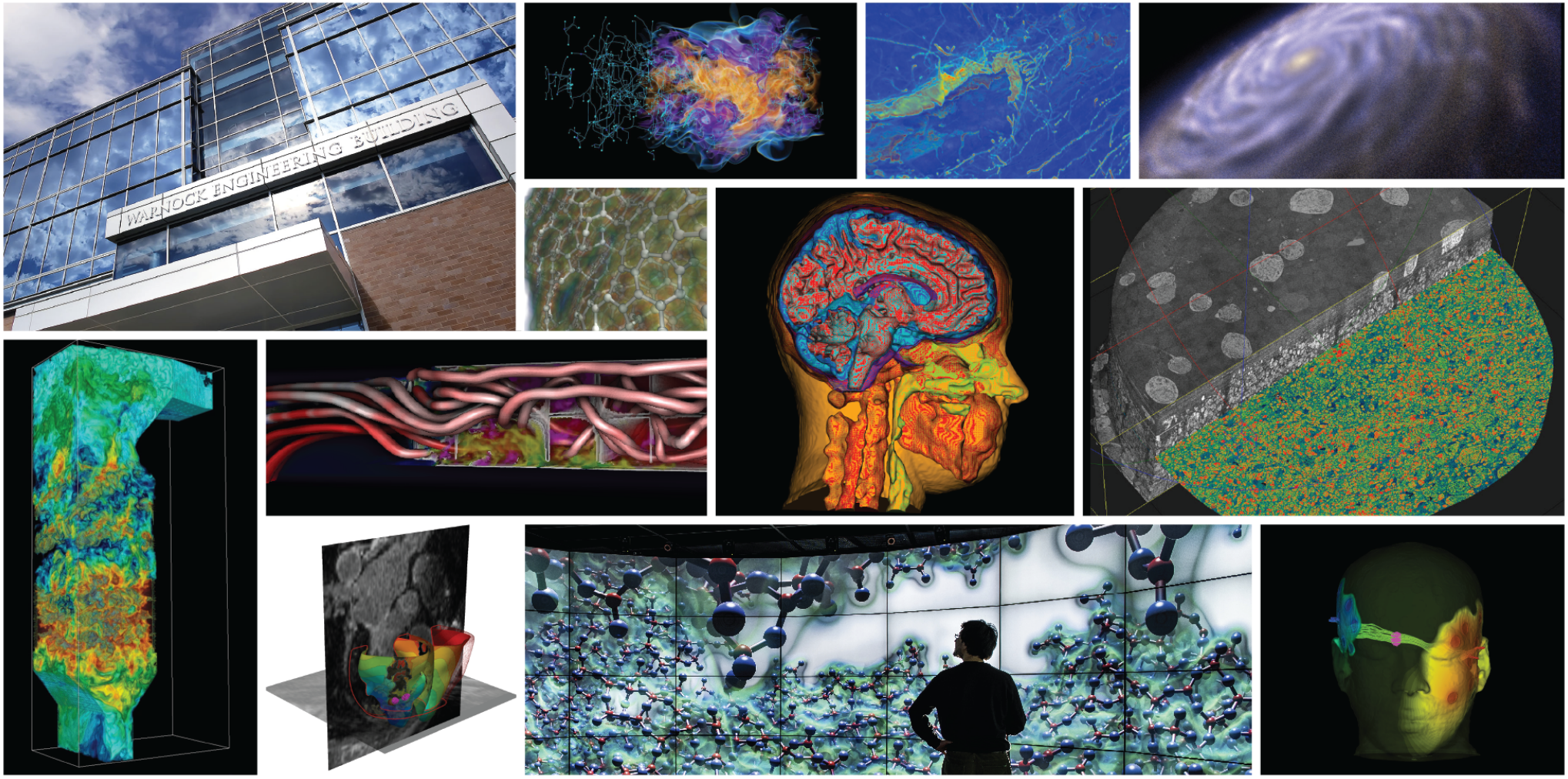
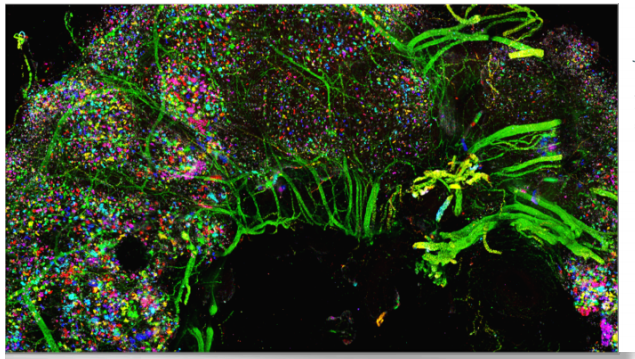
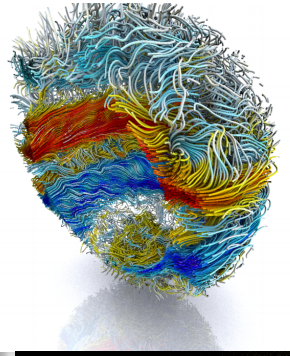
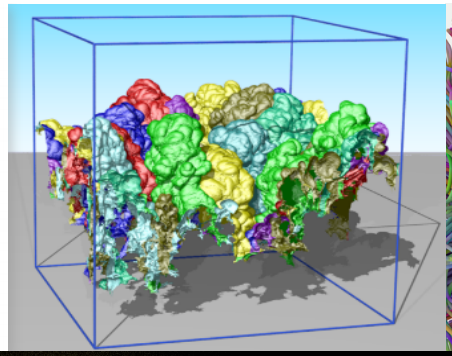
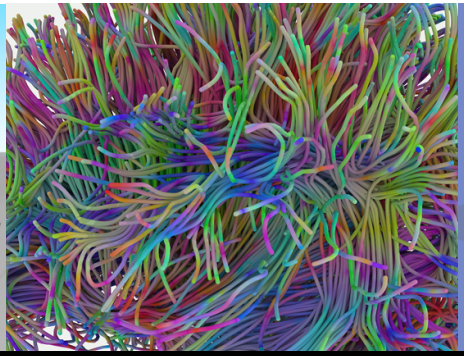
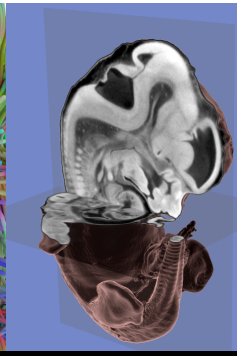
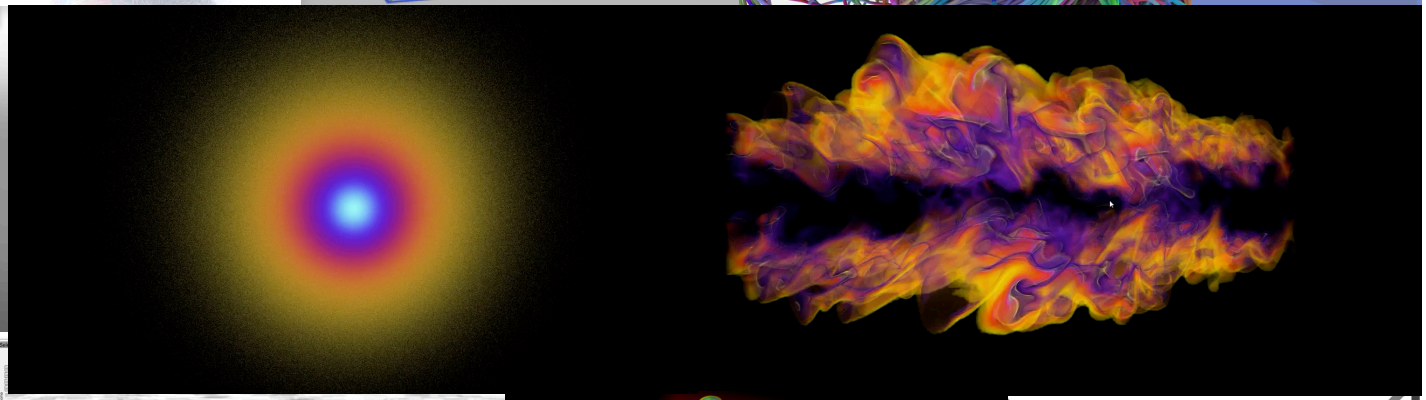
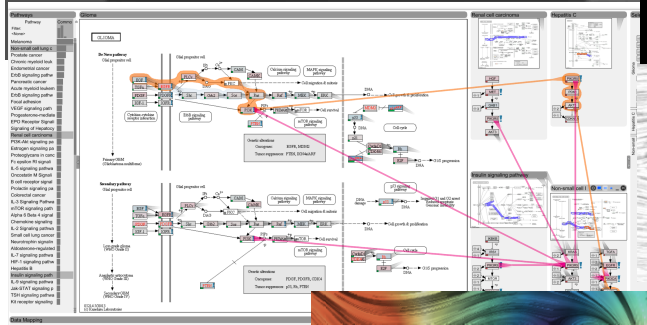
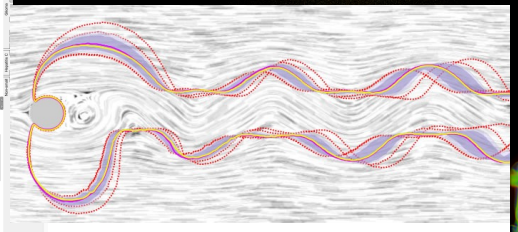
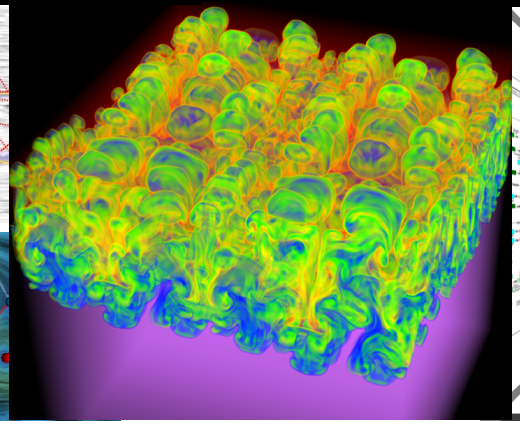
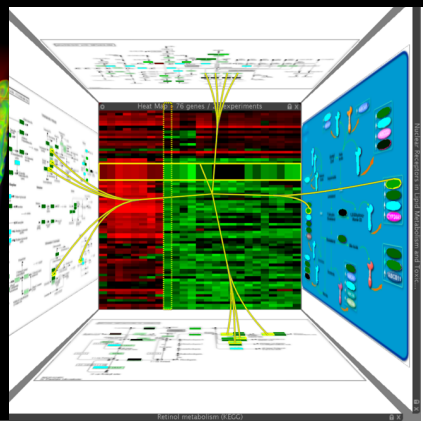
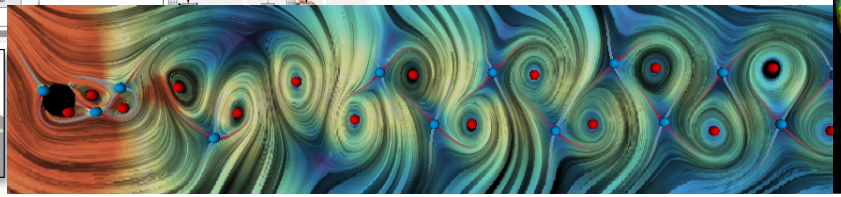
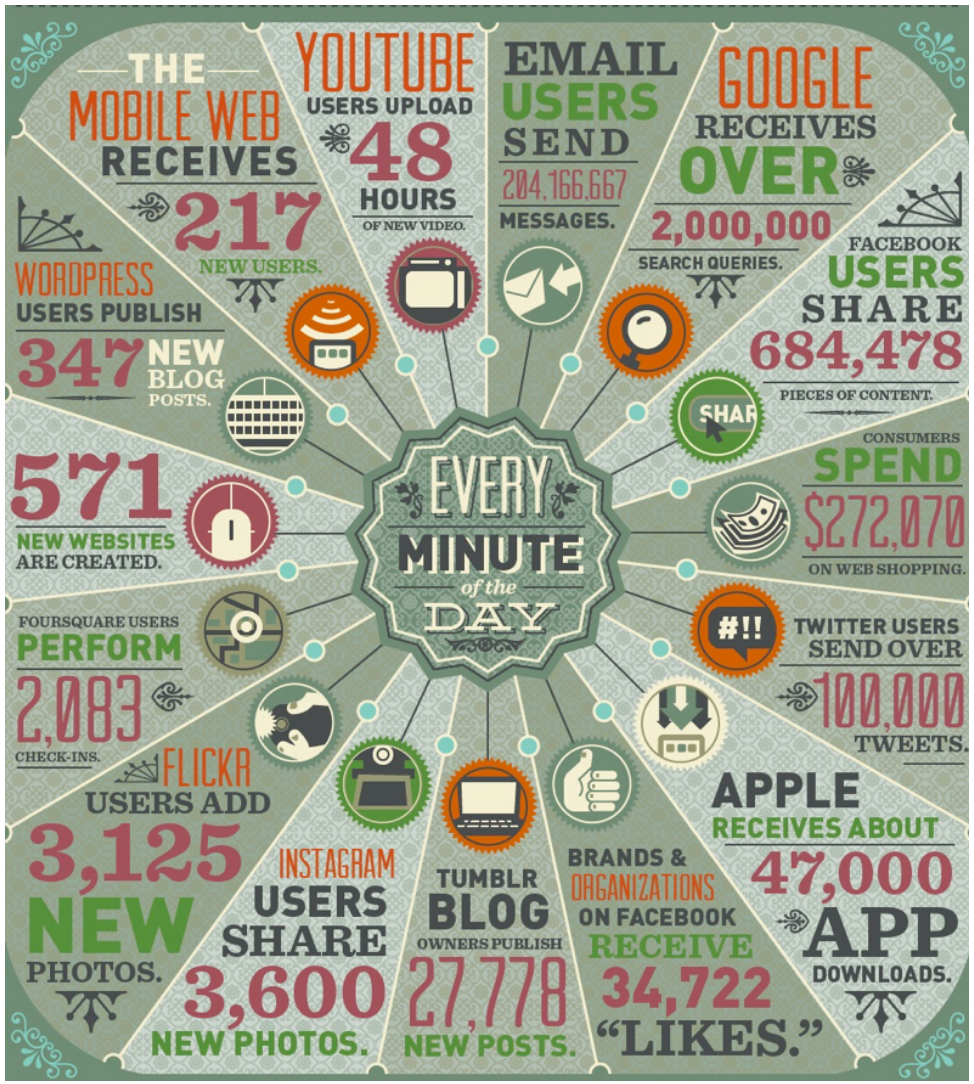
Visual Analysis Challenges in the Age of Data

