CALEYDO

From Visual Exploration to Storytelling and Back Again

Samuel Gratzl

Alexander Lex
@alexander_lex

Nils Gehlenborg
@nils_gehlenborg

Nicola Cosgrove
@nicola_lady

Marc Streit
@marc_streit

THE UNIVERSITY OF UTAH

JOHANNES KEPLER UNIVERSITY LINZ

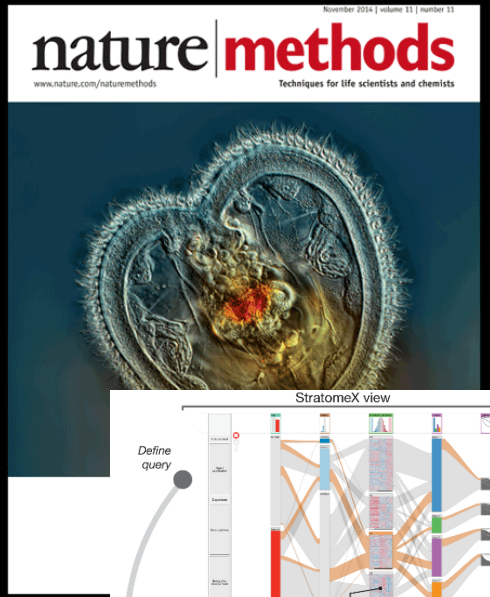
HARVARD MEDICAL SCHOOL

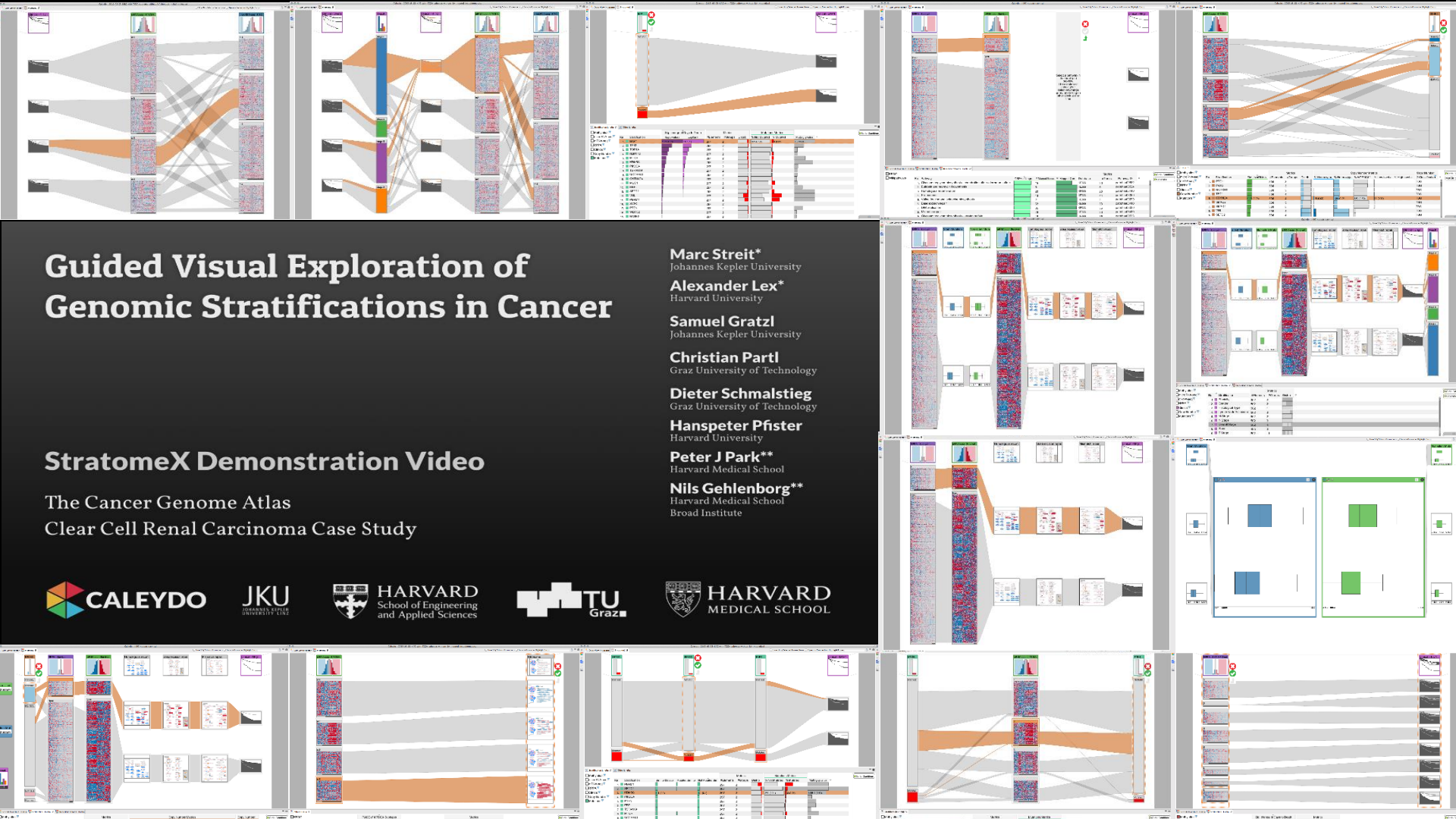
Exploration

Presentation

Guided visual exploration of genomic stratifications in cancer

Marc Streit, Alexander Lex, Samuel Gratzl, Christian Partl, Dieter Schmalstieg, Hanspeter Pfister, Peter J. Park, and Nils Gehlenborg,
Nature Methods 11, 9 (2014), 884–885





Guided Visual Exploration of Genomic Stratifications in Cancer

Marc Streit*
Johannes Kepler University

Alexander Lex*
Harvard University

Samuel Gratzl
Johannes Kepler University

Christian Partl
Graz University of Technology

Dieter Schmalstieg
Graz University of Technology

Hanspeter Pfister
Harvard University

Peter J Park**
Harvard Medical School

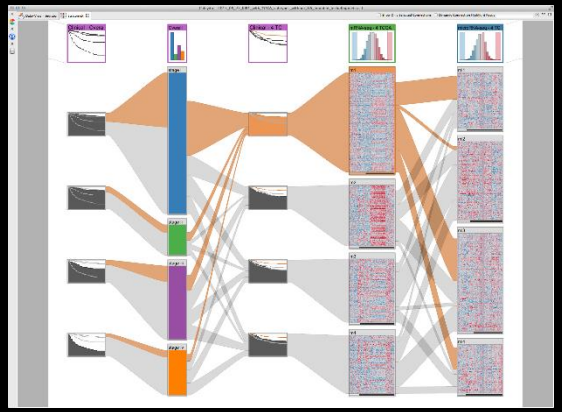
Nils Gehlenborg**
Harvard Medical School
Broad Institute

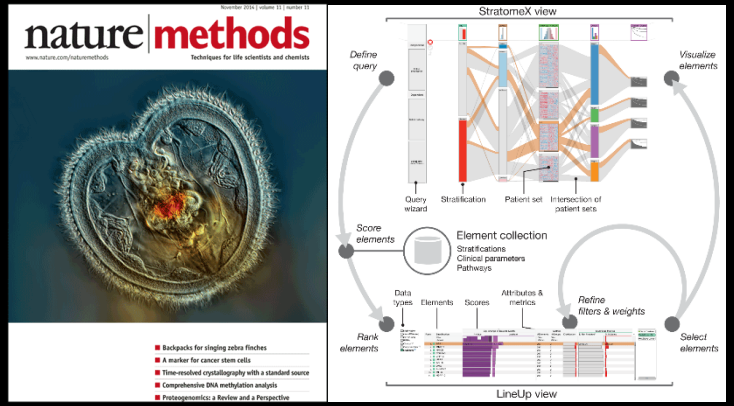
StratomeX Demonstration Video

The Cancer Genome Atlas
Clear Cell Renal Carcinoma Case Study

CALEYDO JKU HARVARD School of Engineering and Applied Sciences TU Graz HARVARD MEDICAL SCHOOL

BBC FOUR
lifespan
75 years
50 years
25 years
JAPAN
USA
BRAZIL
IRAN
1948
income
$400
$4 000
$40 000

