75135	-87.6342
6	HT654968	12/31/2011 22:19	046XX N RACINE AVE	1811	CANNABIS 30GMS OR LESS	SIDEWALK	2311	46	41.96656	-87.6595
7	HT654970	12/31/2011 22:19	012XX N CENTRAL AVE	2025	HALLUCINOGENS	PARKING LOT/GARAGE(NON.RES	2532	37	41.90224	-87.7655
8	HT654946	12/31/2011 22:00	040XX W MADISON ST	2024	HEROIN(WHITE)	SIDEWALK	1115	28	41.88059	-87.7278
9	HT654944	12/31/2011 22:00	0000X E CONGRESS PKWY	2025	HALLUCINOGENS	OTHER	132	2	41.8758	-87.6257
10	HT654912	12/31/2011 21:40	013XX W 95TH ST	1811	CANNABIS 30GMS OR LESS	SIDEWALK	2222	21	41.72141	-87.6564
11	HT654917	12/31/2011 21:39	032XX N KARLOV AVE	1811	CANNABIS 30GMS OR LESS	STREET	1731	31	41.93966	-87.7295
12	HT654935	12/31/2011 21:19	061XX S ARCHER AVE	2020	AMPHETAMINES	STREET	811	23	41.79434	-87.7706
13	HT654893	12/31/2011 20:35	082XX S WOLCOTT AVE	2024	HEROIN(WHITE)	ALLEY	614	18	41.74351	-87.6707
14	HT654840	12/31/2011 19:49	002XX W 111TH ST	1811	CANNABIS 30GMS OR LESS	SIDEWALK	513	34	41.69262	-87.6296
15	HT654826	12/31/2011 19:19	011XX N LOREL AVE	1811	CANNABIS 30GMS OR LESS	STREET	1524	37	41.90177	-87.7594
16	HT654781	12/31/2011 18:45	039XX W CERMAK RD	2024	HEROIN(WHITE)	SIDEWALK	1014	24	41.85166	-87.7242
17	HT654785	12/31/2011 18:24	063XX N CLAREMONT AVE	1811	CANNABIS 30GMS OR LESS	STREET	2413	50	41.99681	-87.6888
18	HT654775	12/31/2011 18:14	001XX N LARAMIE AVE	2027	CRACK	OTHER	1523	28	41.88414	-87.7554
19	HT654777	12/31/2011 18:11	050XX S LA CROSSE AVE	1811	CANNABIS 30GMS OR LESS	STREET	814	23	41.80195	-87.7444
20	HT654772	12/31/2011 18:10	086XX S INGLESIDE AVE	1812	CANNABIS MORE THAN 30GMS	ALLEY	632	8	41.73718	-87.6013
21	HT654784	12/31/2011 18:05	010XX W WILSON AVE	1811	CANNABIS 30GMS OR LESS	SIDEWALK	2313	46	41.9653	-87.6567
22	HT654737	12/31/2011 17:35	055XX W FLOURNOY ST	1811	CANNABIS 30GMS OR LESS	STREET	1522	29	41.87195	-87.7635
23	HT654723	12/31/2011 17:25	059XX S EMERALD AVE	1811	CANNABIS 30GMS OR LESS	SIDEWALK	711	16	41.78596	-87.6439
24	HT654576	12/31/2011 16:45	050XX S WOOD ST	1811	CANNABIS 30GMS OR LESS	STREET	931	16	41.80231	-87.6698
25	HT654489	12/31/2011 14:00	003XX N LOCKWOOD AVE	2024	HEROIN(WHITE)	SIDEWALK	1523	28	41.88649	-87.7579
26	HT654510	12/31/2011 13:07	077XX S EVANS AVE	2027	CRACK	RESIDENCE	624	6	41.75382	-87.6067
27	HT654351	12/31/2011 12:20	012XX N LOCKWOOD AVE	1811	CANNABIS 30GMS OR LESS	SIDEWALK	2532	37	41.90237	-87.7581
28	HT654428	12/31/2011 12:09	006XX N LARAMIE AVE	2028	SYNTHETIC DRUGS	STREET	1532	28	41.89193	-87.7553
29	HT654318	12/31/2011 11:20	055XX W JACKSON BLVD	1811	CANNABIS 30GMS OR LESS	APARTMENT	1522	29	41.87645	-87.7628
30	HT654350	12/31/2011 10:45	011XX N MONTICELLO AVE	2027	CRACK	APARTMENT	1112	27	41.90137	-87.7176
31	HT654157	12/31/2011 8:46	015XX N CENTRAL PARK AVE	2024	HEROIN(WHITE)	STREET	2535	26	41.90898	-87.7169
32	HT654333	12/31/2011 8:15	014XX N LARAMIE AVE	1811	CANNABIS 30GMS OR LESS	RESIDENCE	2532	37	41.9064	-87.756
33	HT654394	12/31/2011 8:00	003XX N PINE AVE	2093	FOUND SUSPECT NARCOTICS	RESIDENCE	1523	28	41.88678	-87.7636
34	HT654080	12/31/2011 5:10	035XX S RHODES AVE	2093	FOUND SUSPECT NARCOTICS	APARTMENT	212	4	41.82964	-87.6146