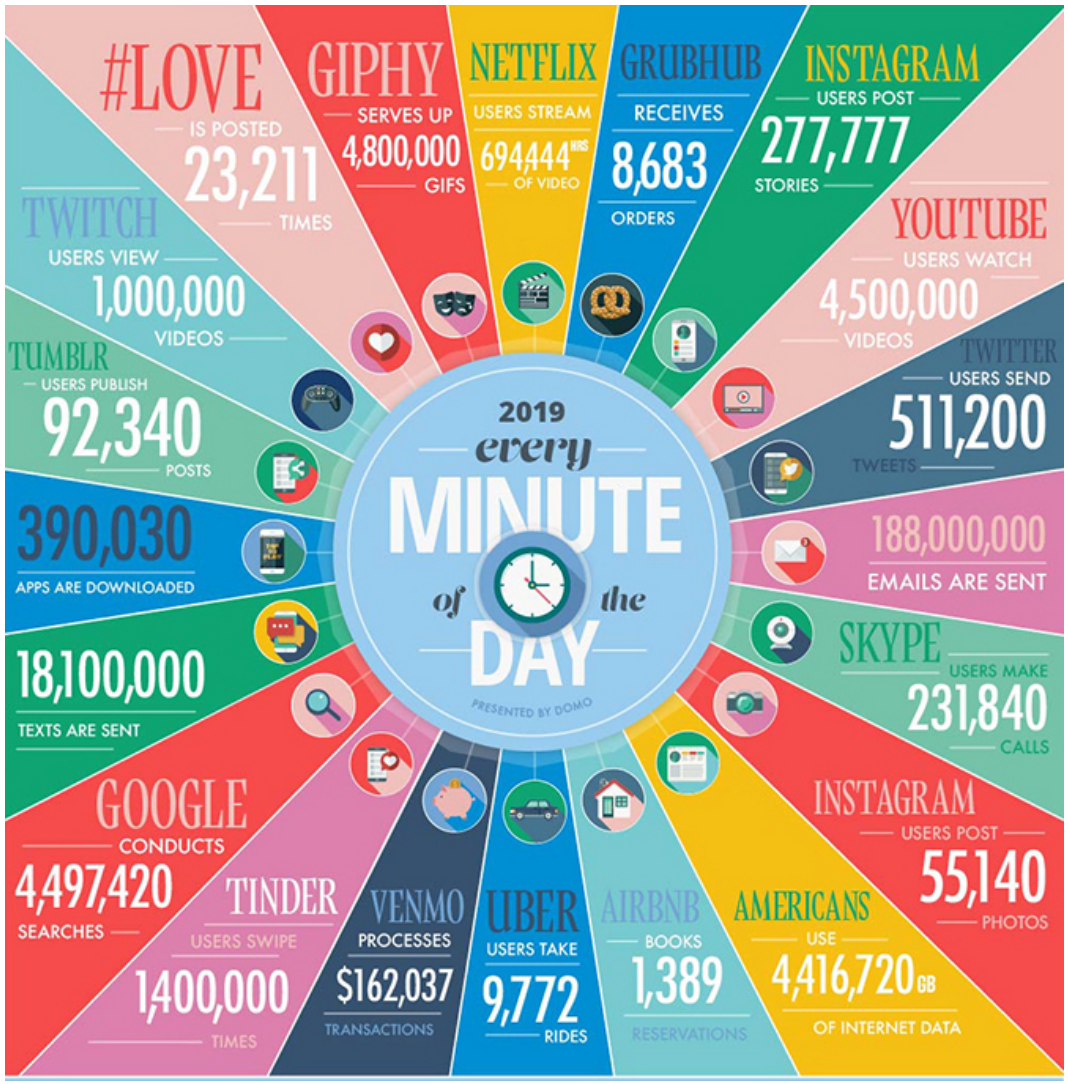


SCI
INSTITUTE


EVERY MINUTE of the DAY
THE MOBILE WEB RECEIVES 217 NEW USERS.
YOUTUBE USERS UPLOAD 48 HOURS OF NEW VIDEO.
EMAIL USERS SEND 204,166,667 MESSAGES.
GOOGLE RECEIVES OVER 2,000,000 SEARCH QUERIES.
FACEBOOK USERS SHARE 684,478 PIECES OF CONTENT.
CONSUMERS SPEND $272,070 ON WEB SHOPPING.
TWITTER USERS SEND OVER 100,000 TWEETS.
APPLE RECEIVES ABOUT 47,000 APP DOWNLOADS.
BRANDS & ORGANIZATIONS ON FACEBOOK RECEIVE 34,722 "LIKES."
TUMBLR BLOG OWNERS PUBLISH 27,778 NEW POSTS.
INSTAGRAM USERS SHARE 3,600 NEW PHOTOS.
FLICKR USERS ADD 3,125 NEW PHOTOS.
FOURSQUARE USERS PERFORM 2,083 CHECK-INS.
571 NEW WEBSITES ARE CREATED.
WORDPRESS USERS PUBLISH 347 NEW BLOG POSTS.


2019 every MINUTE of the DAY
PRESENTED BY DOMO
#LOVE IS POSTED 23,211 TIMES
GIPHY SERVES UP 4,800,000 GIFS
NETFLIX USERS STREAM 694,444 HRS OF VIDEO
GRUBHUB RECEIVES 8,683 ORDERS
INSTAGRAM USERS POST 277,777 STORIES
TWITCH USERS VIEW 1,000,000 VIDEOS
YOUTUBE USERS WATCH 4,500,000 VIDEOS
TUMBLR USERS PUBLISH 92,340 POSTS
TWITTER USERS SEND 511,200 TWEETS
390,030 APPS ARE DOWNLOADED
188,000,000 EMAILS ARE SENT
18,100,000 TEXTS ARE SENT
SKYPE USERS MAKE 231,840 CALLS
GOOGLE CONDUCTS 4,497,420 SEARCHES
INSTAGRAM USERS POST 55,140 PHOTOS
TINDER USERS SWIPE 1,400,000 TIMES
VENMO PROCESSES $162,037 TRANSACTIONS
UBER USERS TAKE 9,772 RIDES
AIRBNB BOOKS 1,389 RESERVATIONS
AMERICANS USE 4,416,720 GB OF INTERNET DATA

Large-Scale Visualization Challenges

Analysis

Scalability

Users and Usability

Large-Scale Visualization Challenges

Analysis

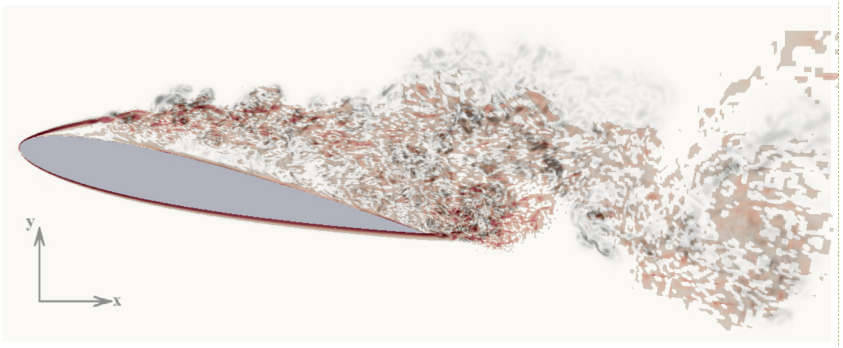
- High Order Simulation
- Uncertainty Quantification and Visualization
- Topological Data Analysis

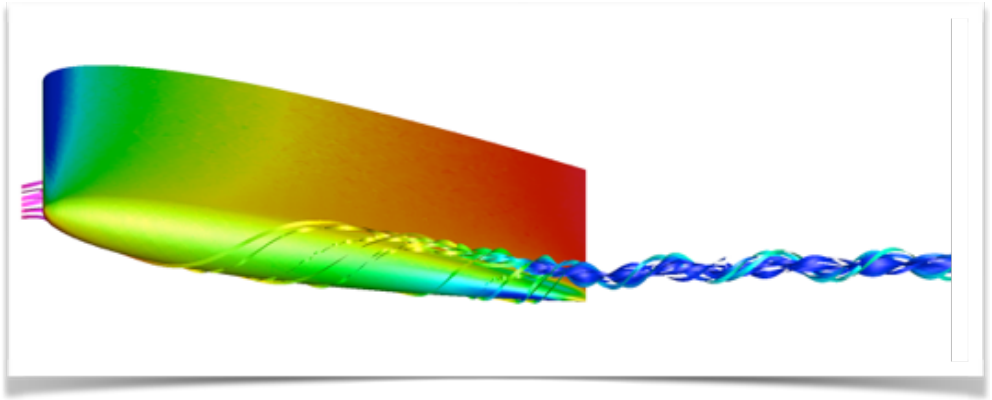
SCI INSTITUTE

Formula-1 Racing & Simulation

Nektar++: an open-source scalable C++ software library for high-order Finite Element methods.

See: www.nektar.info

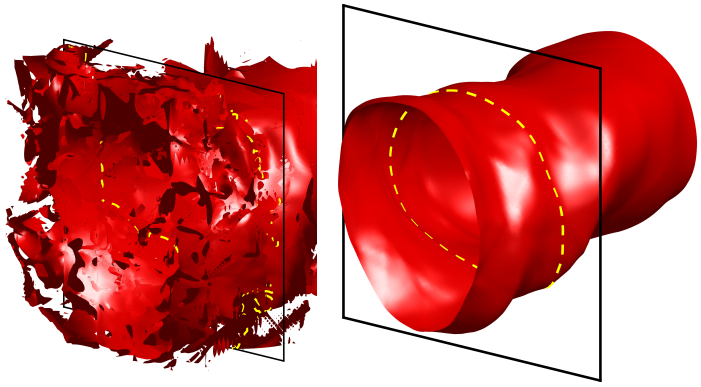




SCI INSTITUTE

Mike Kirby and Spencer Sherwin

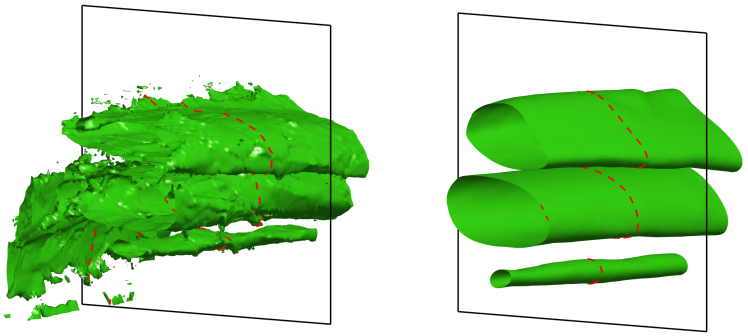
Filtering for High-Order FEM Visualization



FEM Vorticity

LSIAC Filtered Vorticity

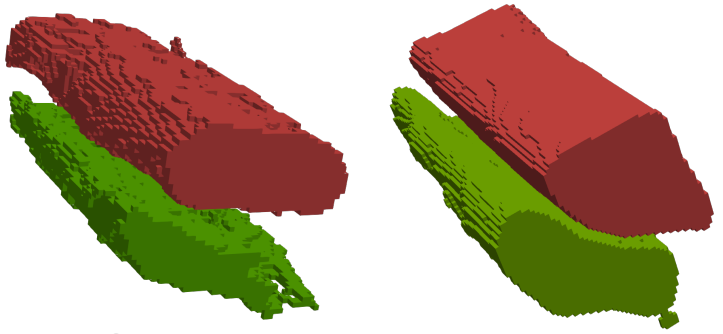
Counter-Rotating Vortex Vorticity



Sampled

LSIAC Filtered

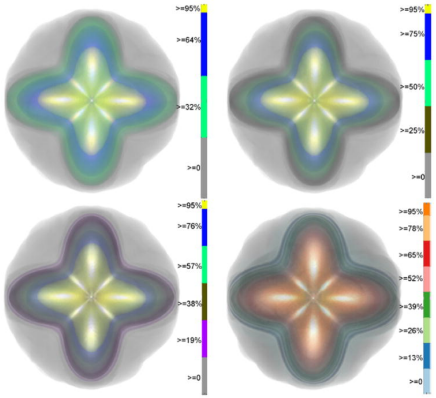
Topological Analysis of Fields

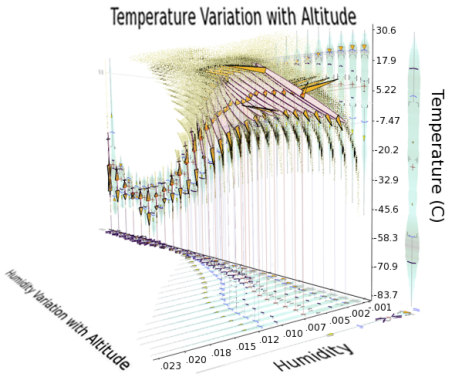


Sampled

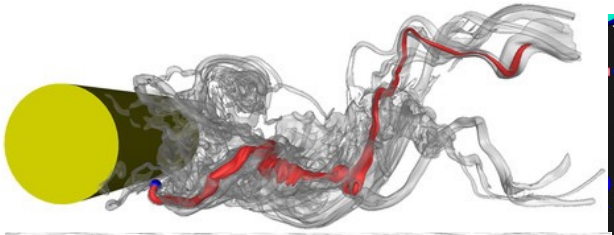
LSIAC Filtered

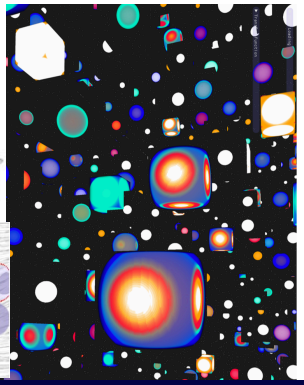
Uncertainty Visualization

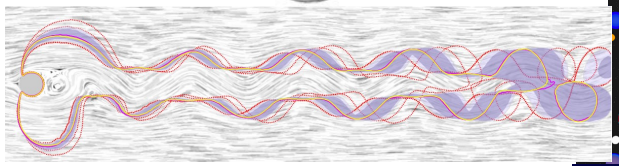


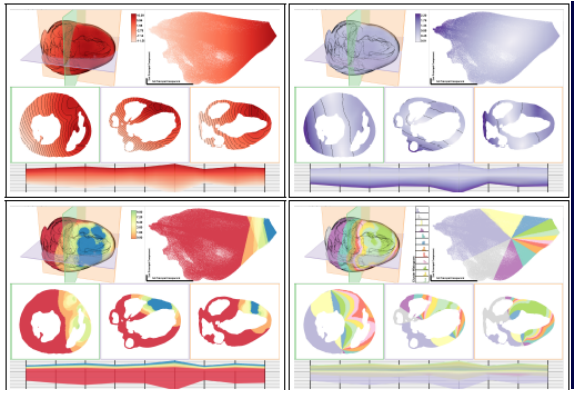


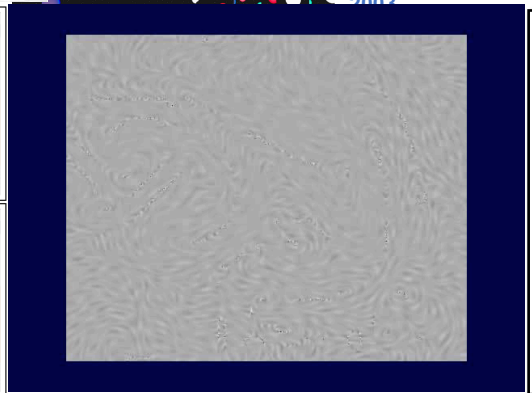
When is the last time you've seen an error bar on an isosurface?

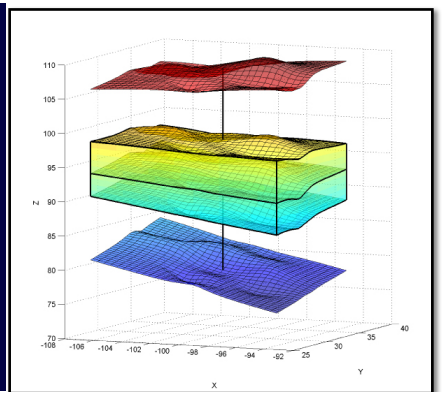












G.P. Bonneau, H.C. Hege, C.R. Johnson, M.M. Oliveira, K. Potter, P. Rheingans, T. Schultz. "Overview and State-of-the-Art of Uncertainty Visualization," In *Scientific Visualization: Uncertainty, Multifield, Biomedical, and Scalable Visualization*, Edited by M. Chen and H. Hagen and C.D. Hansen and C.R. Johnson and A. Kauffman, Springer-Verlag, pp. 3-27. 2014.

M.G. Genton, C.R. Johnson, K. Potter, G. Stenchikov, Y. Sun. "Surface boxplots," In *Stat Journal*, Vol. 3, No. 1, pp. 1-11. 2014.

K. Potter, P. Rosen, C.R. Johnson. "From Quantification to Visualization: A Taxonomy of Uncertainty Visualization Approaches," In *Uncertainty Quantification in Scientific Computing*, IFIP Series, Vol. 377, Springer, pp. 226-249. 2012.

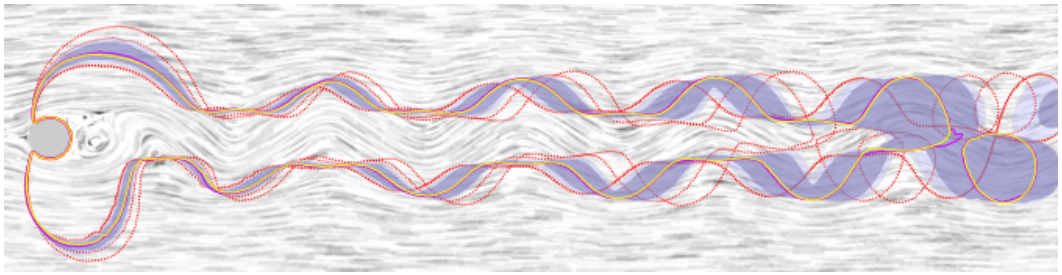
K. Potter, A. Wilson, P.-T. Bremer, D. Williams, C. Doutriaux, V. Pascucci, C.R. Johnson. "Ensemble-Vis: A Framework for the Statistical Visualization of Ensemble Data," In *Proceedings of the 2009 IEEE International Conference on Data Mining Workshops*, pp. 233-240. 2009.