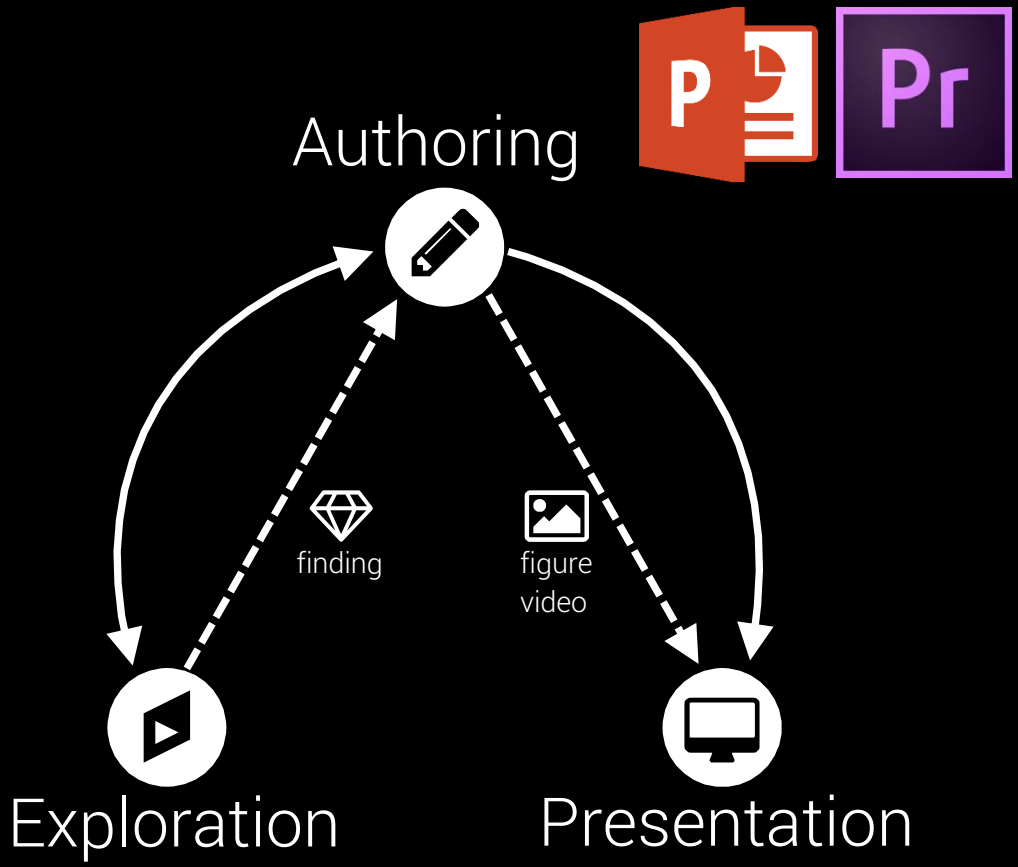


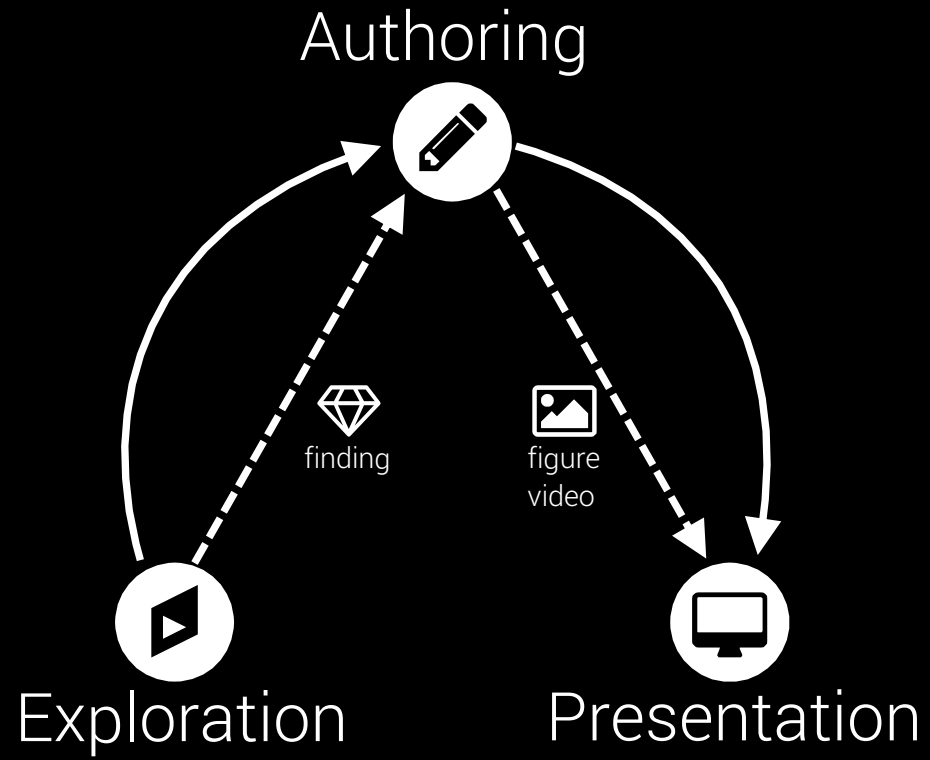


Exploration

Presentation

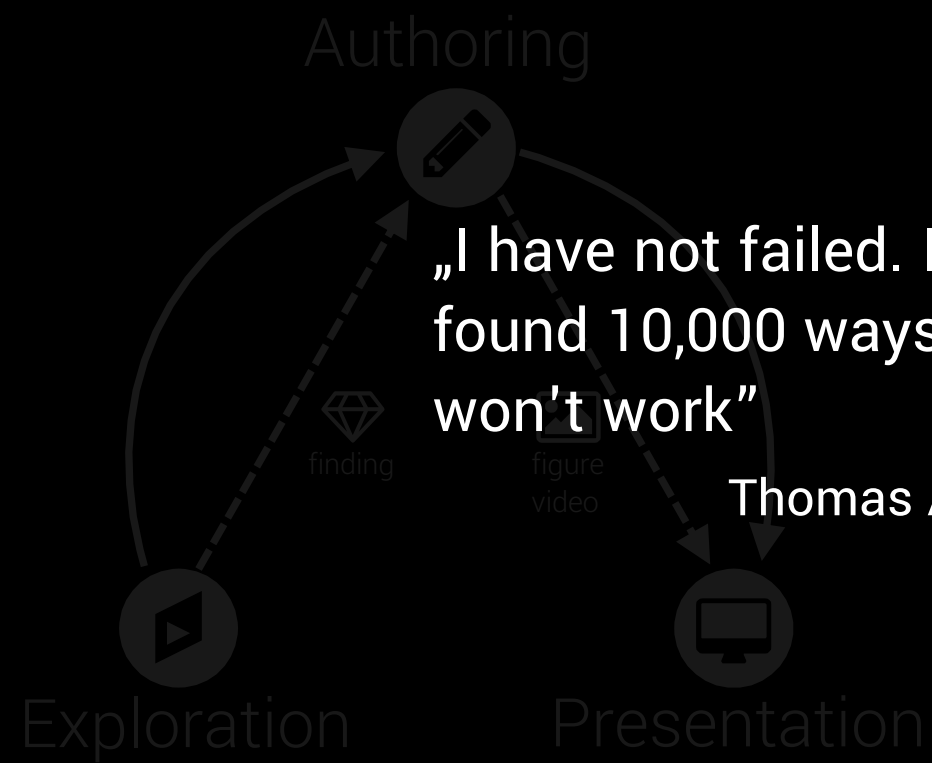
CALEYDO


Authoring
finding
figure
video
Exploration
Presentation



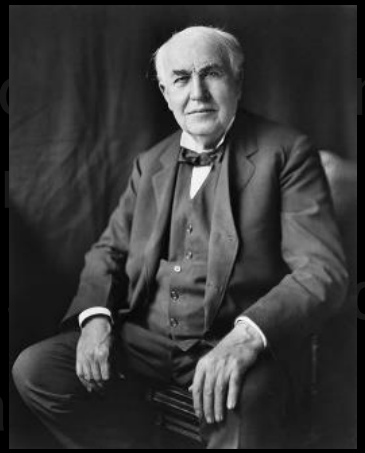
Issues

- fraction/selection of the exploration
- tell just the „good“ parts
- no back link