Chicago Narcotics arrests Jan 2011 - July 2012

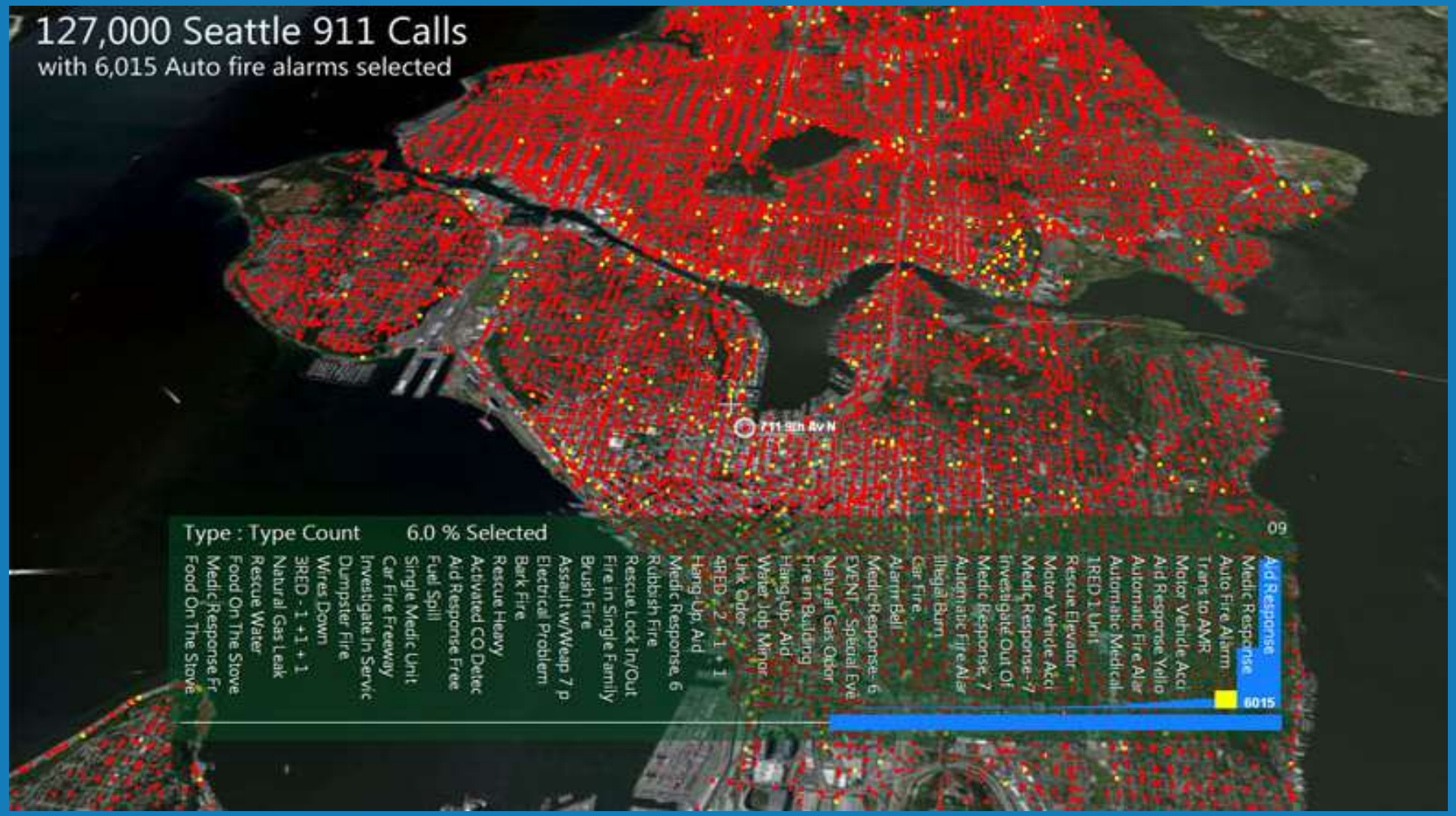
Exploring linked visualizations in WWT

- 2011



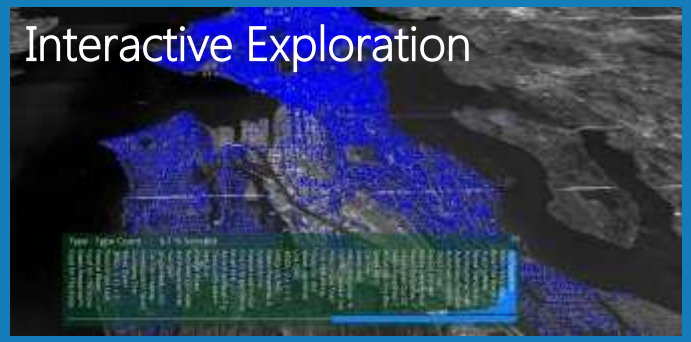
Linked data visualization explorations

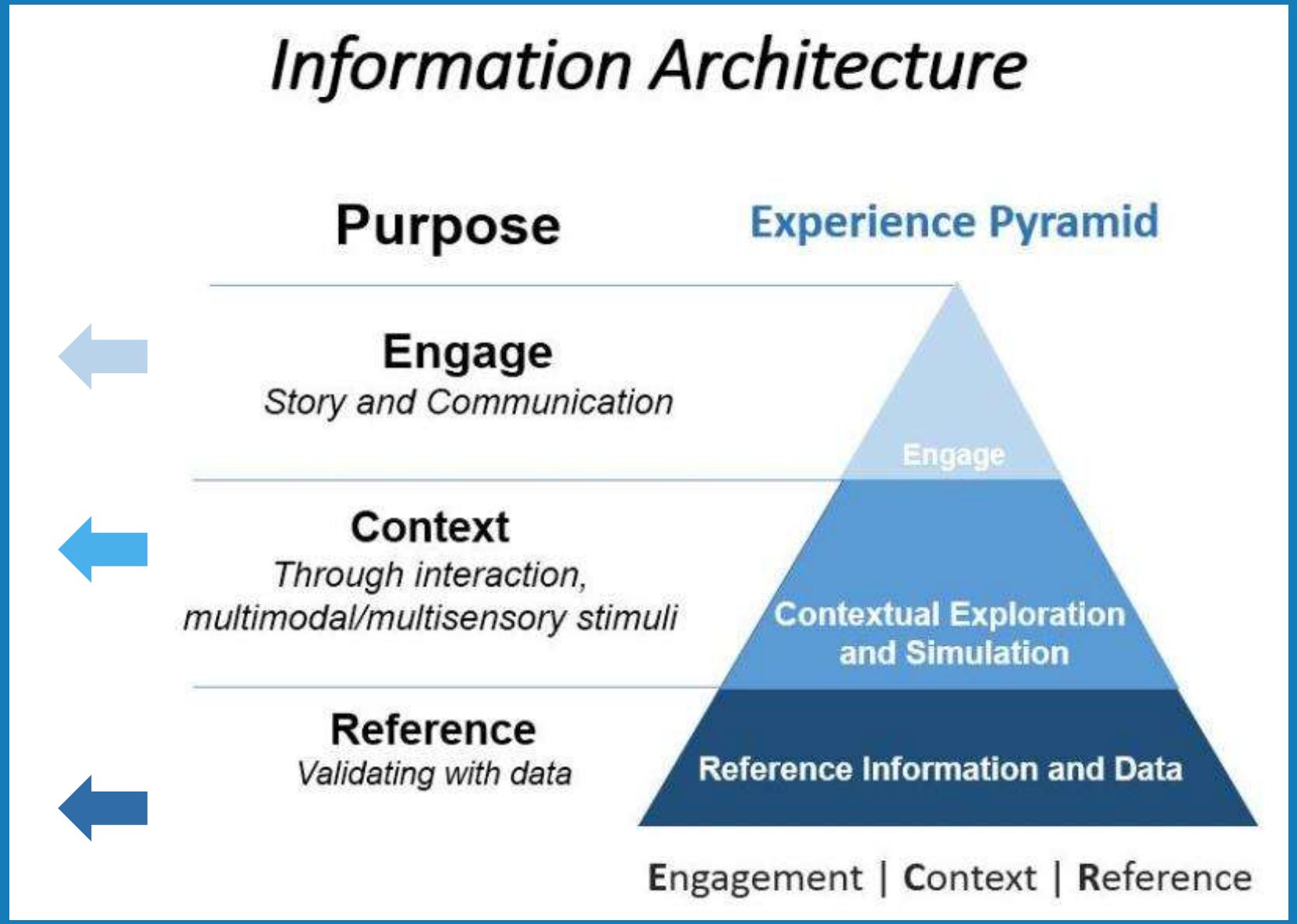
- 2011



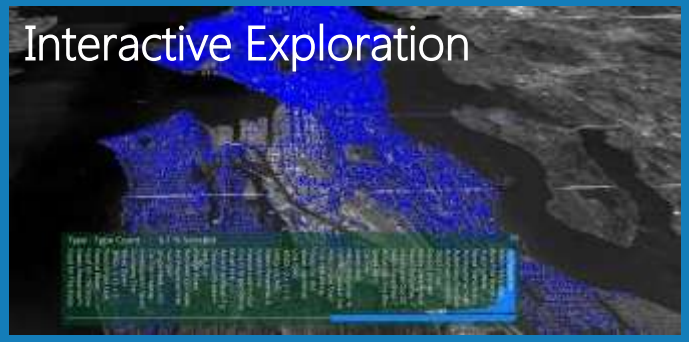
ECR | An Information Architecture for Visualization

- 2011





Bigger data visualization opportunities



- GPU's in the cloud
  - Move visualization and analytics to the highly structured data
  - Guided Tour to the data
    - Defined path in the data sets
    - Rendered in real time as HD video
    - Fully Interactive anytime unlike true video
    - Useful for sharing data insight