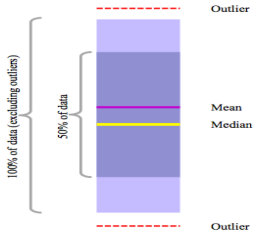
C.R. Johnson, A.R. Sanderson. "A Next Step: Visualizing Errors and Uncertainty," In *IEEE Computer Graphics and Applications*, Vol. 23, No. 5, pp. 6-10. September/October, 2003.

SCI INSTITUTE

Contour Box Plots

$$S \in \mathrm{sB}\left(S_1, \ldots S_j\right) \iff \bigcap_{k=1}^{j} S_k \subset S \subset \bigcup_{k=1}^{j} S_k.$$



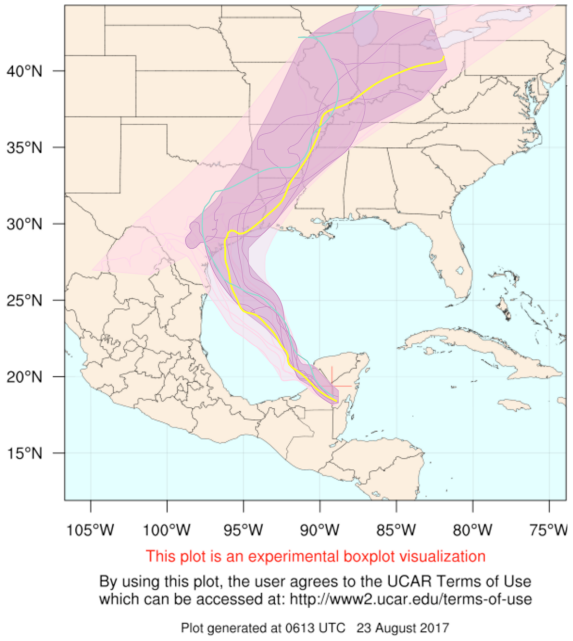


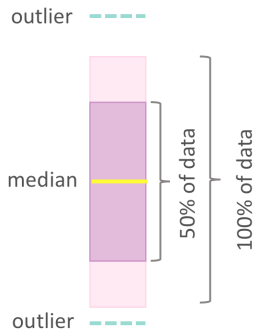
Whitaker, Mirzargar, Kirby, *IEEE Transactions on Visualization and Computer Graphics*, Vol. 19, No. 12, pp. 2713--2722, 2013.

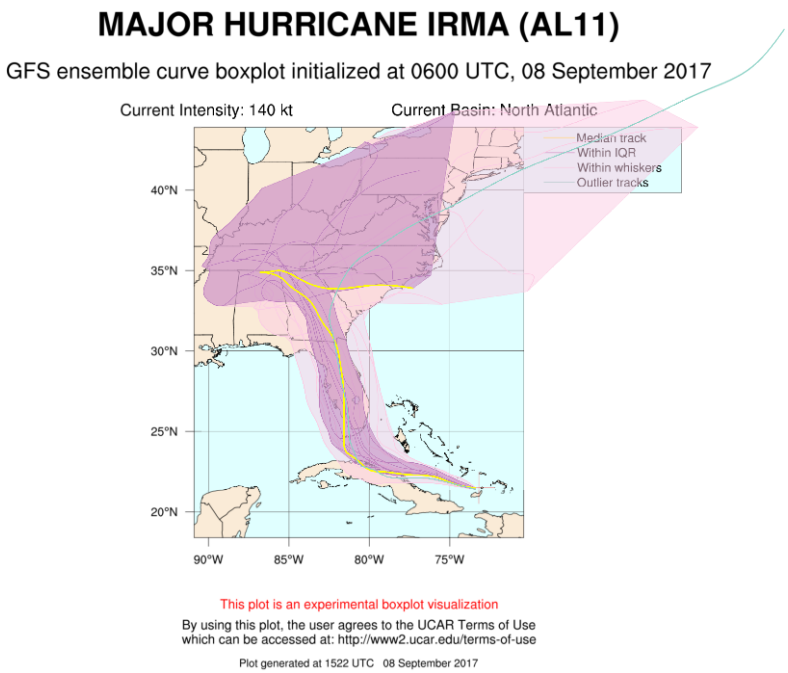
M.G. Genton, C.R. Johnson, K. Potter, G. Stenchikov, Y. Sun. "Surface boxplots," In *Stat Journal*, Vol. 3, No. 1, pp. 1-11. 2014.

SCI INSTITUTE

Ensemble Curved Boxplot







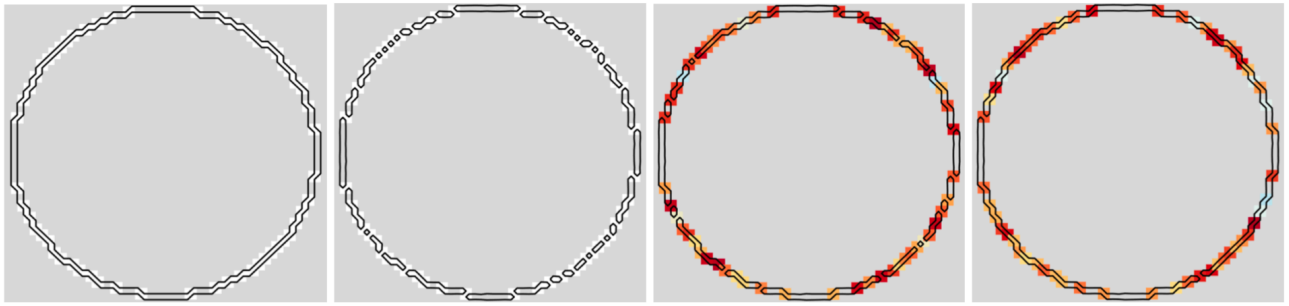
NCAR

M. Mirzargar, R. Whitaker, R. M. Kirby. "Curve Boxplot: Generalization of Boxplot for Ensembles of Curves," IEEE Transactions on Visualization and Computer Graphics, Vol. 20, No. 12, IEEE, pp. 2654-63. December, 2014.

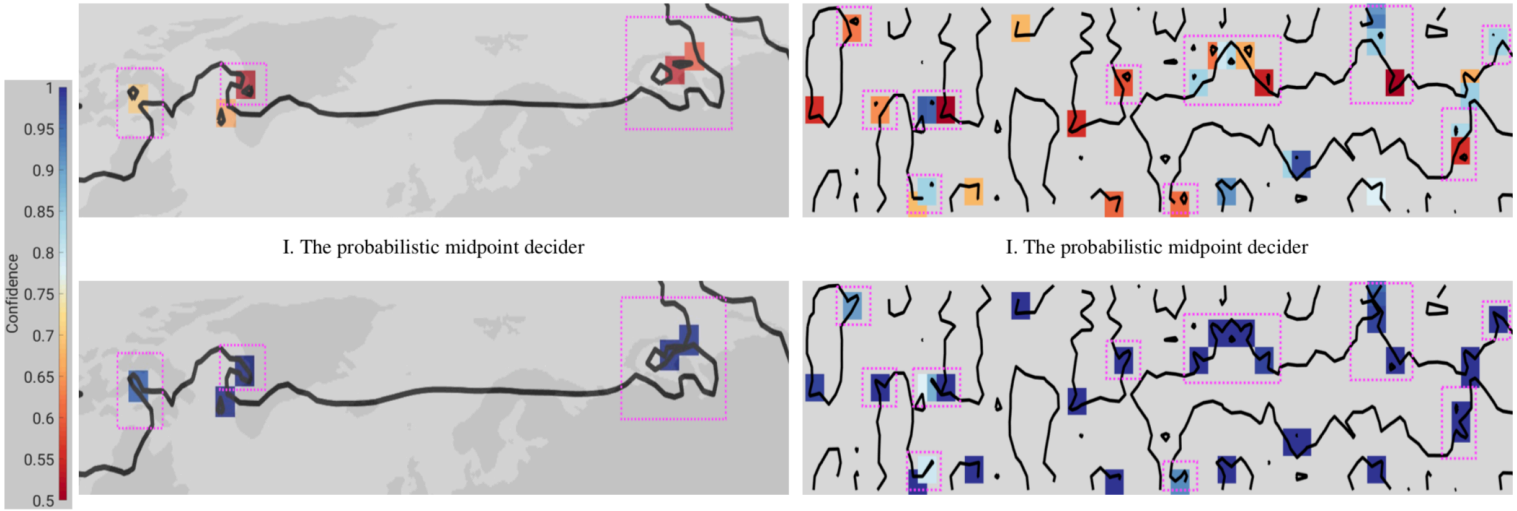
SCI INSTITUTE

Probabilistic Asymptotic Decider for Topological Ambiguity Resolution in Level-Set Extraction for Uncertain 2D Data

Tushar Athawale and Chris R. Johnson



(a) The isocontour topology in the (b) The asymptotic decider in the mean (c) The probabilistic midpoint decider (d) The probabilistic asymptotic decider



I. The probabilistic midpoint decider

I. The probabilistic midpoint decider

II. The probabilistic asymptotic decider

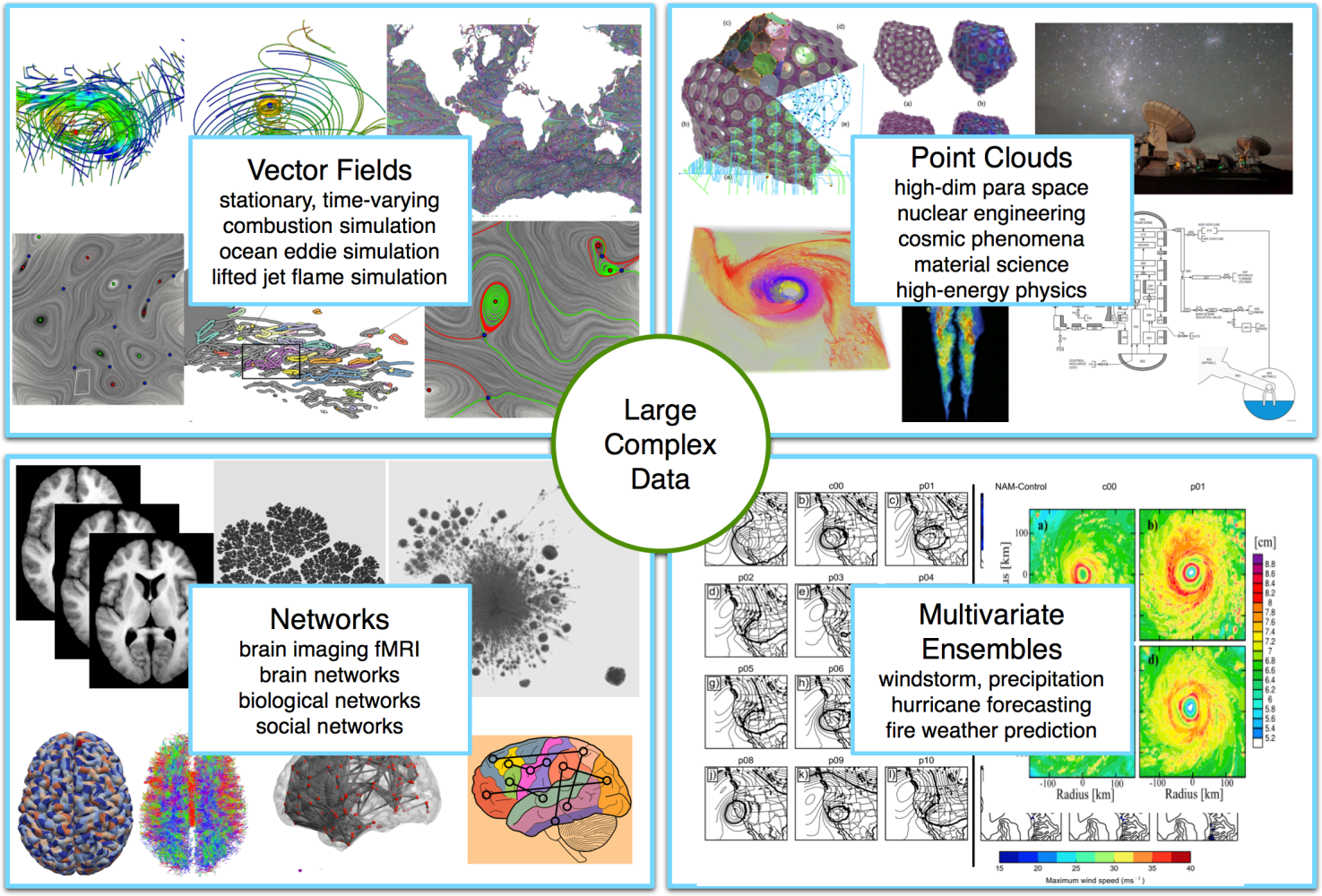
II. The probabilistic asymptotic decider

(a) The temperature field

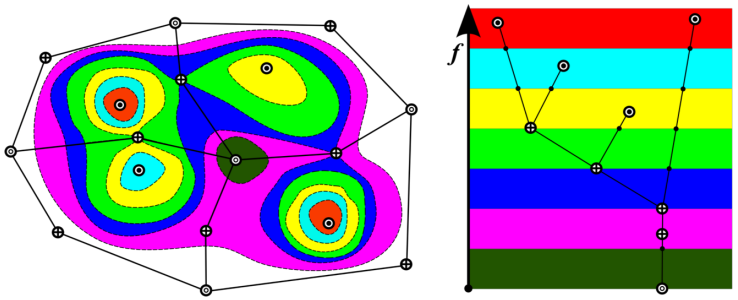
(b) The velocity field for the Kàrmàn vortex street

SCI INSTITUTE

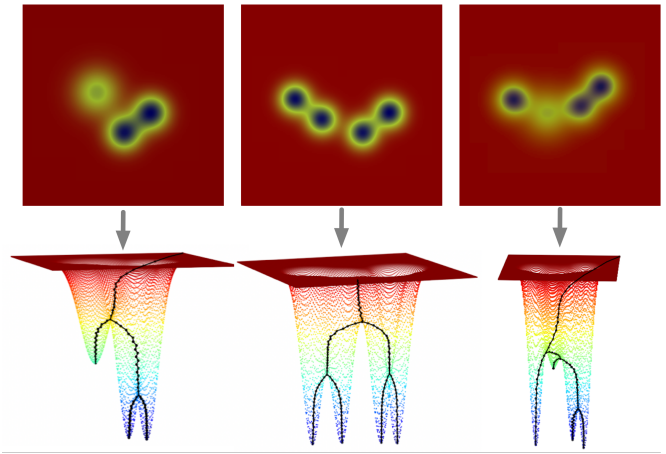
Topological Data Analysis and Visualization



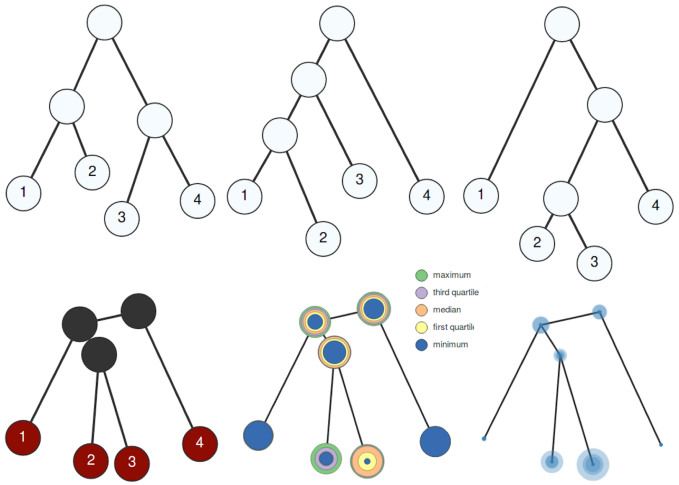
Visualizing uncertainty in topological structures



Merge Tree: a topological summary of scalar fields



Merge trees that arise from an ensemble of scalar fields

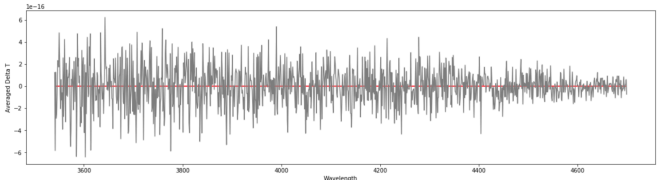


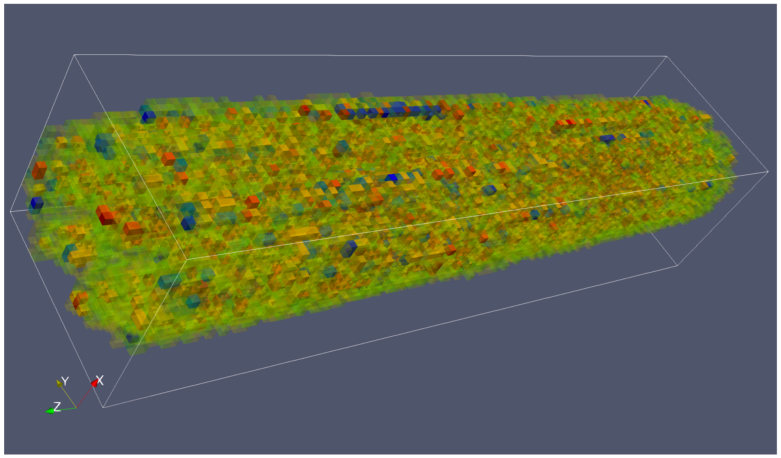
1. Compute an average merge tree from an ensemble
2. Uncertainty visualization of the average tree captures structural variations among the ensembles

Lin Yan, Yusu Wang, Elizabeth Munch, Ellen Gasparovic, **Bei Wang**.