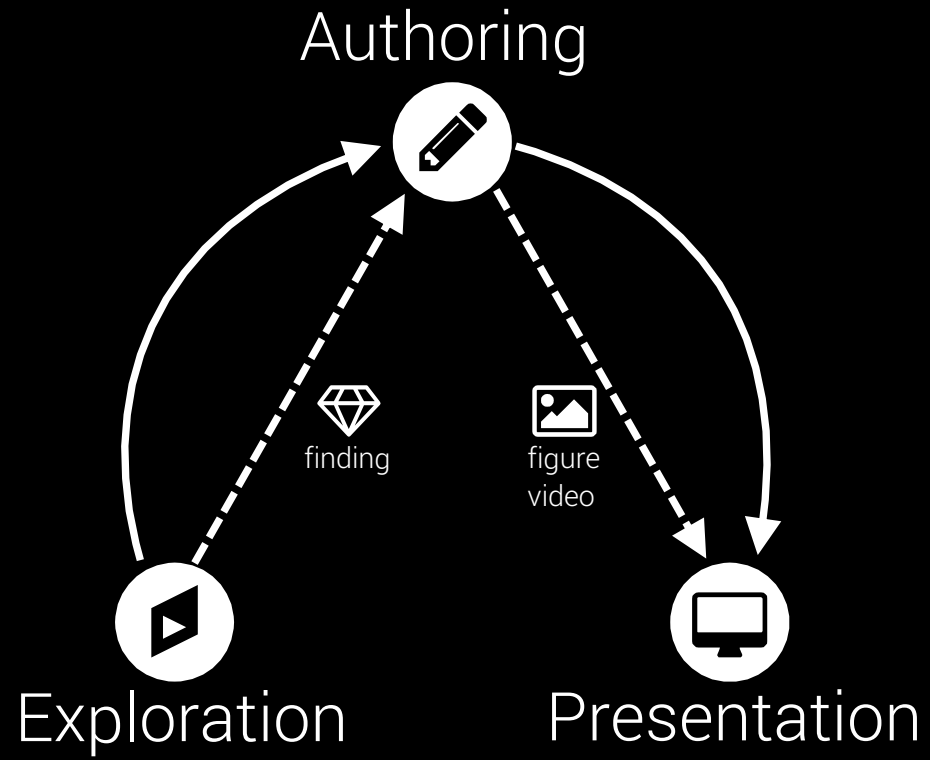


„I have not failed. I've just found 10,000 ways that won't work”

Thomas A. Edison

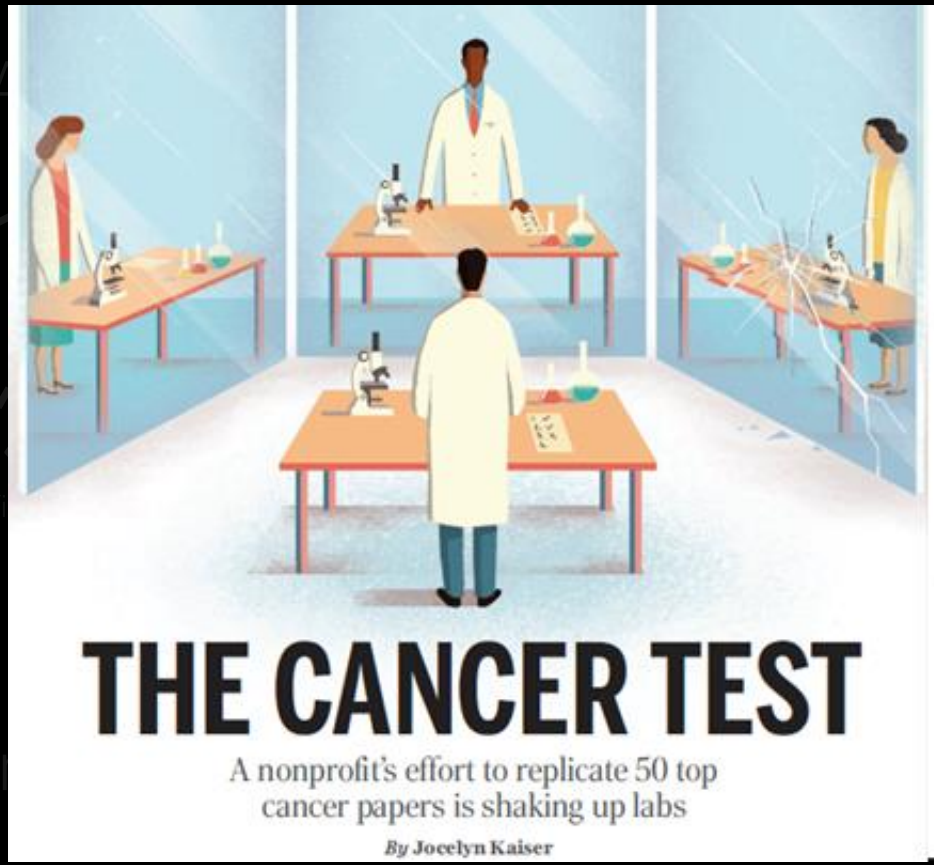


https://en.wikipedia.org/wiki/File:Thomas_Edison2.jpg



Issues

- fraction/selection of the exploration
- tell just the „good“ parts
- no back link



THE CANCER TEST

A nonprofit's effort to replicate 50 top cancer papers is shaking up labs

By Jocelyn Kaiser

Repeat failures

6 of 53
Cancer papers that Amgen could reproduce

14 of 67
Biomedical papers that Bayer completely reproduced

55%
MD Anderson researchers who could not reproduce a published study

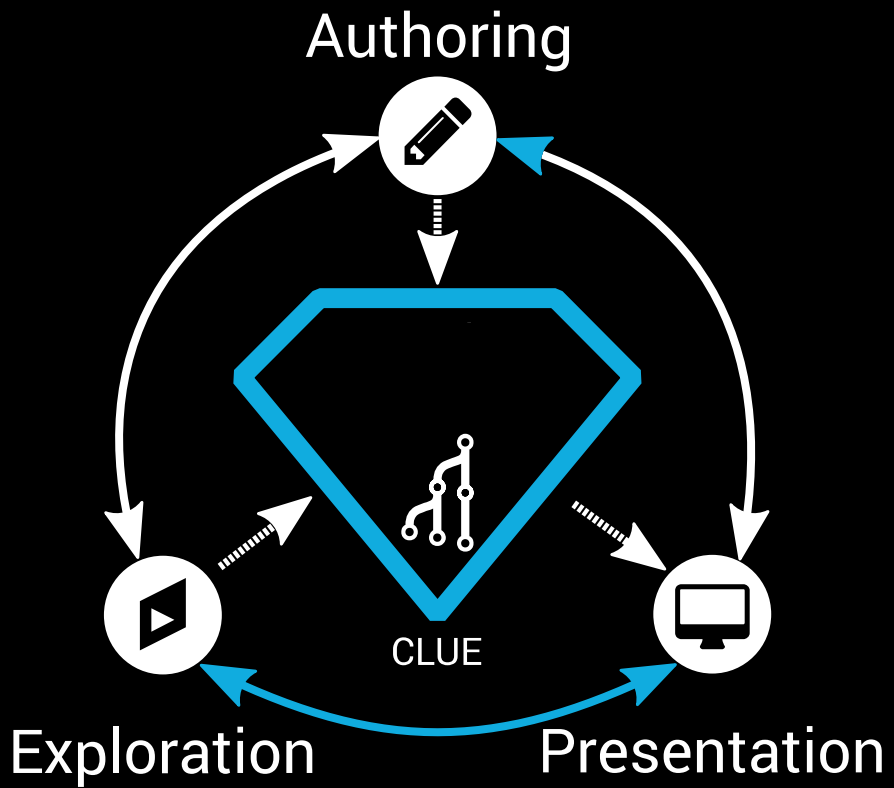
J. Kaiser, "The Cancer Test," Science, vol. 348, no. 6242, pp. 1411–1413, 2015.

continue

How to reproduce the presented exploration?

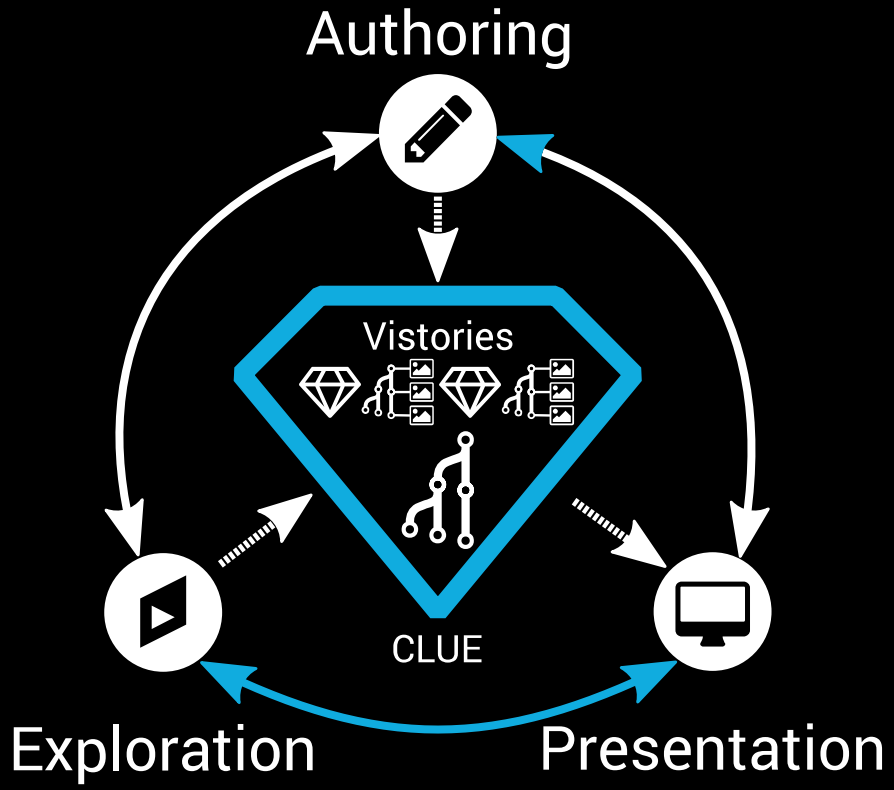
alter

How to go from
visual exploration to storytelling
and back again?



