

HiPiler:



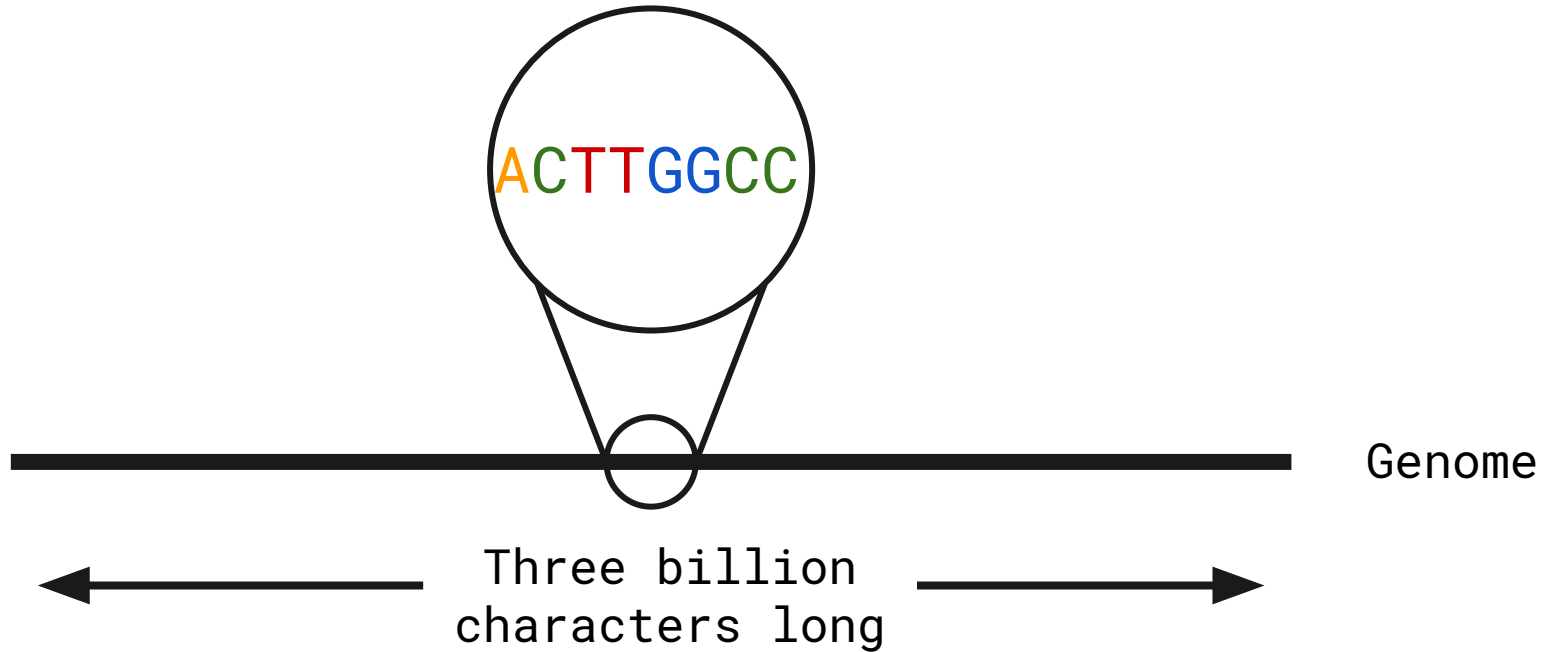
Visual exploration of large
genome interaction matrices with
interactive ***small multiples***

F. Lekschas, B. Bach, P. Kerpedjiev, N. Gehlenborg and H. Pfister
IEEE Trans. on Vis. and Com. Grphx. (Jan. 2018)
<http://hipiler.higlass.io/>

Presented by **Baraa Orabi**
CPSC 547 Information Visualization
The University of British Columbia
2019-11-19