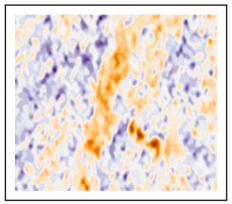
SCI INSTITUTE

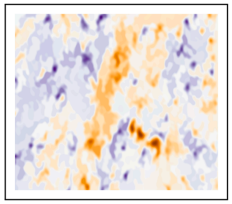
Topological Data Analysis for Astronomical Data Cubes

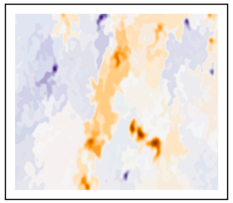


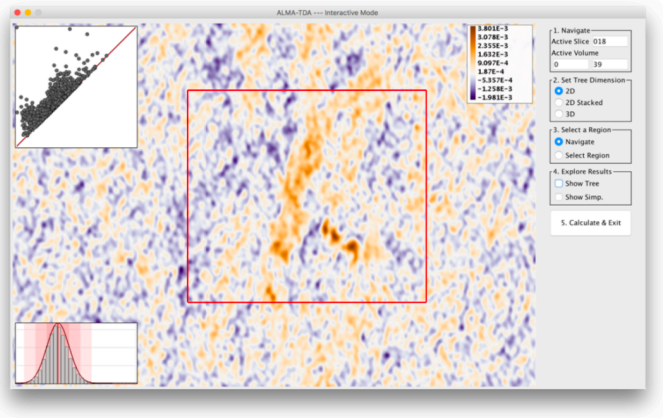


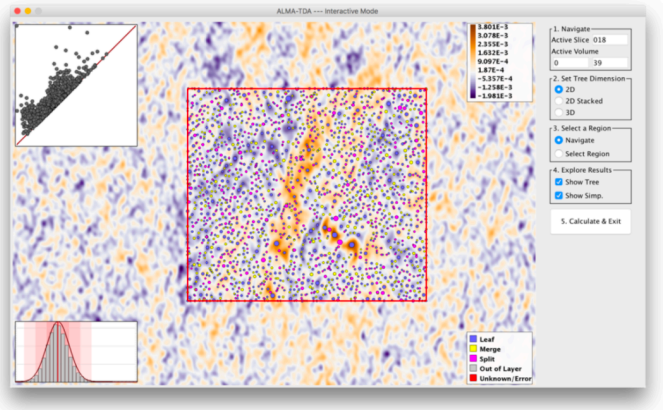
Analysis of cosmic voids







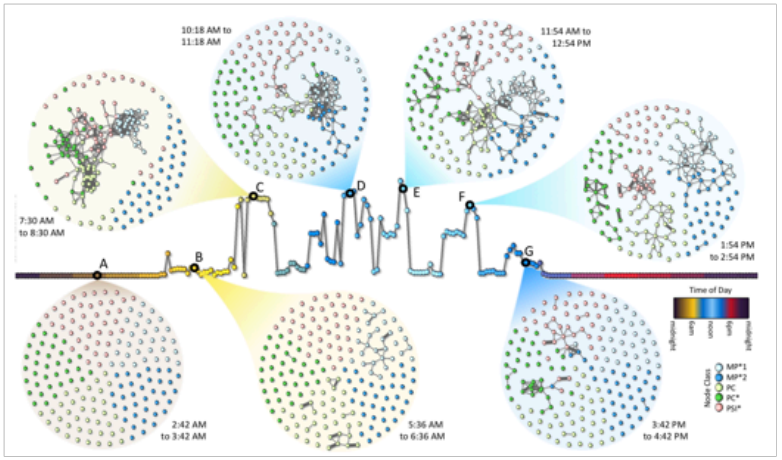


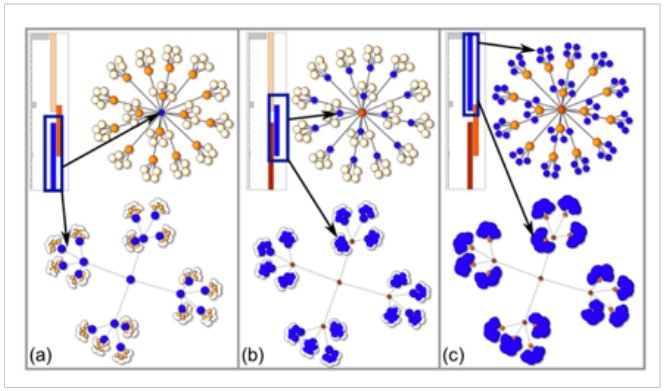


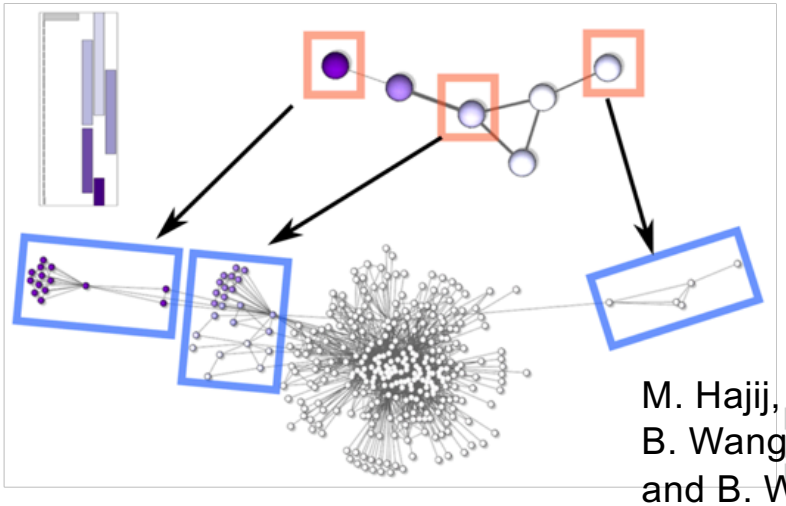
Using Contour Trees in the Analysis and Visualization of Radio Astronomy Data Cubes

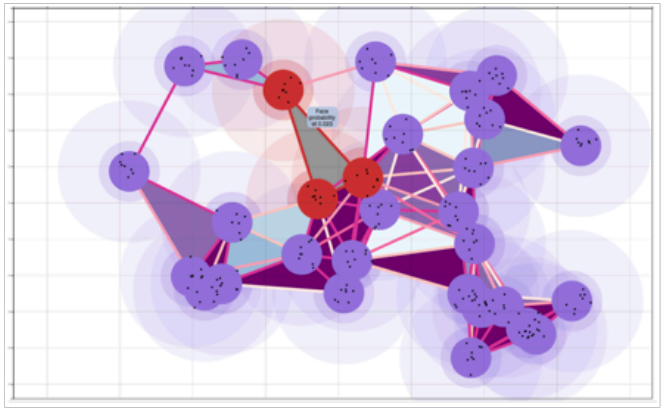
Yulong Liang, Vikranta Kamble, Helion Dumas Desbourboux, Lin Yan, Mengjiao Han, Kyle Dawson, Nicholas Boardman, Gail Zasowski, Anil Seth, Joel Brownstein, Paul Rosen, Juna A. Kollmeier, Guillermo Blanc, **Bei Wang**

Topological Analysis of Time Dependent Networks









M. Hajij, B. Wang, C. Scheidegger, P. Rosen, 2018; M. Hajij, B. Wang, P. Rosen, 2018; T. Sodergren, J. Hair, J. M. Phillips and B. Wang, 2017.

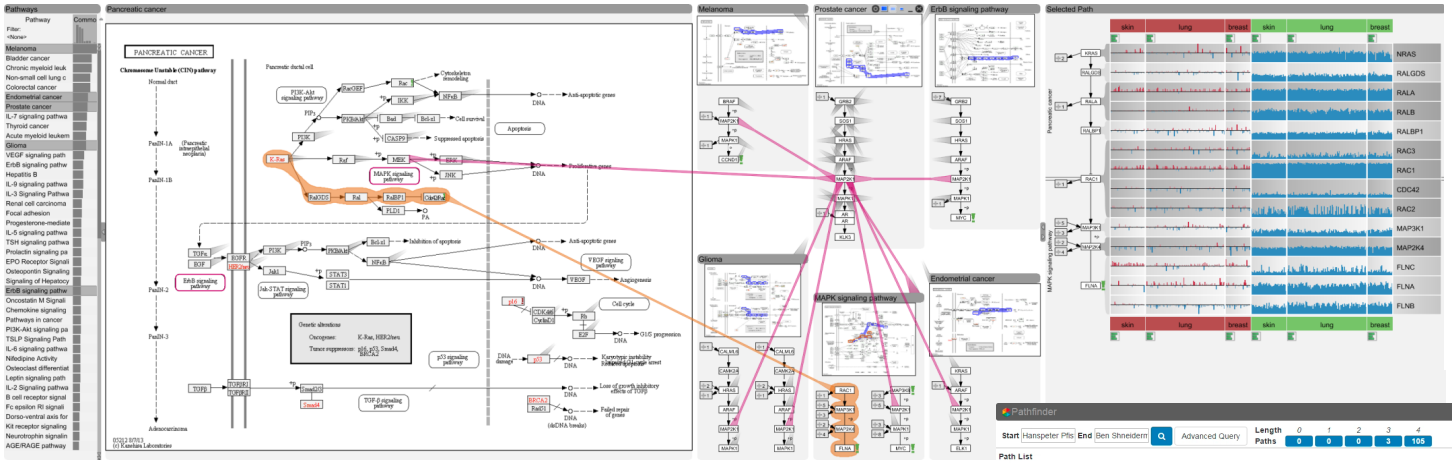
Large-Scale Visualization Challenges

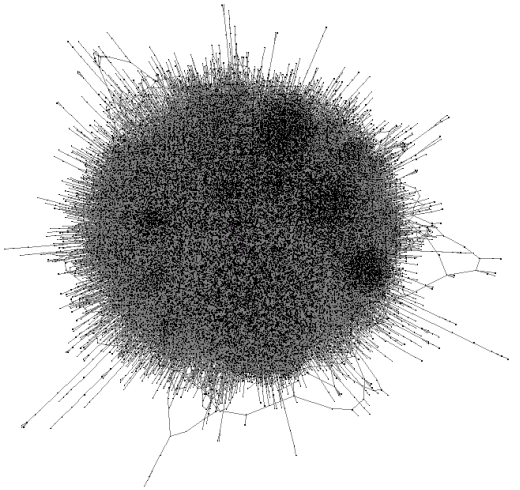
Scalability

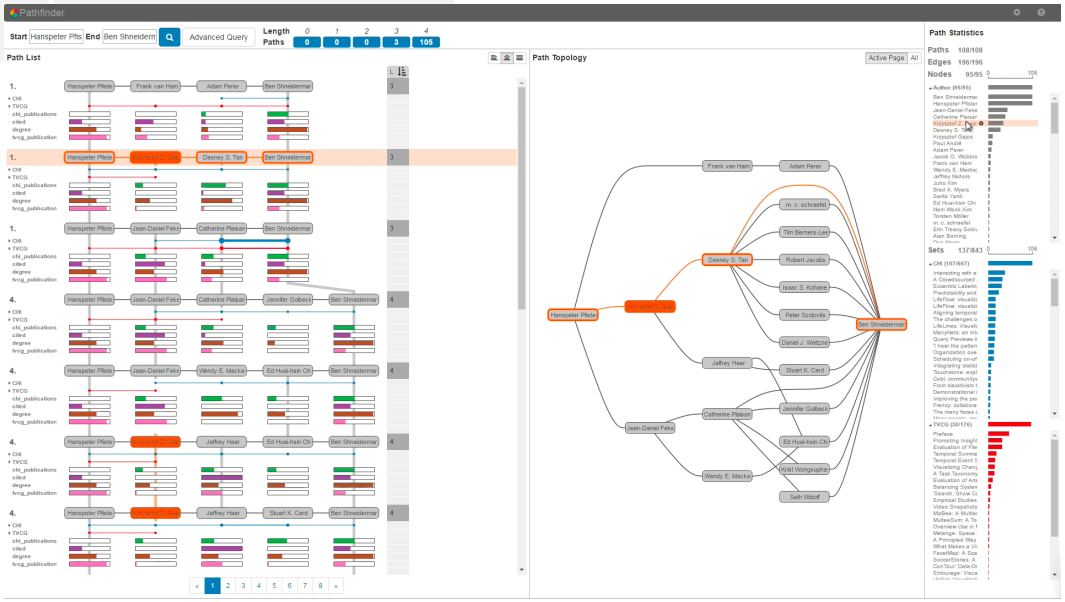
- High Dimensional Visualization
- In Situ and Streaming Visualization
- Hierarchy