C apture
L abel
U nderstand
E xplain

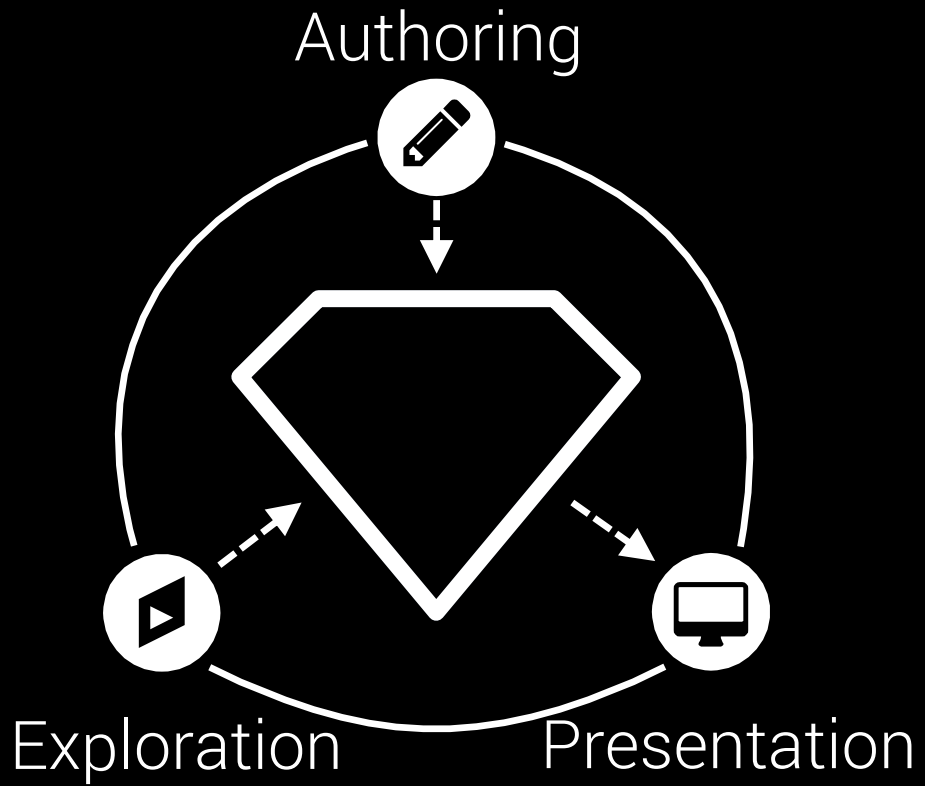
visualization driven explorations based on automated recorded provenance data.