    - Showing data in context
  - Guided Tour is metaview about the data
  - Interactive Exploration of outliers
  - Linking to other related data and metadata
    - Comparison and context

Opportunities and challenges

- Dirty Data
- Poorly structured data
- Analytics
- Machine Learning
- N-dimensional slicers
- Spatial Queries
- Managing complexity | Ease of use
- Different or non-existing spatial coordinates
- API's
- etc…

Other Explorations

- Very large tiled displays – 35 to 100 megapixels
- Natural User Interaction
  - Speech, Gesture, Touch, Multimodal
  - Devices
- New graphics rendering architecture – Project vX

Integrated vs. Connected Visualization Applications

- Integrated – each has some of what you need
  - (ESRI, Spotfire)
- Hubs – message passing btw programs (SAMP)
  - Aladin/WWT/TOPCAT
- Programming tools to make it all work
  - Matlab, Python, R, etc.

Summary of applicable visualization programs

Astron. Nachr. / AN xxx, No. xx, 001–10 (2012) / DOI *please set DOI!*

Principles of High-Dimensional Data Visualization in Astronomy

Alyssa A. Goodman[1,*]

Harvard-Smithsonian Center for Astrophysics, Cambridge, MA, USA

Received 20 Apr 2012, accepted 20 Apr 2012
Published online later

Key words techniques: image processing; methods: data analysis; techniques: radial velocities; cosmology: large-scale structure; ISM: clouds