SCI INSTITUTE

Large Multivariate Networks

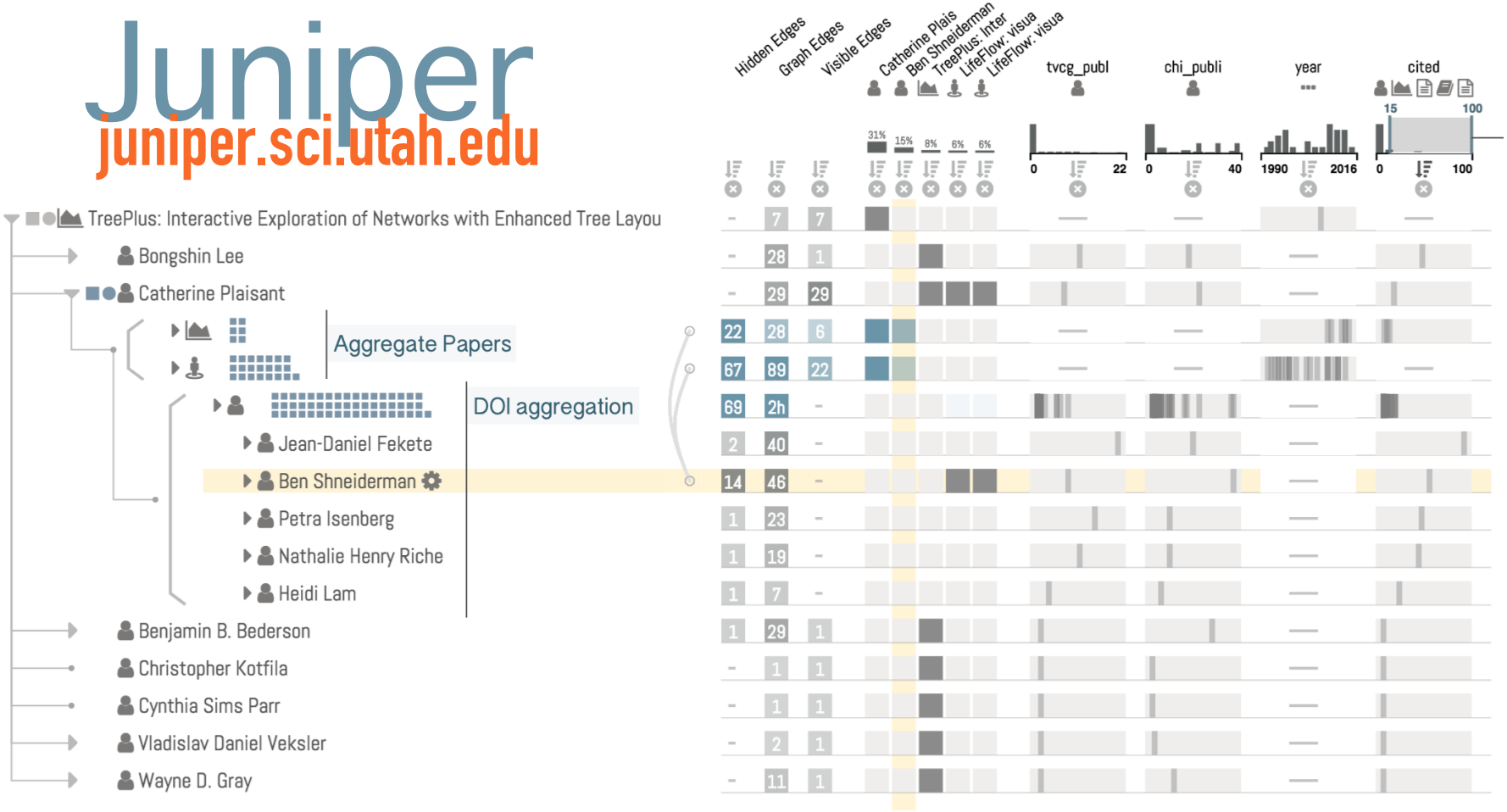




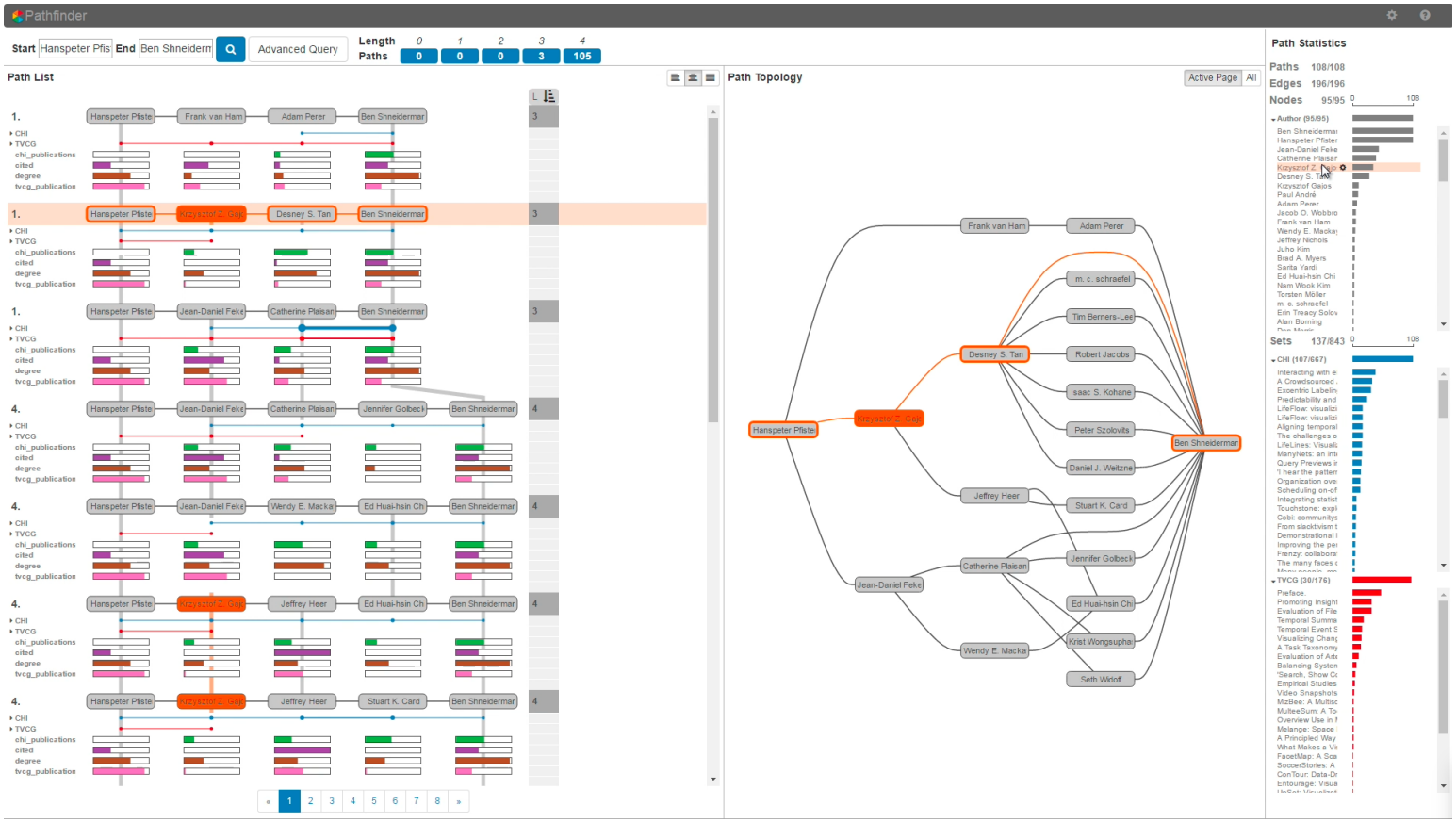


Juniper

juniper.sci.utah.edu



Pathfinder: Visual Analysis of Paths in Graphs

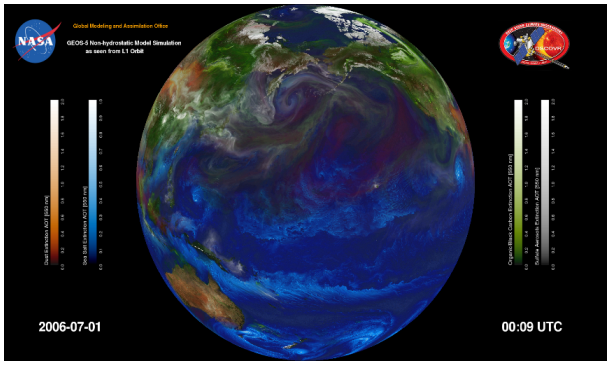


C. Partl, S. Gratzl, M. Streit, A. Wassermann, H. Pfister, D. Schmalstieg, A. Lex. **"Pathfinder: Visual Analysis of Paths in Graphs,"** In *Computer Graphics Forum (EuroVis '16)*, Vol. 35, No. 3, pp. 71-80, 2016.

Scalable Deployment: Exploration of 3.5PB of NASA Weather/Climate Data in Real Time

Workflow

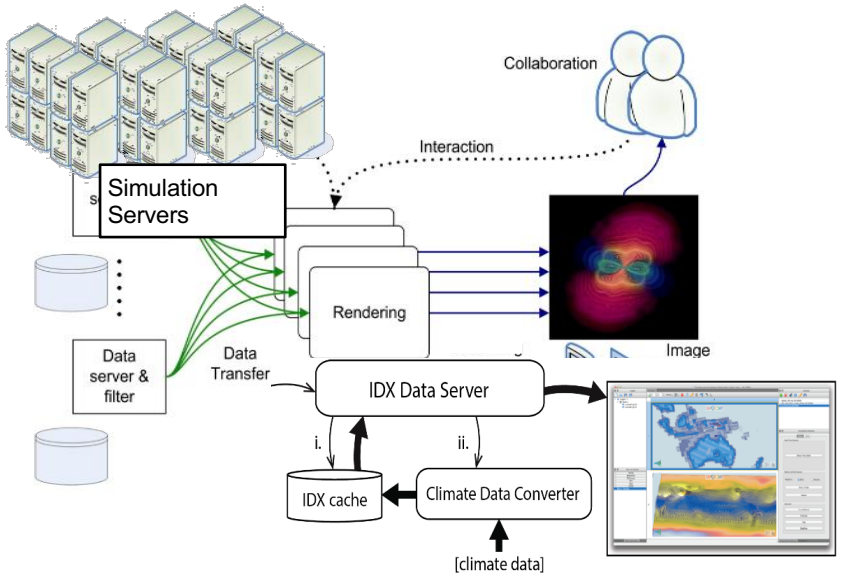
- *Data creation*
- *Data Management*
  - Processing
  - Analysis
  - Visualization

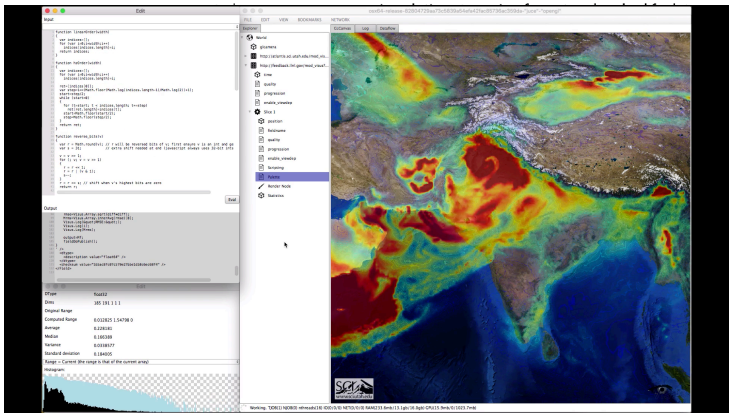


- 7km GEOS-5 “Nature Run”
- 1 dataset, 3.5 PB
- theoretically: openly accessible
- practically: precomputed pics

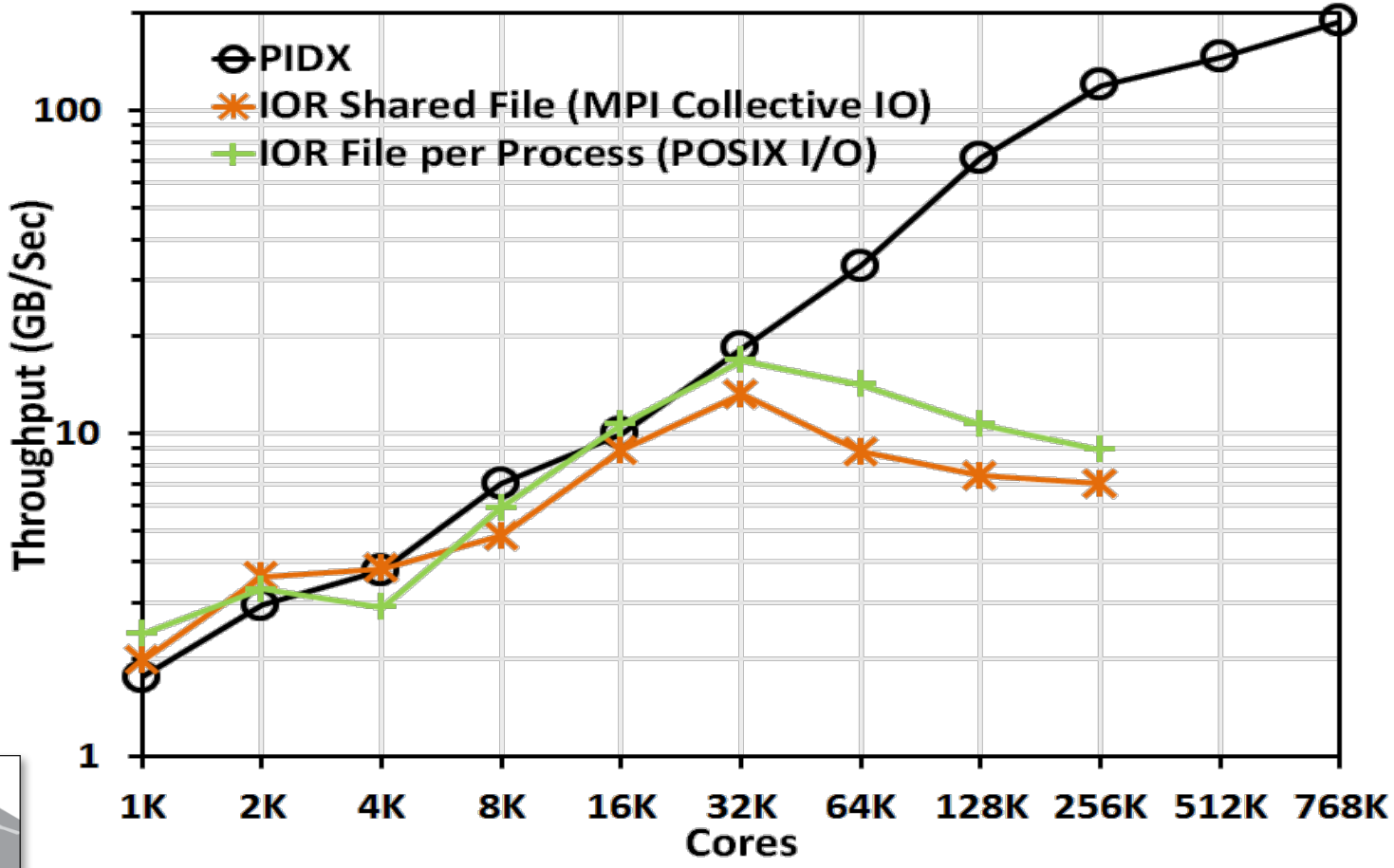
Distributed Resources

- 3.5 PB of *data store in NASA*
- *Primary ViSUS server in LLNL*
- *Secondary ViSUS server in Utah*
- *Clients connect remotely*
- *Work without additional HPC resources*





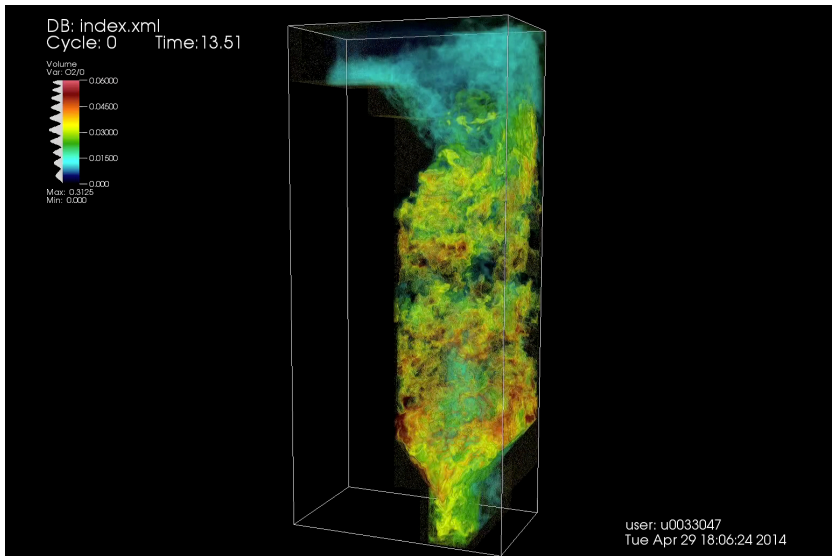
High Performance Data Movement for Real-Time Monitoring of Large Scale Simulations