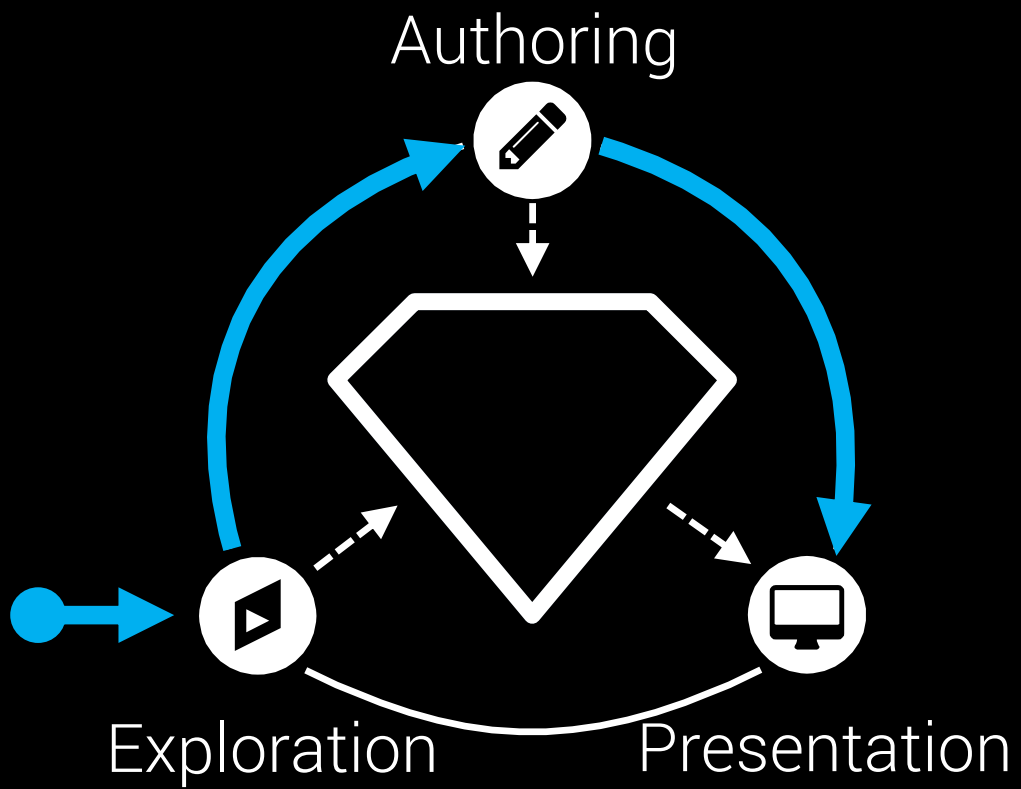


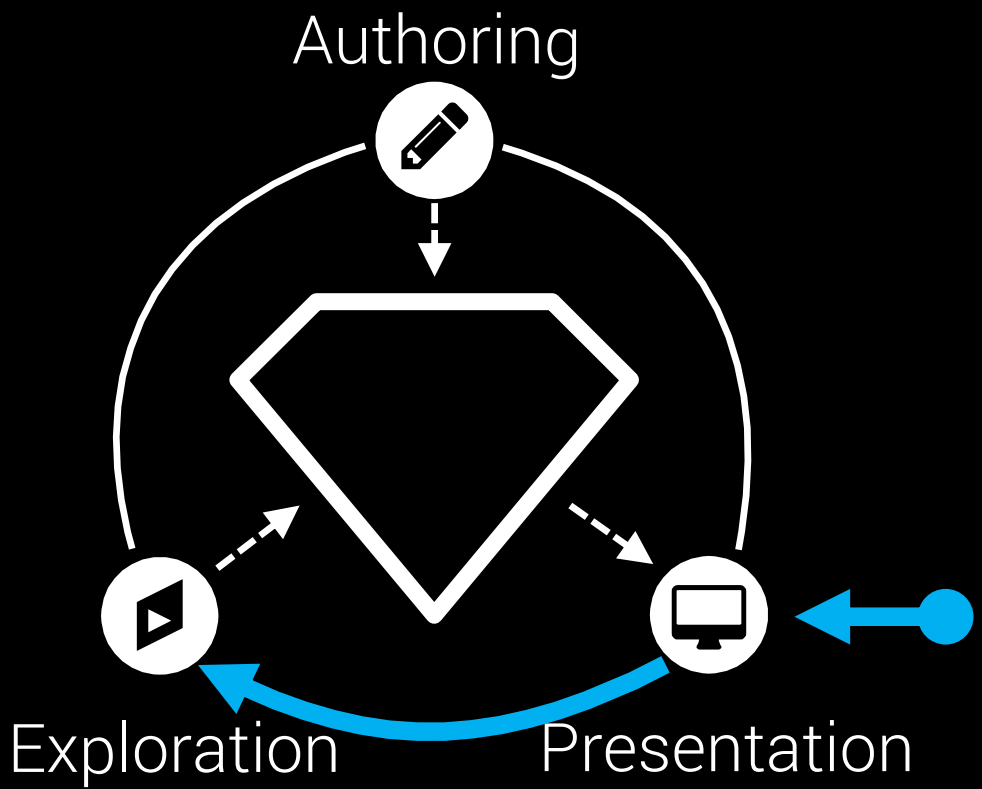
Vistories

- visual data stories
- visual histories

stories consisting of authored subset of the provenance graph

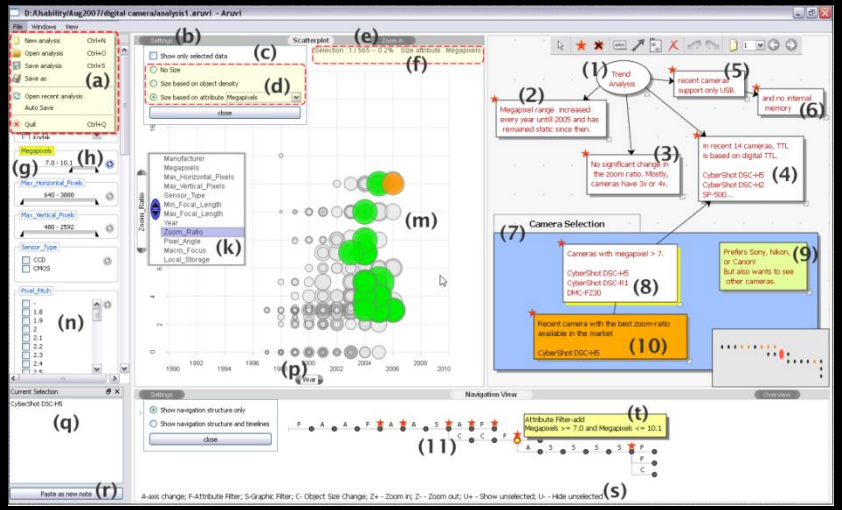

Authoring
Exploration
Presentation


Authoring
Exploration
Presentation


Authoring
Exploration
Presentation

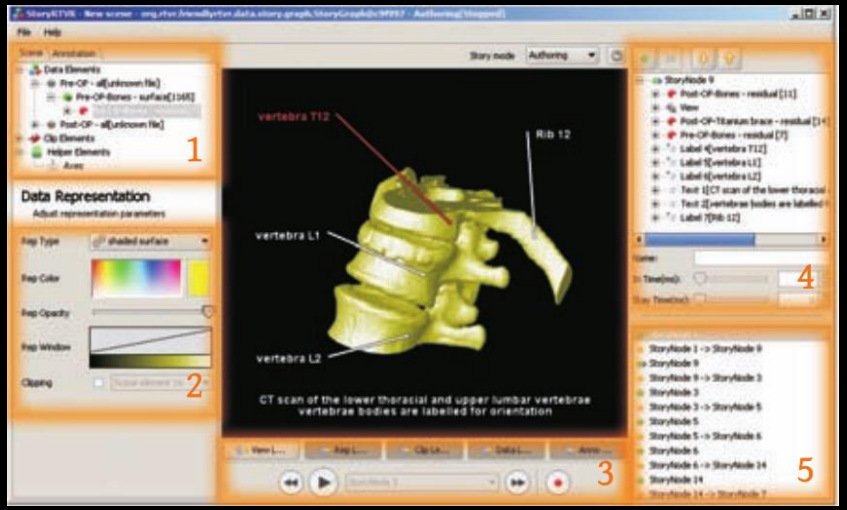
Shrinivasan and van Wijk, 2008

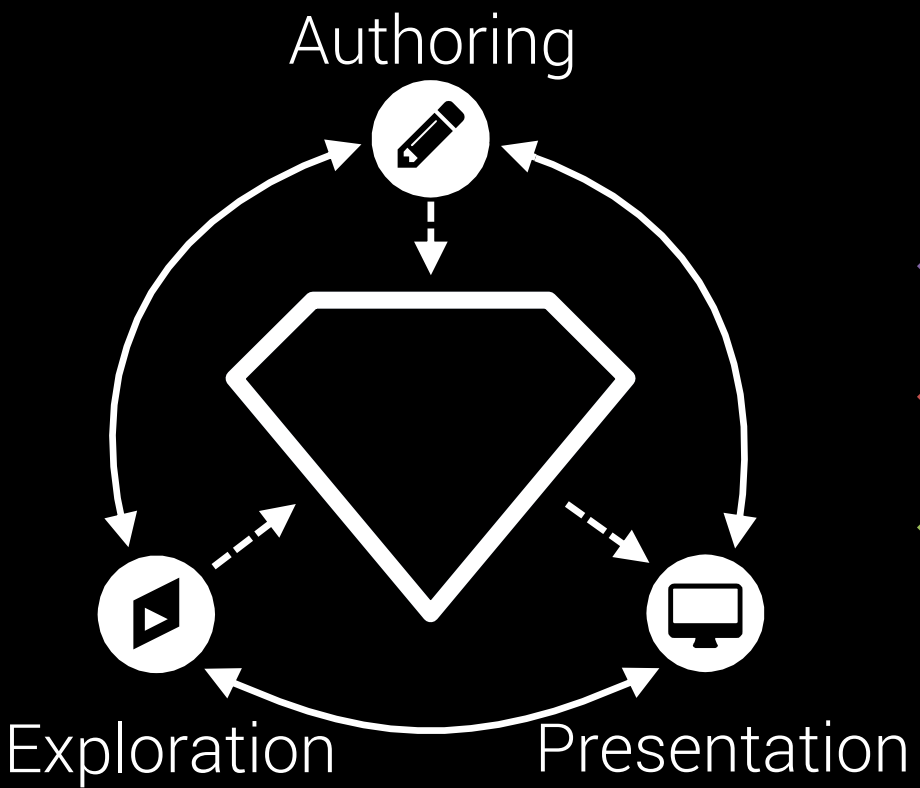
Supporting the Analytical Reasoning Process in Information Visualization