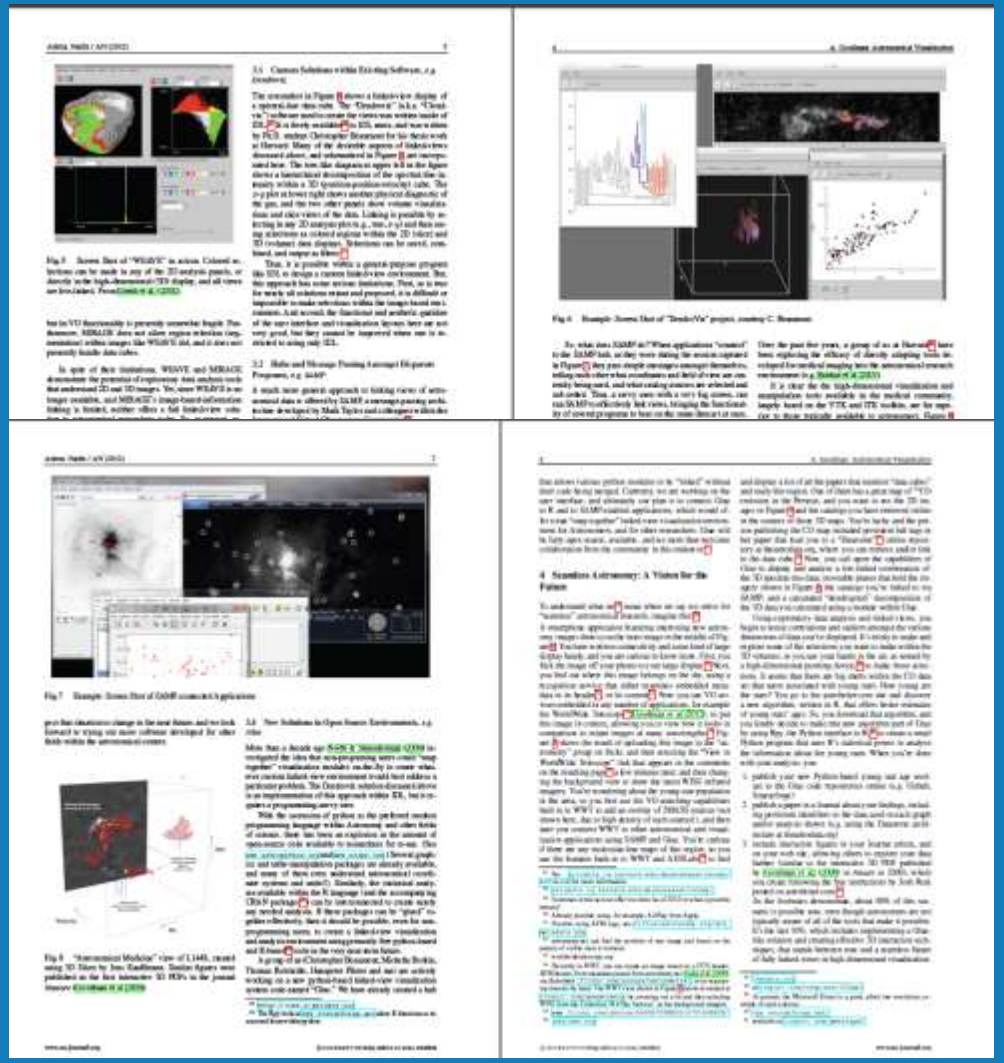
Astronomical researchers often think of analysis and visualization as separate tasks. In the case of high-dimensional data sets, though, interactive *exploratory data visualization* can give far more insight than an approach where data processing and statistical analysis are followed, rather than accompanied, by visualization. This paper attempts to charts a course toward "linked view" systems, where multiple views of high-dimensional data sets update live as a researcher selects, highlights, or otherwise manipulates, one of several open views. For example, imagine a researcher looking at a 3D volume visualization of simulated or observed data, and simultaneously viewing statistical displays of the data set's properties (such as an x-y plot of temperature vs. velocity, or a histogram of vorticities). Then, imagine that when the researcher selects an interesting group of points in any one of these displays, that the same points become a highlighted subset in all other open displays. Selections can be graphical or algorithmic, and they can be combined, and saved. For tabular (ASCII) data, this kind of analysis has long been possible, even though it has been under-used in much of science. The bigger issue for Astronomy and other "high-dimensional" fields, though, is that no extant system allows for full integration of images and data cubes within a linked-view environment. The paper concludes its history and analysis of the present situation with suggestions that look toward cooperatively-developed open-source modular software as a way to create an evolving, flexible, high-dimensional, linked-view visualization environment useful in astrophysical research.

1 Introduction

Historically, Astronomy has been a visual science. Thousands of years ago observations were carried out with the naked eye; hundreds of years ago telescopes augmented the eye; and during the last century sensitive film and CCD recording devices enhanced what the eye could see. More capabilities to statistical principles are, and should continue to be, on the rise within modern astronomy.

2 Data-Dimensions-Display

There are three simple words to keep in mind when one sets out to explore and/or explain high-dimensional infor-



http://projects.iq.harvard.edu/seamlessastronomy/files/heidelberg_ag.pdf

Microsoft®
Research