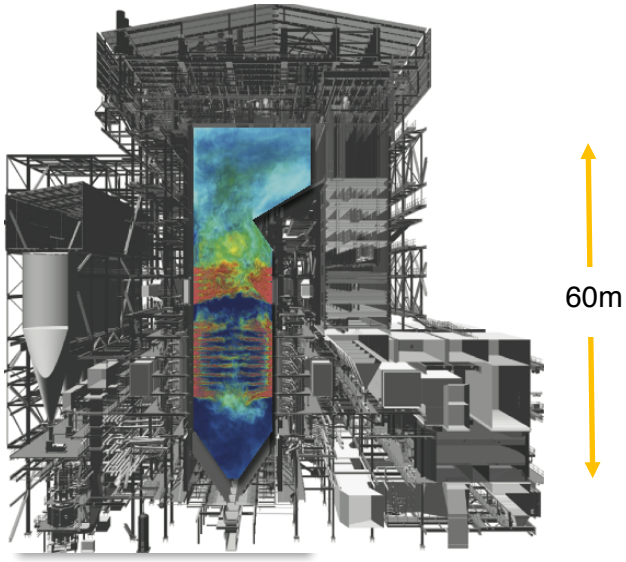


SCI INSTITUTE

DOE PSAAP2 Simulations of GE Clean(er) Coal Boilers

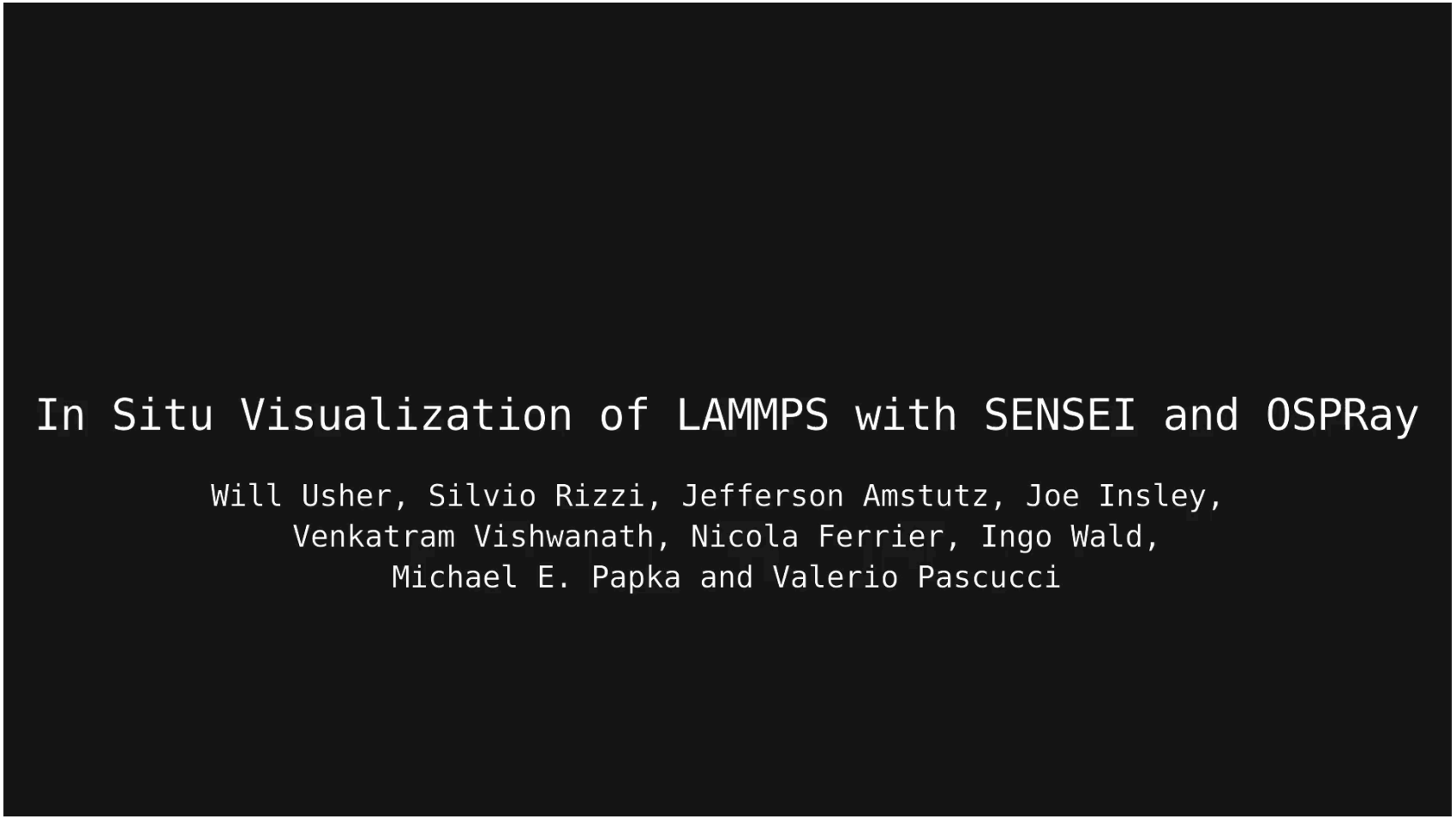
- Large scale turbulent combustion needs mm scale grids
  10^14 mesh cells 10^15 variables (1000x more than now)
- Structured, high order finite-volume discretization
- Mass, momentum, energy conservation
- LES closure, tabulated chemistry
- PDF mixing models
- DQMOM (many small linear solves)
- Uncertainty quantification





- Low Mach number approx. (pressure Poisson solve up to 10^12 variables. 1M patches 10 B variables
- Radiation via Discrete Ordinates – many hypre solves Mira (cpus) or ray tracing Titan (gpus strong and weak scaling via AMR).
- FAST I/O needed PIDX for scalability

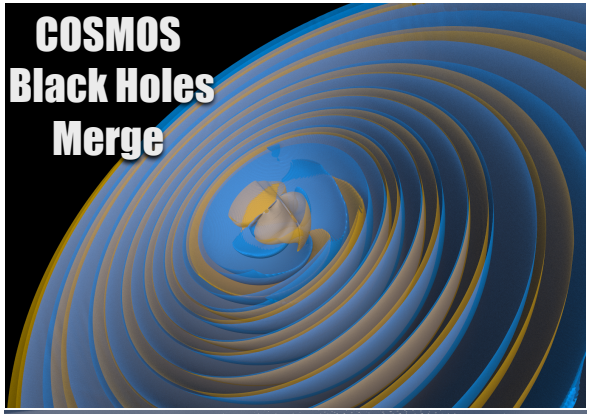
LAMMPS In-Situ Rendering with OSPRay



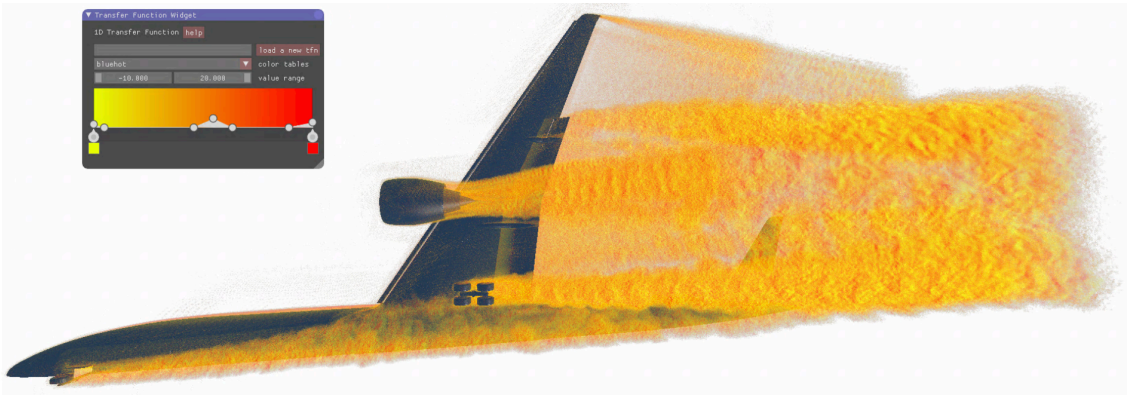
SCI INSTITUTE

AMR Data Visualization

- Block-structured AMR
- Tree-based (unstructured) AMR



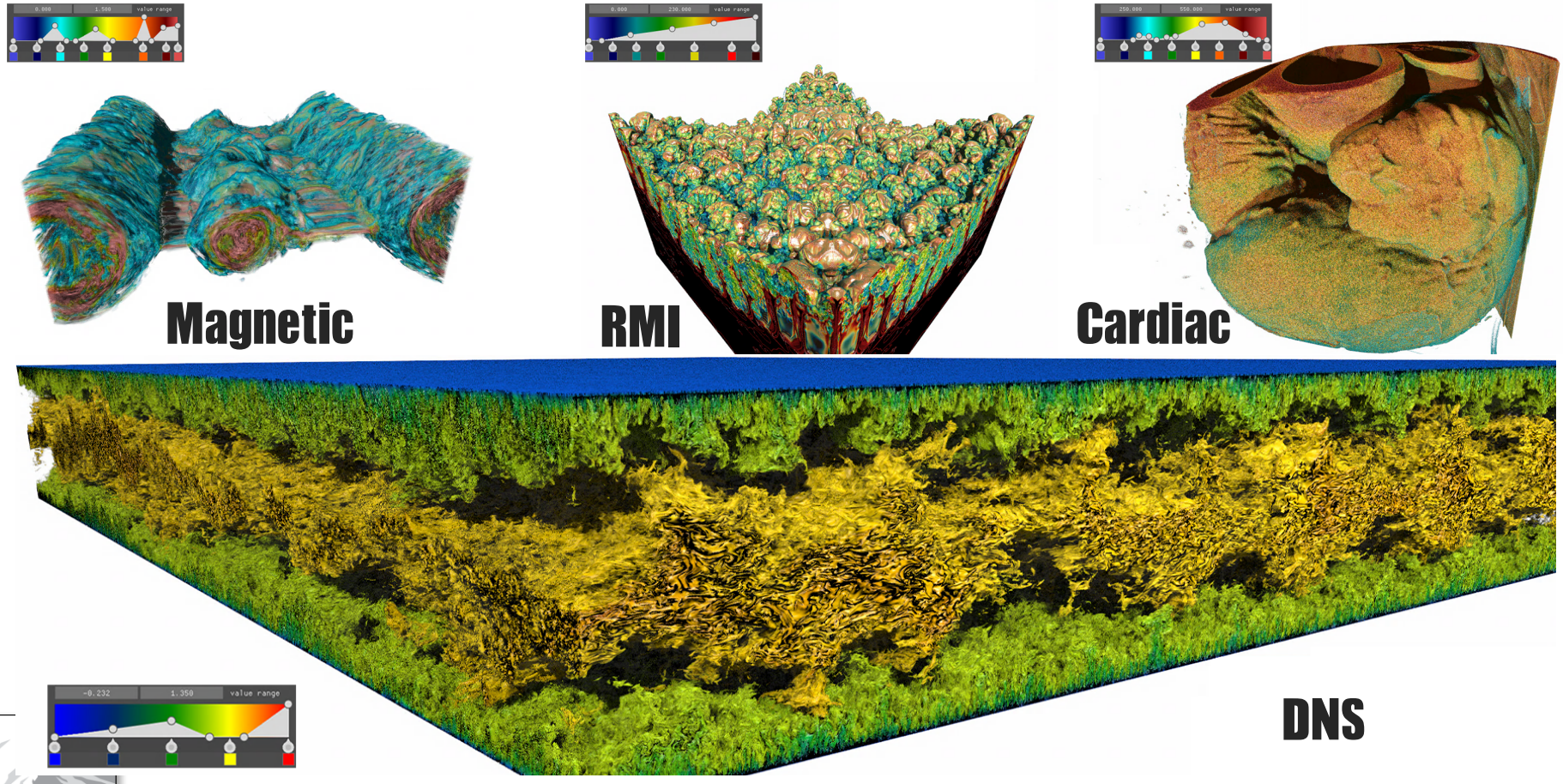
NASA LandingGear



NASA Exajet

SCI INSTITUTE

Interactive Visualization of Large-Scale Datasets

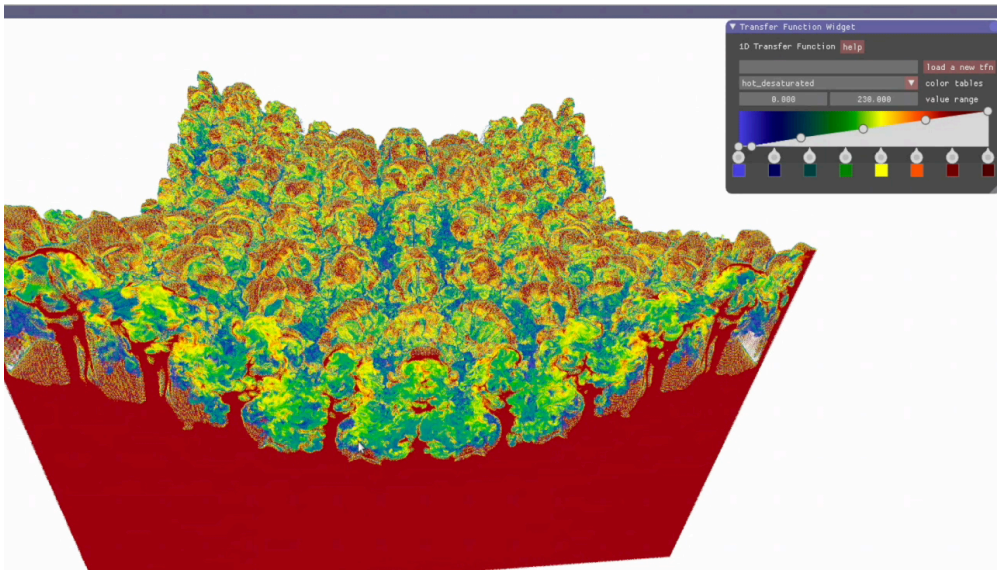


SCI INSTITUTE

Ray-guided Progressive Rendering

Progressive sampling

- Hierarchical representation
- On-demand loading
- Independent data-streaming threads
- Visualize coarse data as a approximate and gradually refine it



SCI INSTITUTE

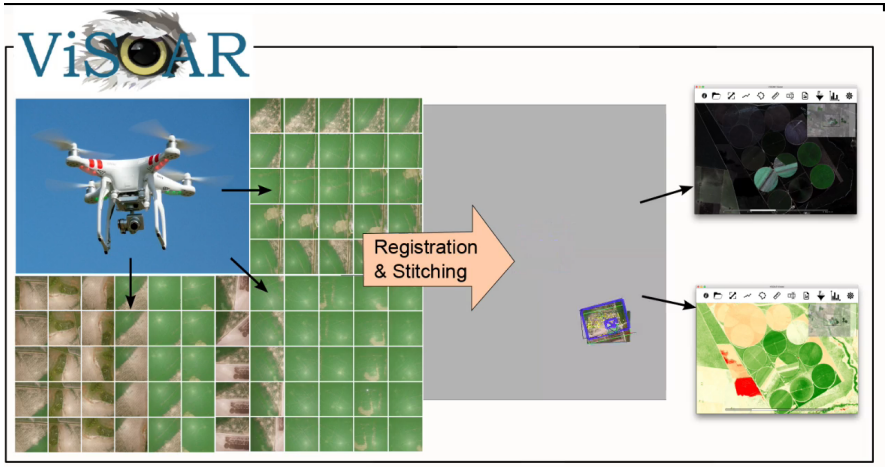
Large-Scale Visualization Challenges

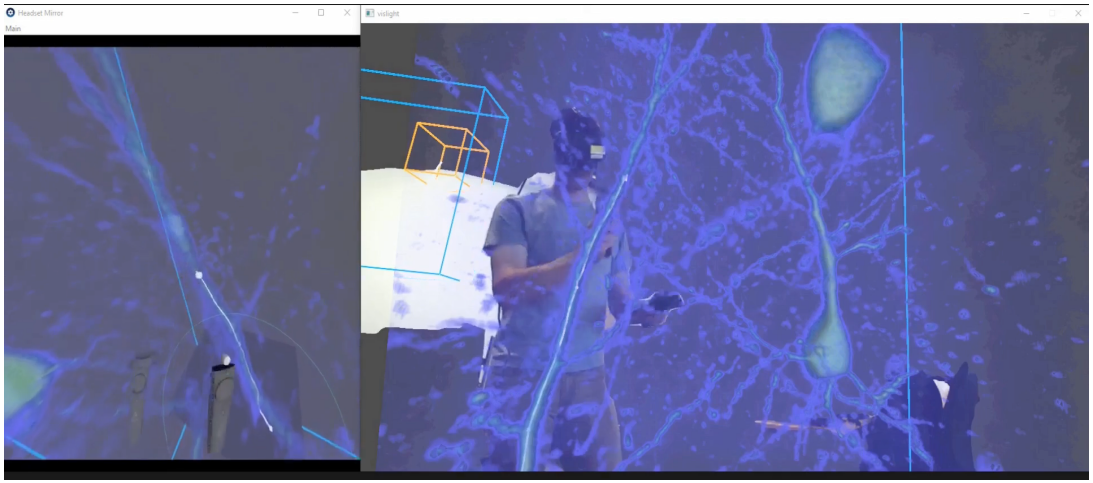
Users and Usability

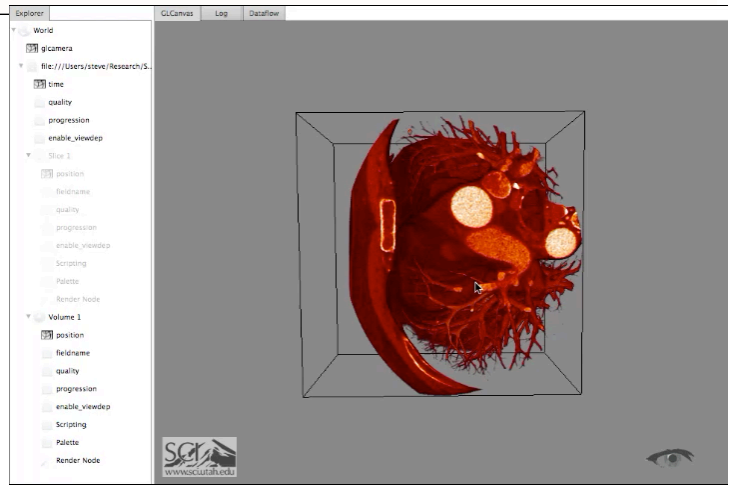
- Interactivity
- Collaboration
- Design
- Reproducible Visualization