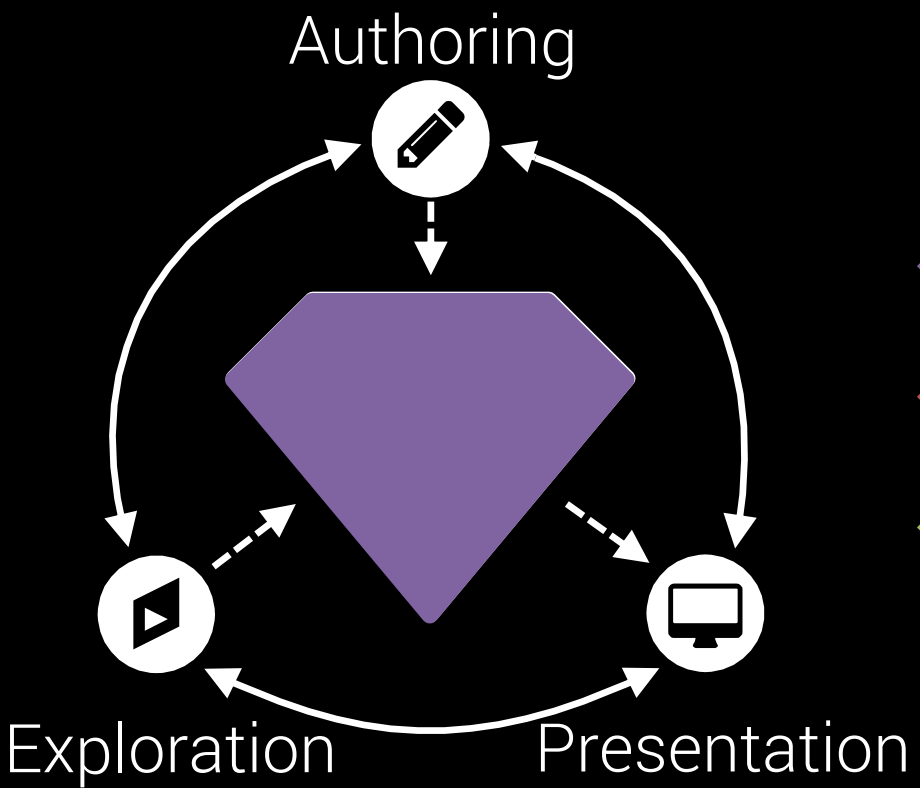
Wohlfahrt and Hauser, 2007

Story Telling for Presentation in Volume Visualization

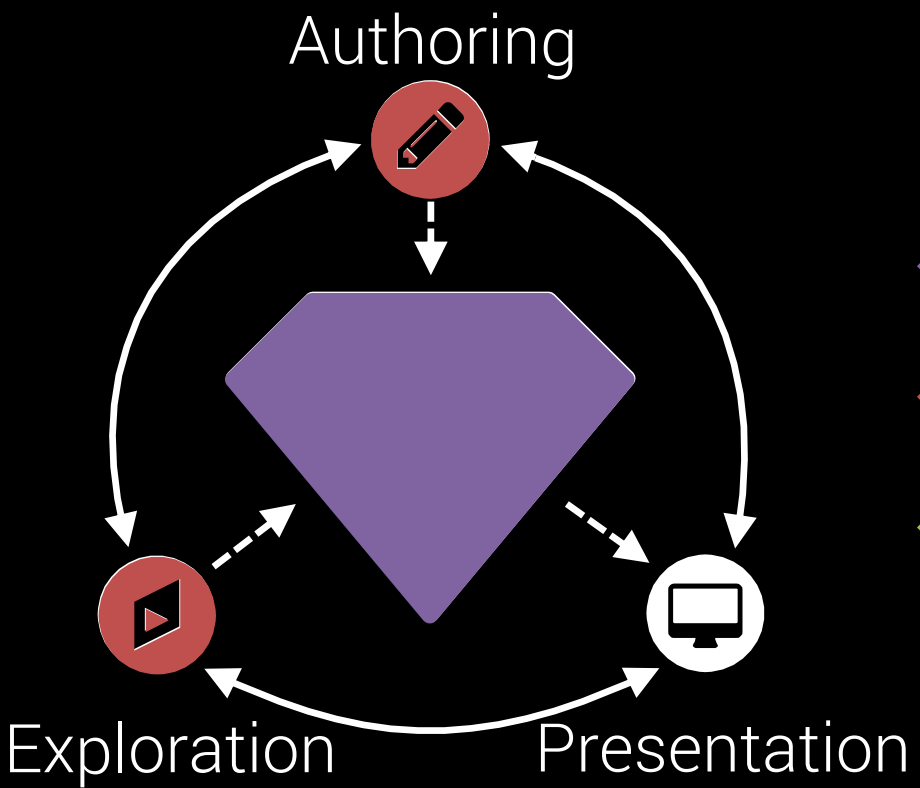




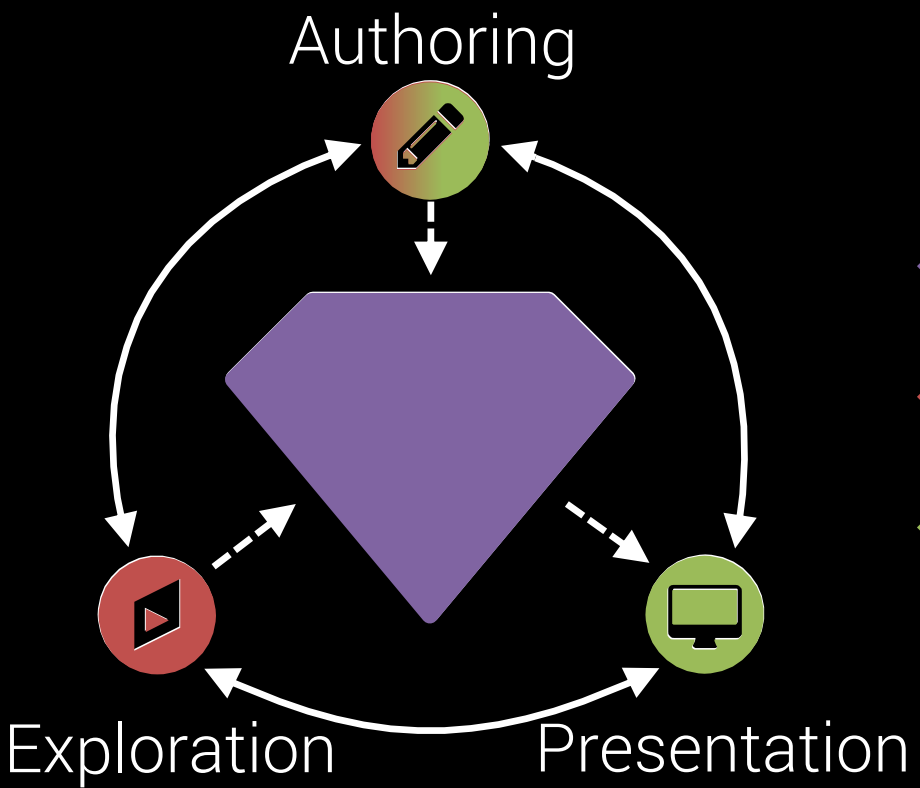
- Provenance Graph Recording
- Provenance Graph Visualization
- Story Editor Visualization



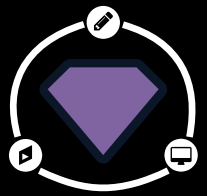
- Provenance Graph Recording
- Provenance Graph Visualization
- Story Editor Visualization



- Provenance Graph Recording
- Provenance Graph Visualization
- Story Editor Visualization



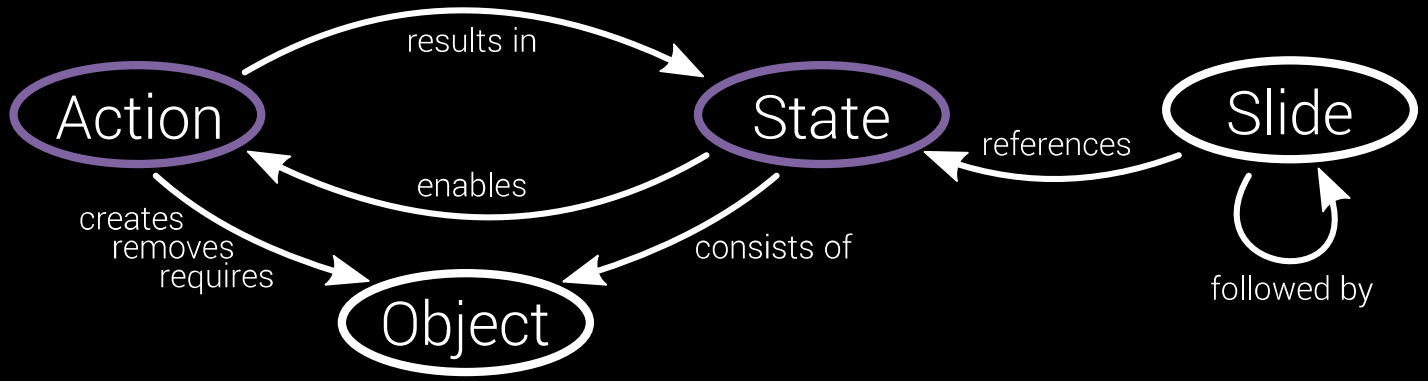
- Provenance Graph Recording
- Provenance Graph Visualization
- Story Editor Visualization



Provenance Graph Recording

CALEYDO

- action-centered bi-partite graph
- a state is determined by the action sequence from the root