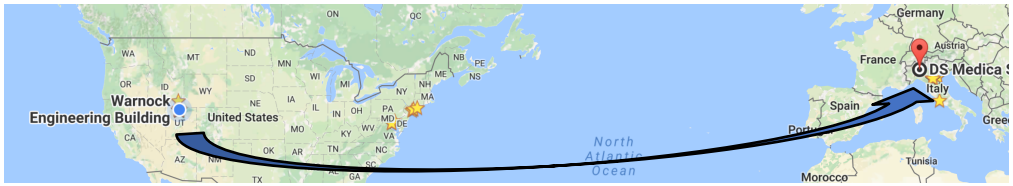
Interactive Streamline Exploration and Manipulation using Deformation

Xin Tong[1], John Edwards[2], Chun-Ming Chen[1],
Han-Wei Shen[1], Chris R. Johnson[2], Pak Chung Wong[3]
[1]The Ohio State University
[2]Scientific Computing and Imaging Institute, University of Utah
[3]Pacific Northwest National Laboratory


ViSOAR
Registration
& Stitching


Headset Mirror
Main
vrlight

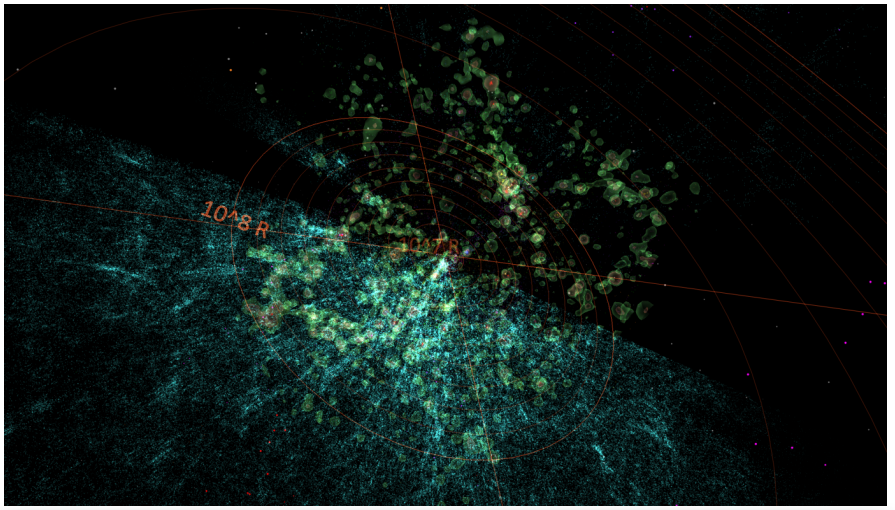

Explorer
GLCanvas
Log
Dataflow
World
glcamera
time
quality
progression
enable_viewdep
Volume 1
position
fieldname
Scripting
Palette
Render Node

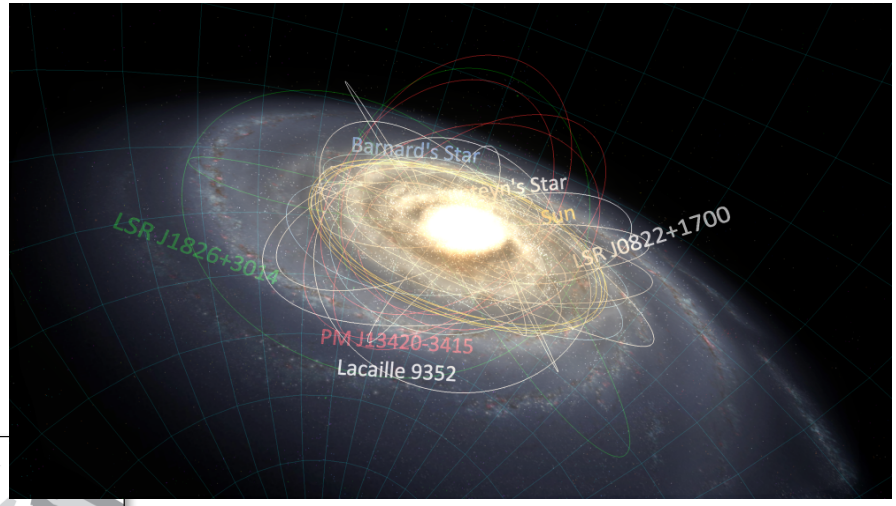

Warnock
Engineering Building
United States
North
Atlantic
Ocean
Germany
Austria
France
Italy
Spain
Tunisia
Morocco

OpenSpace

Platform for:

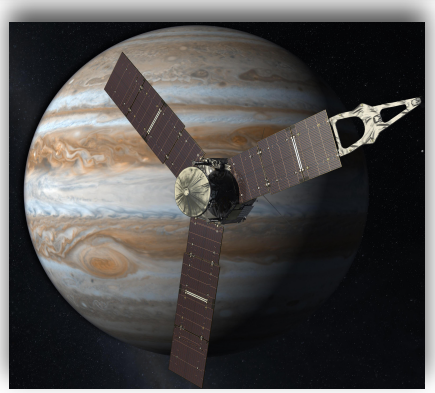
- **Visualization Research**
- **Space & Astro Research**
- **Science Communication**