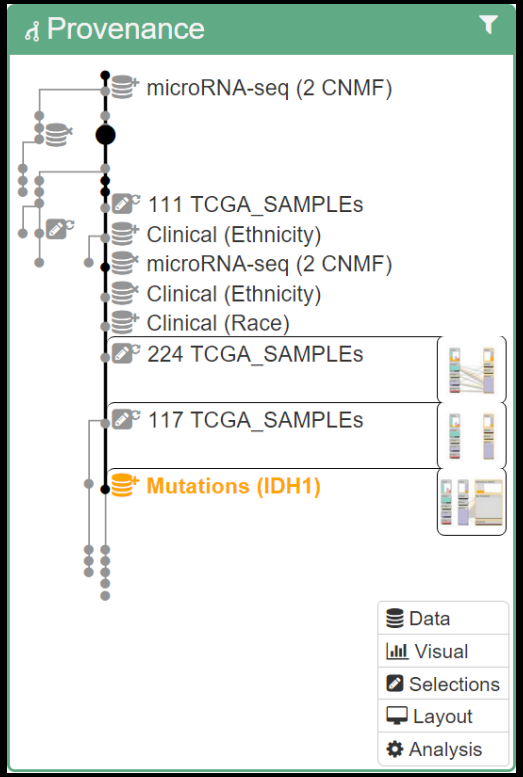
Provenance Graph View

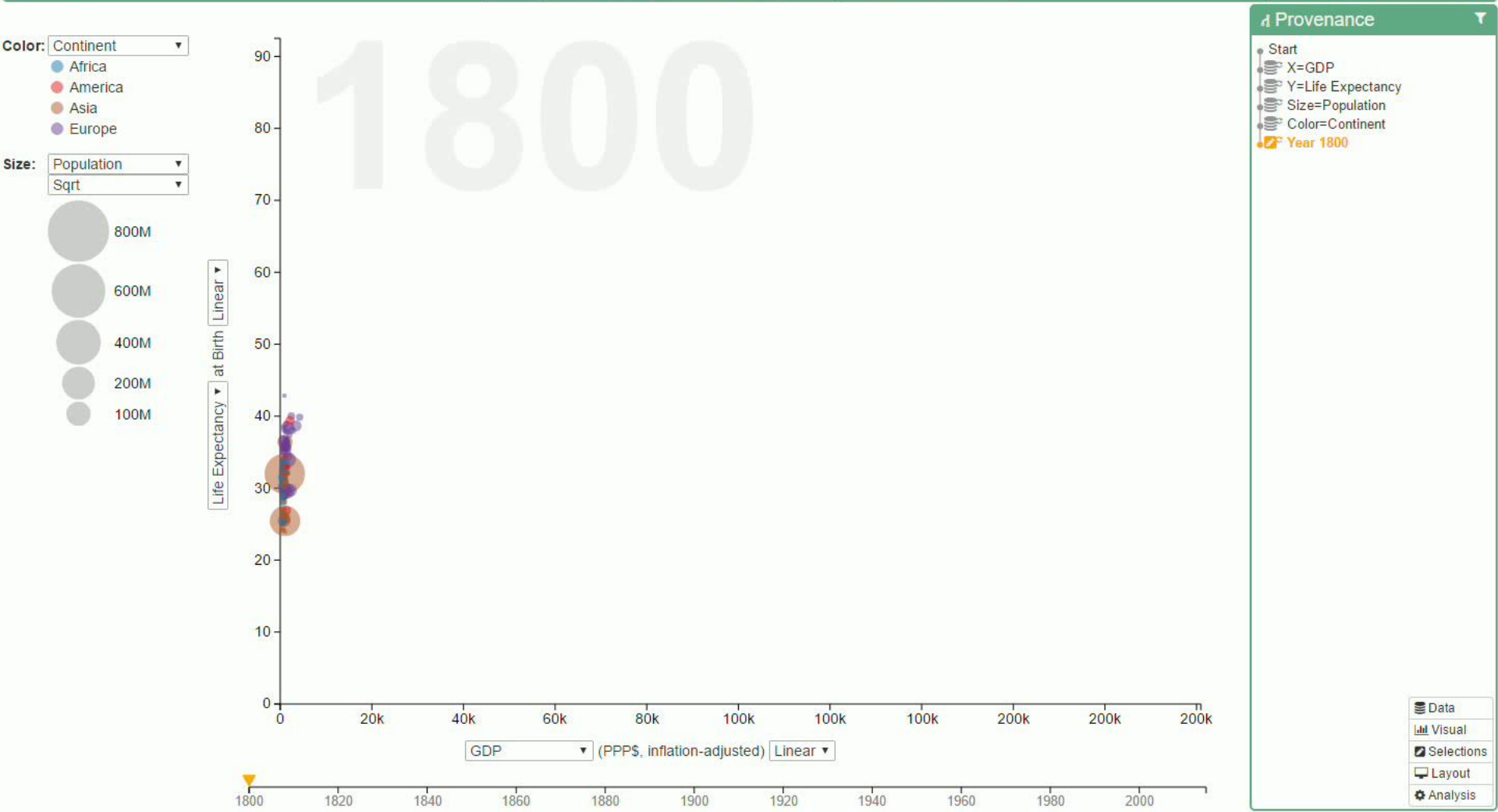
degree-of-interest driven skewed node-link diagram

- node = state
- link = action

purpose

- **exploration**: orientation, navigation
- **authoring**: selection





Provenance
- Start
- X=GDP
- Y=Life Expectancy
- Size=Population
- Color=Continent
- Year 1800

Data
Visual
Selections
Layout
Analysis

Story Editor View

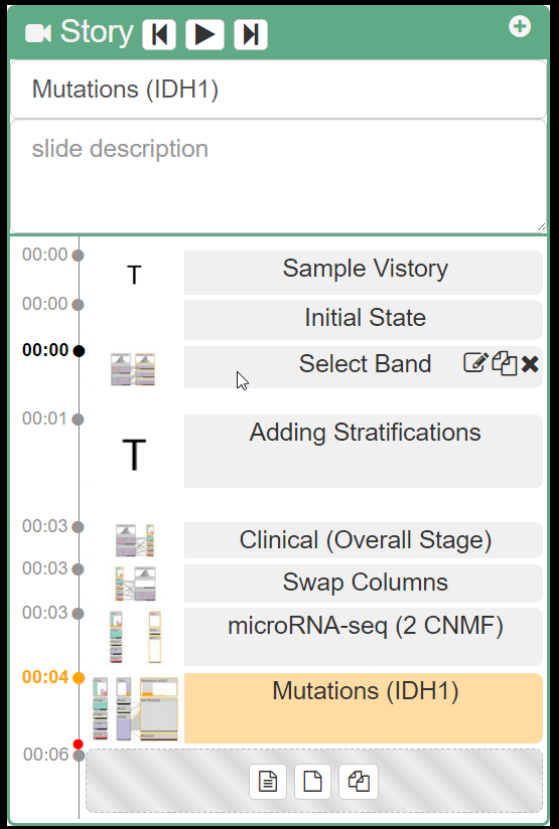
define one or multiple Vistories

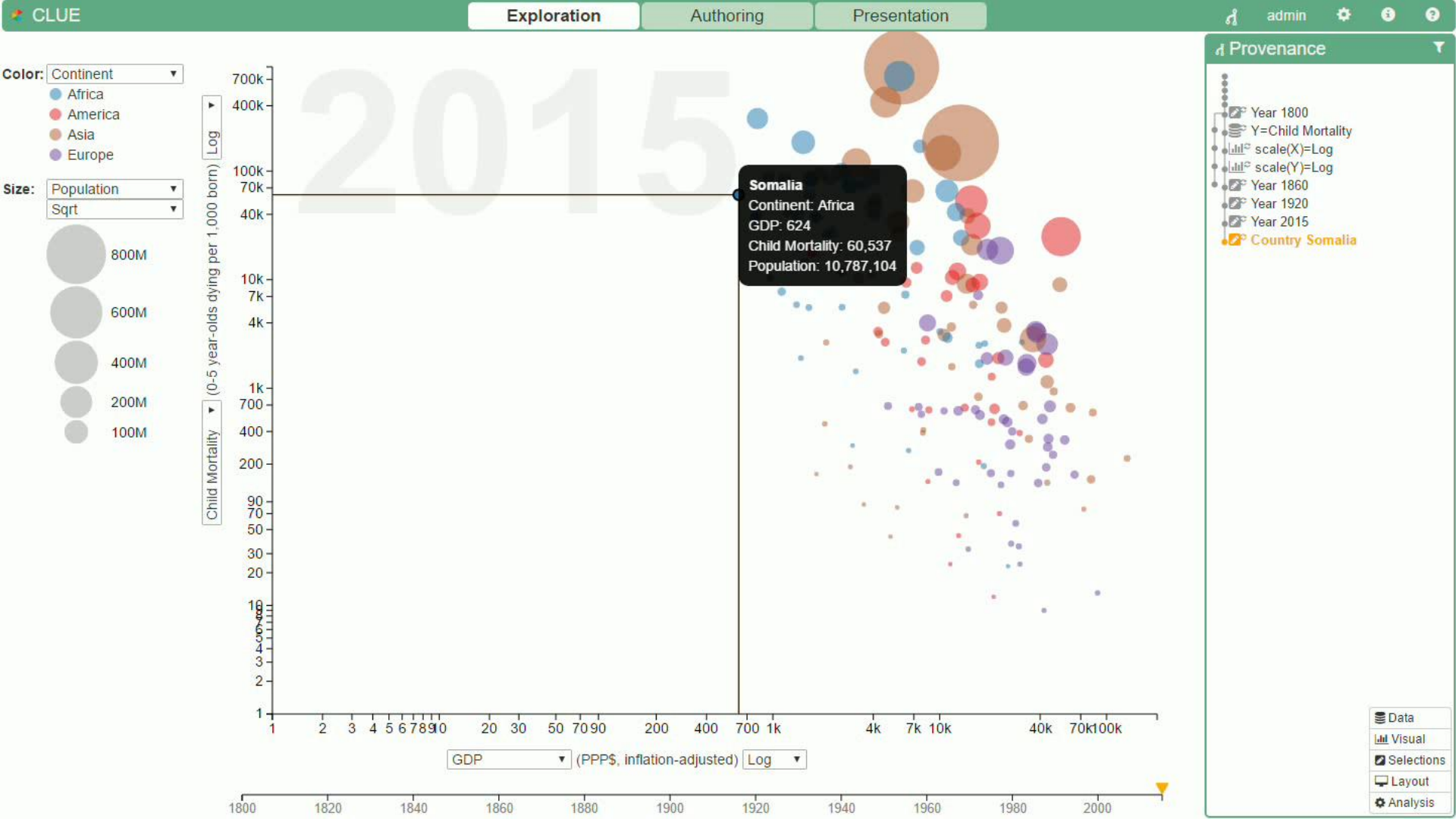
- select provenance states
- annotations (text, frame, arrow)
- timing information

purpose

authoring: definition, annotation

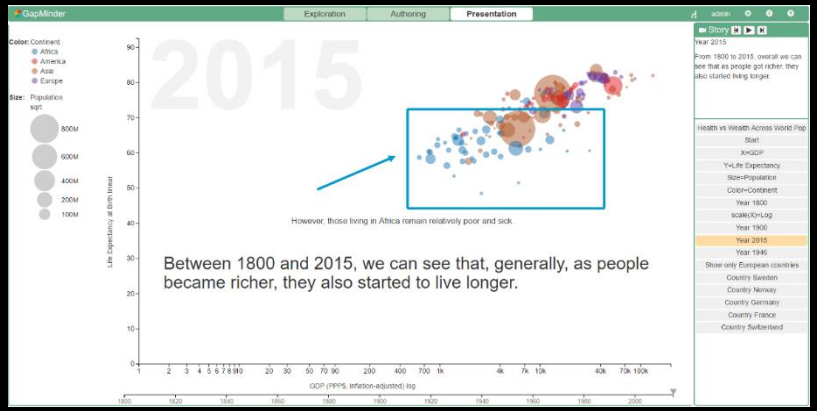
presentation: orientation, navigation




CLUE
Exploration
Authoring
Presentation
admin
Color: Continent
Africa
America
Asia
Europe
Size: Population
Sqrt
800M
600M
400M
200M
100M
2015
Child Mortality (0-5 year-olds dying per 1,000 born) Log
700k
400k
100k
70k
40k
10k
7k
4k
1k
700
400
200
90
70
50
30
20
10
1
Somalia
Continent: Africa
GDP: 624
Child Mortality: 60,537
Population: 10,787,104
1 2 3 4 5 6 7 8 9 10 20 30 50 70 90 200 400 700 1k 4k 7k 10k 40k 70k 100k
GDP (PPP$, inflation-adjusted) Log
1800 1820 1840 1860 1880 1900 1920 1940 1960 1980 2000
Provenance
Year 1800
Y=Child Mortality
scale(X)=Log
scale(Y)=Log
Year 1860
Year 1920
Year 2015
Country Somalia
Data
Visual
Selections
Layout
Analysis