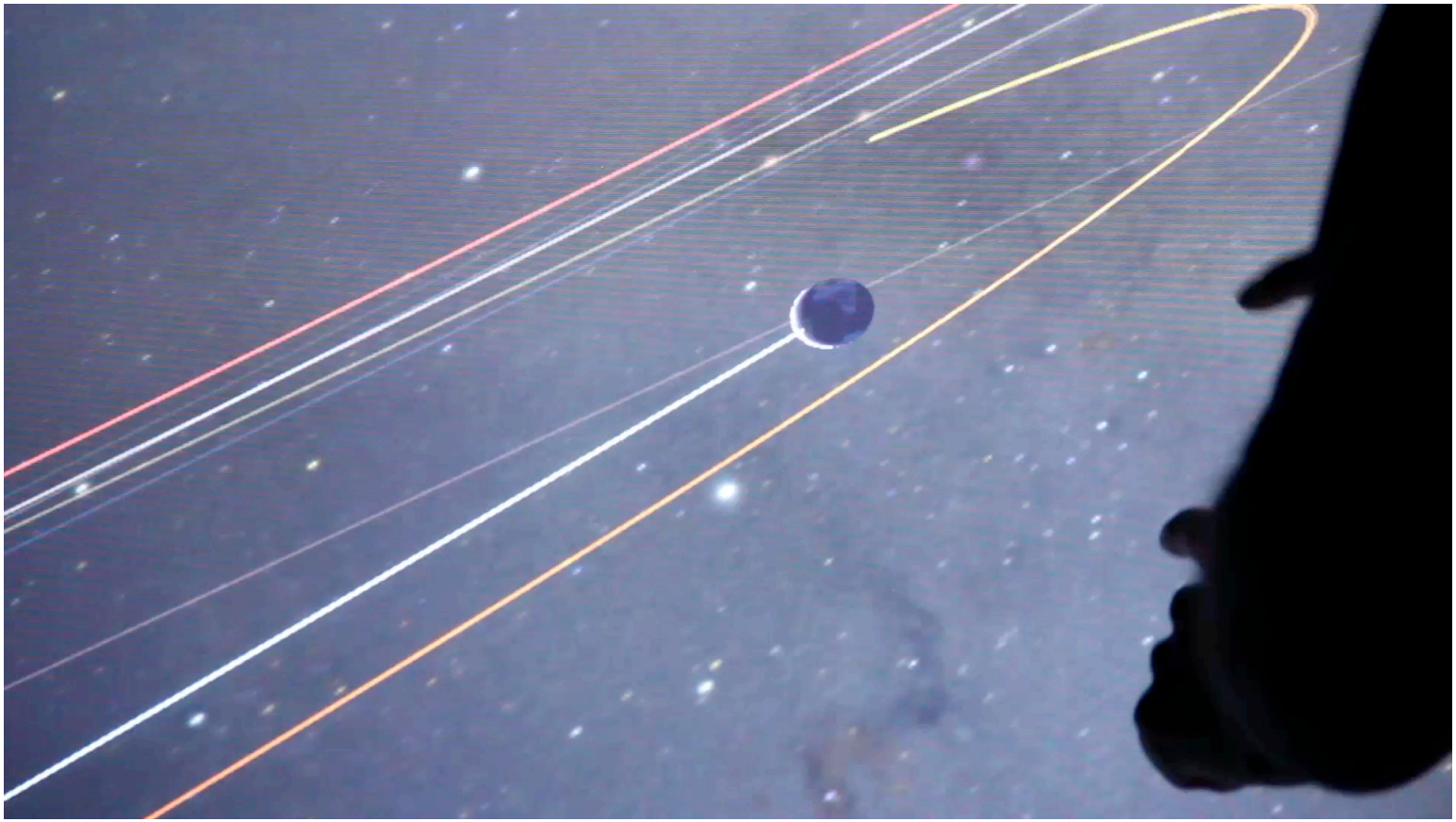
SCI INSTITUTE

OpenSpace Team

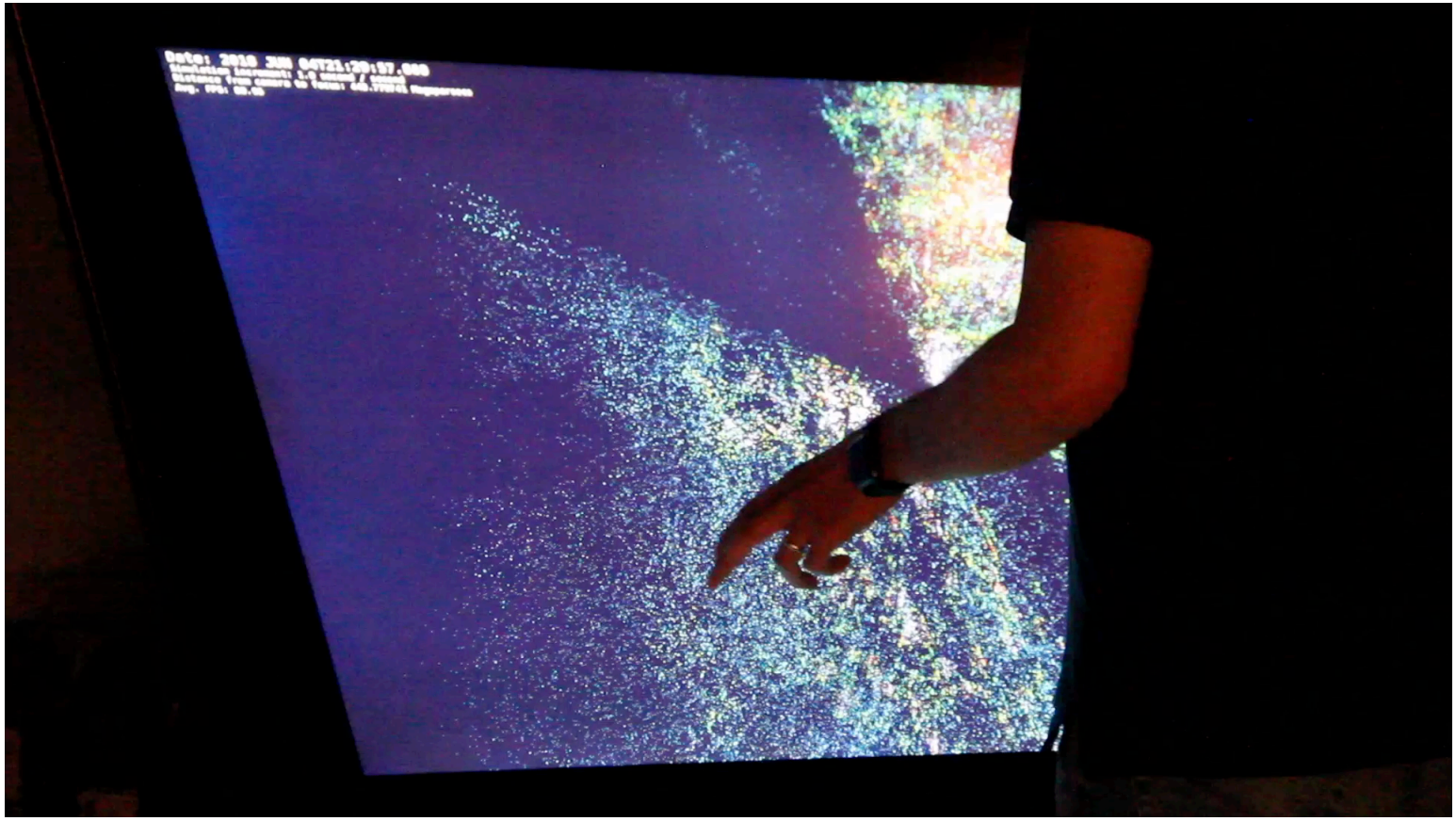


SCI INSTITUTE

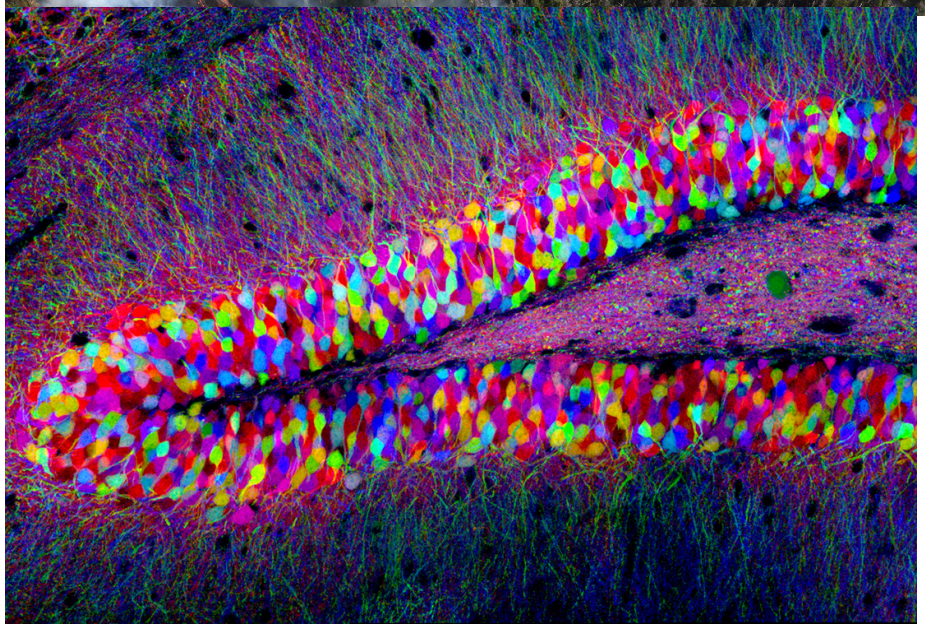
http://openspaceproject.com



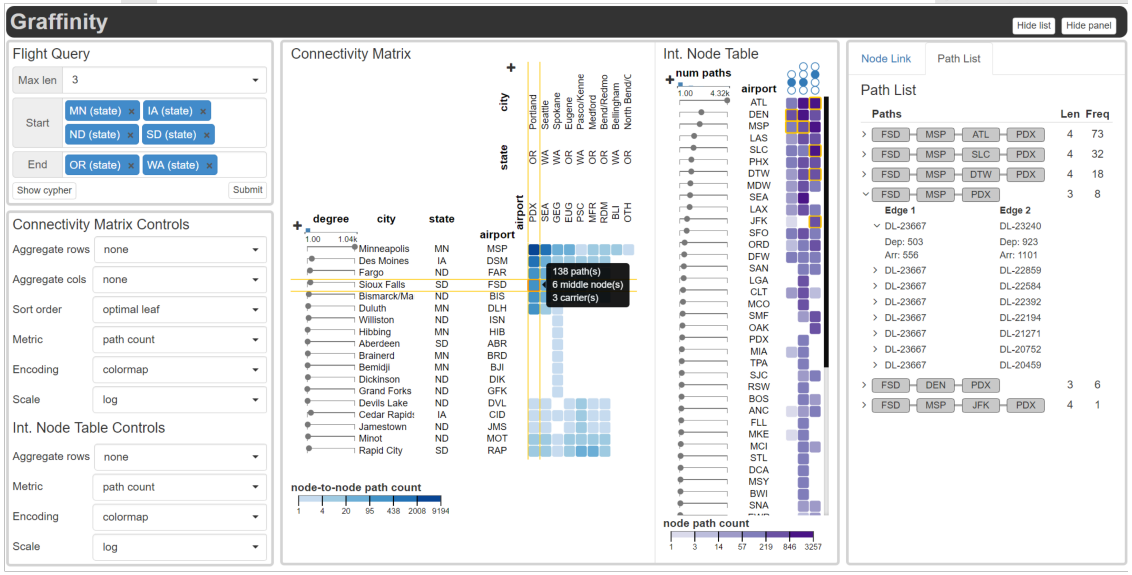
SCI INSTITUTE



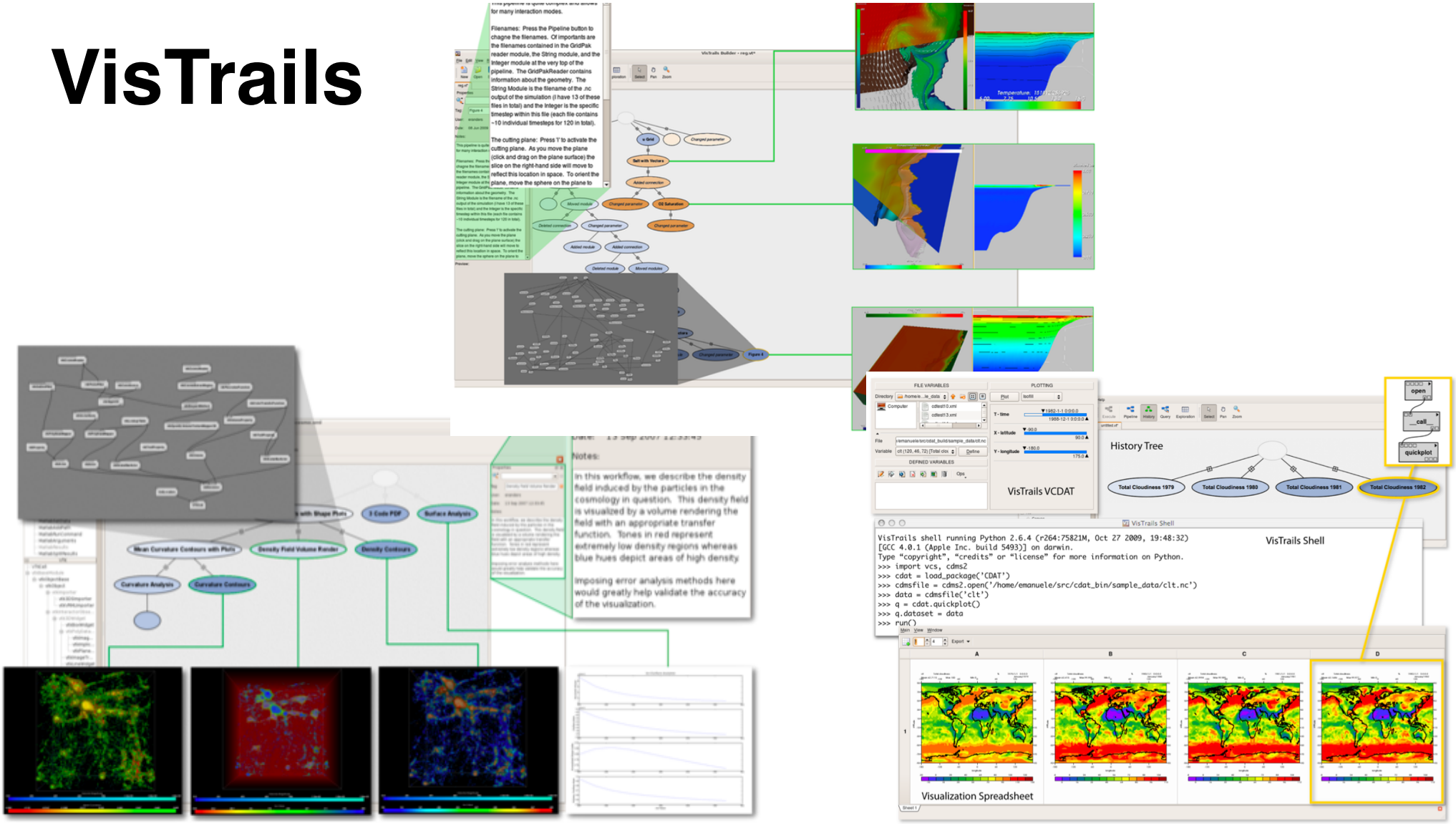
SCI INSTITUTE



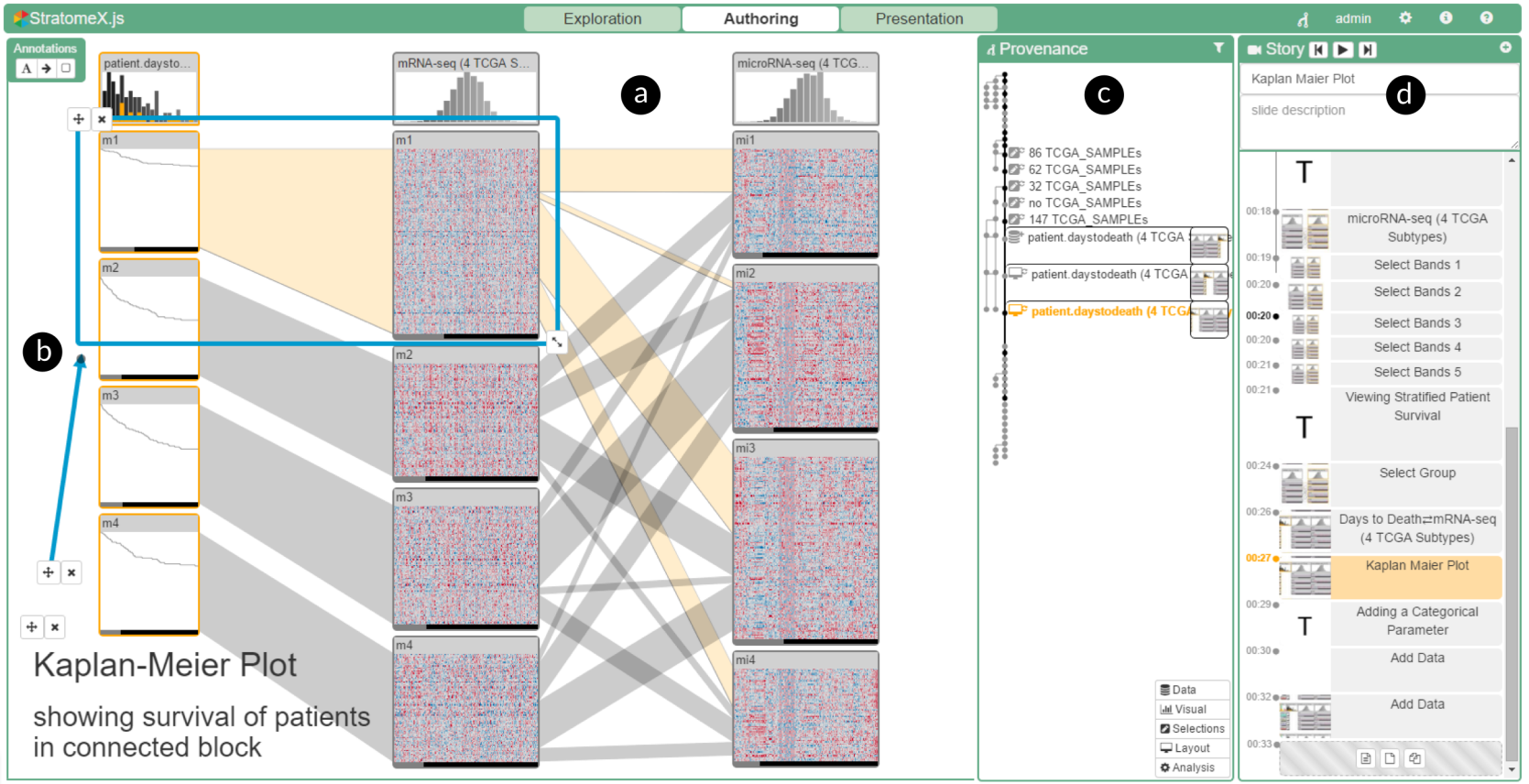
WeaVER
Deterministic
Stat Field
MNSD
Direct Ensemble
Probability


Graffinity
Hide list
Hide panel
Flight Query
Max len
3
Start
MN (state)
IA (state)
ND (state)
SD (state)
End
OR (state)
WA (state)
Show cypher
Submit
Connectivity Matrix Controls
Aggregate rows
none
Aggregate cols
none
Sort order
optimal leaf
Metric
path count
Encoding
colormap
Scale
log
Int. Node Table Controls
Aggregate rows
none
Metric
path count
Encoding
colormap
Scale
log
Connectivity Matrix
city
state
airport
degree
city
state
airport
Minneapolis MN MSP
Des Moines IA DSM
Fargo ND FAR
Sioux Falls SD FSD
Bismarck/Ma ND BIS
Duluth MN DLH
Williston ND ISN
Hibbing MN HIB
Aberdeen SD ABR
Brainerd MN BRD
Bemidji MN BJI
Dickinson ND DIK
Grand Forks ND GFK
Devils Lake ND DVL
Cedar Rapids IA CID
Jamestown ND JMS
Minot ND MOT
Rapid City SD RAP
138 path(s)
6 middle node(s)
3 carrier(s)
node-to-node path count
1 4 20 95 438 2008 9194
Int. Node Table
num paths
airport
ATL DEN MSP LAS SLC PHX DTW MDW SEA LAX JFK SFO ORD DFW SAN LGA CLT MCO SMF OAK PDX MIA TPA SJC RSW BOS ANC FLL MKE MCI STL DCA MSY BWI SNA
node path count
1 3 14 57 219 846 3257
Node Link
Path List
Path List
Paths
Len
Freq
FSD MSP ATL PDX 4 73
FSD MSP SLC PDX 4 32
FSD MSP DTW PDX 4 18
FSD MSP PDX 3 8
Edge 1
Edge 2
DL-23667 DL-23240
Dep: 503 Arr: 556
Dep: 923 Arr: 1101
DL-23667 DL-22859
DL-23667 DL-22584
DL-23667 DL-22392
DL-23667 DL-22194
DL-23667 DL-21271
DL-23667 DL-20752
DL-23667 DL-20459
FSD DEN PDX 3 6
FSD MSP JFK PDX 4 1

VisTrails


Notes:
In this workflow, we describe the density field induced by the particles in the cosmology in question. This density field is visualized by a volume rendering the field with an appropriate transfer function. Tones in red represent extremely low density regions whereas blue hues depict areas of high density.
Imposing error analysis methods here would greatly help validate the accuracy of the visualization.
VisTrails VCDAT
History Tree
Total Cloudiness 1979
Total Cloudiness 1980
Total Cloudiness 1981
Total Cloudiness 1982
open
__call__
quickplot
VisTrails Shell
VisTrails shell running Python 2.6.4 (r264:75821M, Oct 27 2009, 19:48:32)
[GCC 4.0.1 (Apple Inc. build 5493)] on darwin.
Type "copyright", "credits" or "license" for more information on Python.
>>> import vcs, cdms2
>>> cdat = load_package('CDAT')
>>> cdmsfile = cdms2.open('/home/emanuele/src/cdat_bin/sample_data/clt.nc')
>>> data = cdmsfile('clt')
>>> q = cdat.quickplot()
>>> q.dataset = data
>>> run()
Visualization Spreadsheet

Reproducibility in Computational Workflows



SCI INSTITUTE

SCI INSTITUTE

Acknowledgments

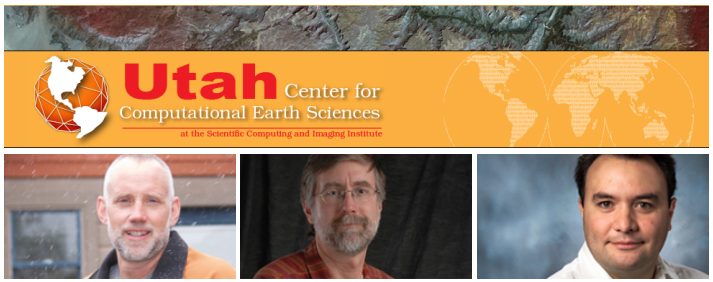
NIH/NIGMS Center for Integrative Biomedical Computing

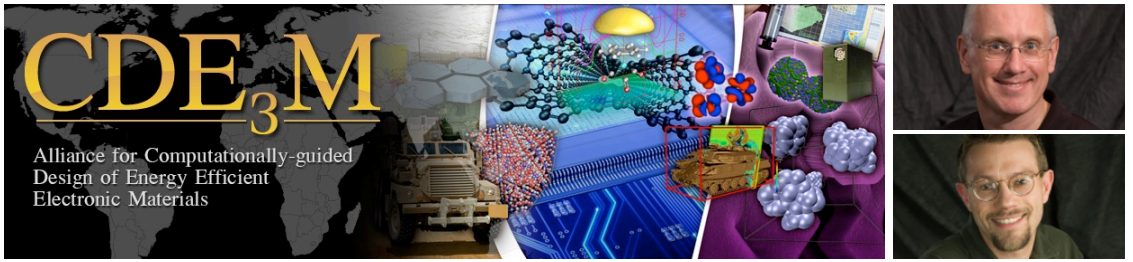
CEDMAV

UNIVERSITY OF UTAH

CENTER FOR EXTREME DATA MANAGEMENT, ANALYSIS, AND VISUALIZATION

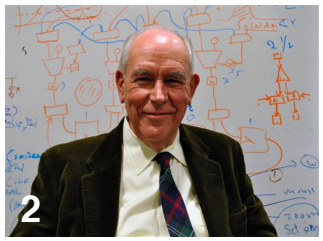
intel Graphics and Visualization Institute

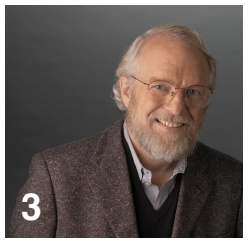


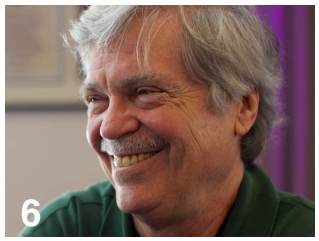


SCI INSTITUTE

Graphics, VR, and the Internet at Utah







1, 2. David Evans /Ivan Sutherland
- Founded CS Dept at the UofU in 1968
- Ivan Sutherland - Turing award
- Founded Evans & Sutherland Company

- **3. John Warnock**
- Worked at Evans & Sutherland
- Founded Adobe
- Hidden Line Removal Algorithm
- Helped invent Postscript @ Adobe

4. Ed Catmull
- Worked at Lucas Film
- Co-Founded Pixar
- President of Disney Animation Studios
- Chair of CoE External Advisory Board

5. Jim Clark
- Founded SGI, Netscape, Healtheon
- Work in Geometry Pipelines

6. Alan Kay
- Personal Computer
- Turing Award Winner
- Object Oriented Languages

7. Nolan Bushnell
- Invented Pong
- Founded Atari

8. Jim Kajyia
- Rendering Equation
- VP Research at Microsoft

9. Tom Stockham
- Known for work in Signal Processing
- Helped to invent the CD Player

10. Jim Blinn
- Invented Blinn-Phong Shading Model

11. Henri Gouraud
- Invented Gouraud Shading Model

12. Bui Tuong Phong
- Invented Phong Reflection and Shading Models

13. Allen Ashton
- Word Perfect
- My CFO Founder

SCI INSTITUTE

Productivity Machines

SCI INSTITUTE

IEEE VIS 2020

Salt Lake City!