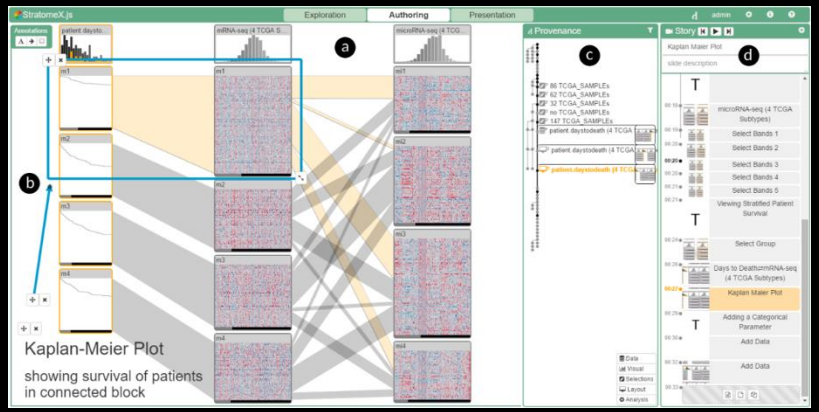
Usage Scenarios

Gapminder



http://vistories.org/v/gapminder

StratomeX



http://vistories.org/v/stratomex

Fork me on GitHub

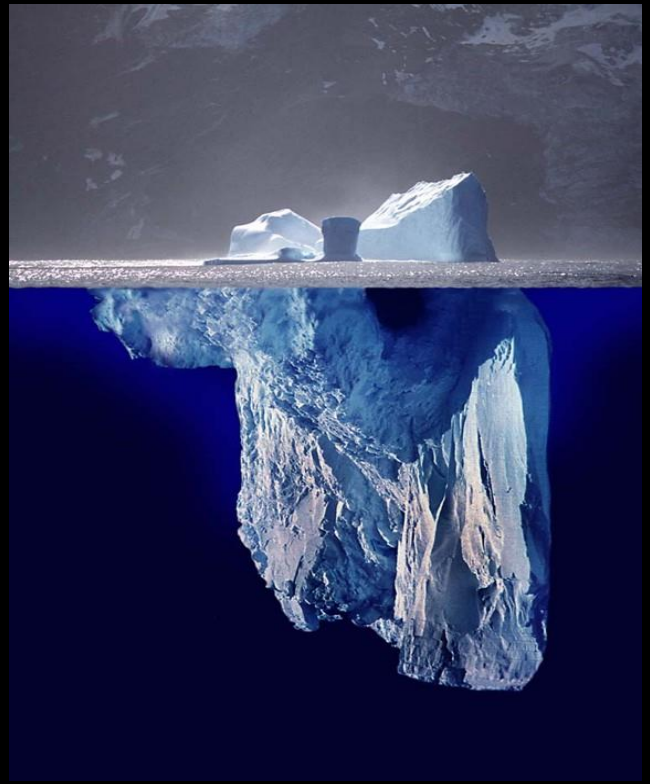
Implementation

CALEYDO

wrapper library around Caleydo Web applications

requires: command design pattern

open source (BSD license)
http://github.com/Caleydo/caleydo_clue



what is tracked

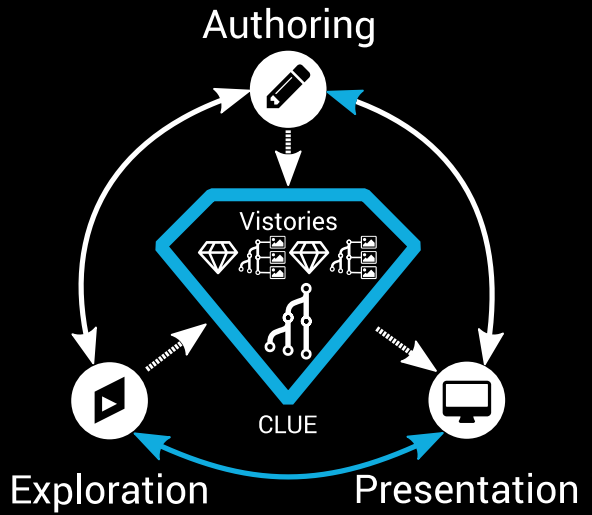
- visual exploration actions

what not

- datasets
- data processing / wrangling
- application
- meta provenance

https://en.wiktionary.org/wiki/File:Iceberg.jpg

C apture
L abel
U nderstand
E xplain



a framework for reproducible, shareable, and communicable visual explorations

Provenance Graph Analysis

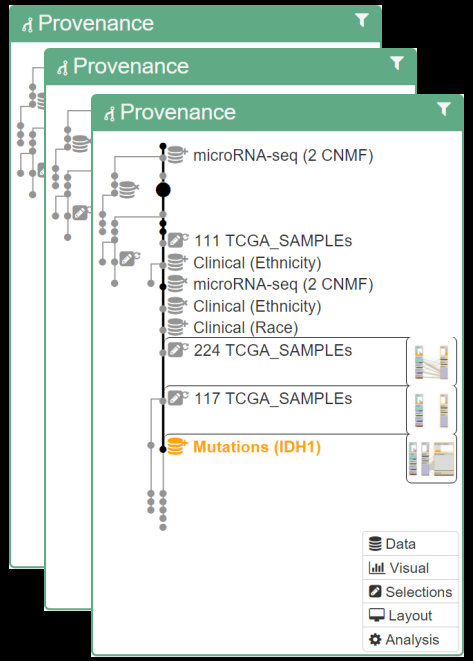
CALEYDO

support user

- similar states
- possible next actions

insights

- usage patterns
- exploration strategies



vistories.org

- visual explorations tracked like source code
- provenance part of the supplementary material
- use case figures have a DOI linking to their Vistory

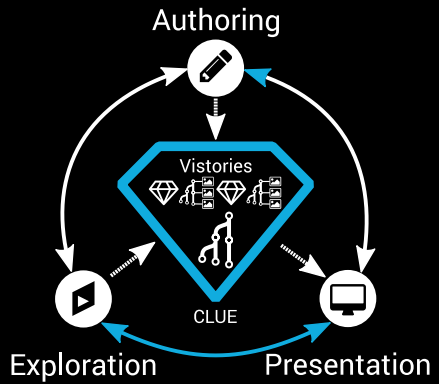
efficient

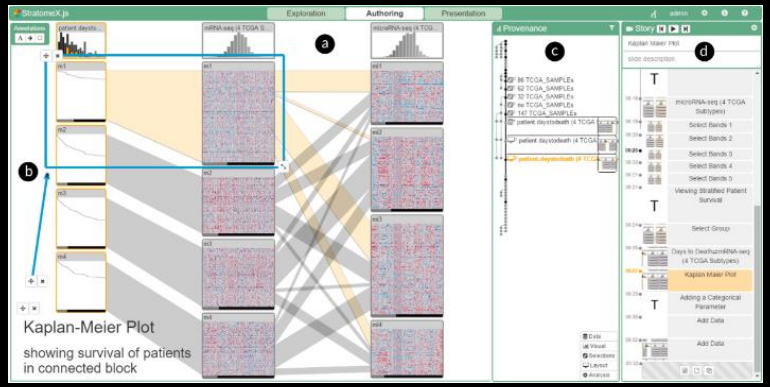
collaborative

sustainable

From Visual Exploration to Storytelling and Back Again

Samuel Gratzl, Alexander Lex, Nils Gehlenborg, Nicola Cosgrove, and Marc Streit





contact@caleydo.org
http://clue.caleydo.org

JOHANNES KEPLER UNIVERSITY LINZ

THE UNIVERSITY OF UTAH

HARVARD MEDICAL SCHOOL

caleydo.org

github.com/Caleydo

twitter.com/caleydo_org

FWF Der Wissenschaftsfonds.

NIH