Some background

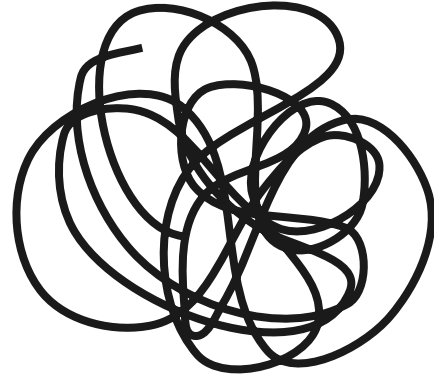
Genome



Genome is 3D

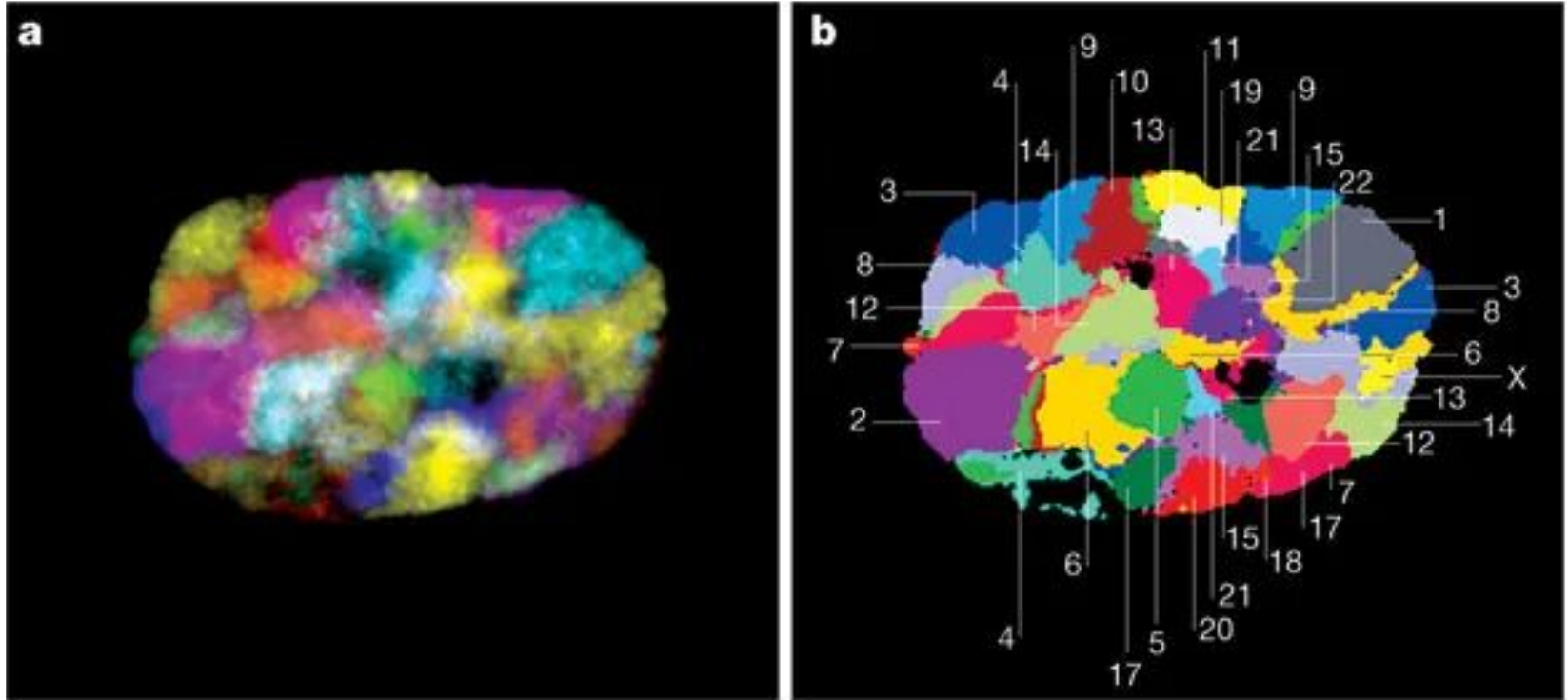


You wish -_-



More like it :-/

3D Genome of single* cell reconstructed

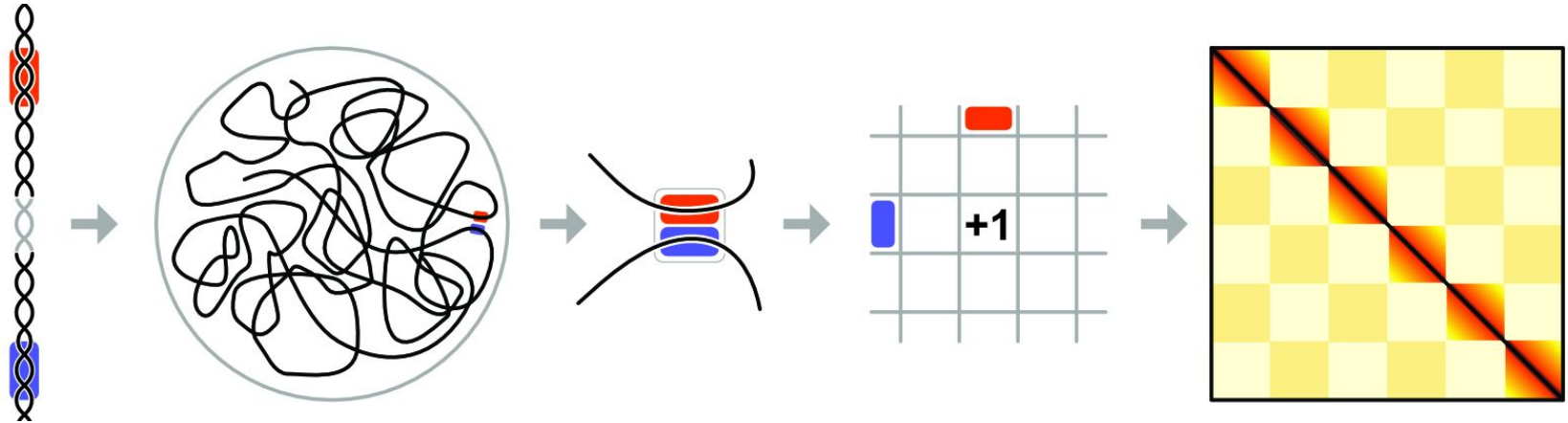


Speicher & Cartel Nat. Rev. Genet . 5: 782 (2005)

What data are we dealing with?

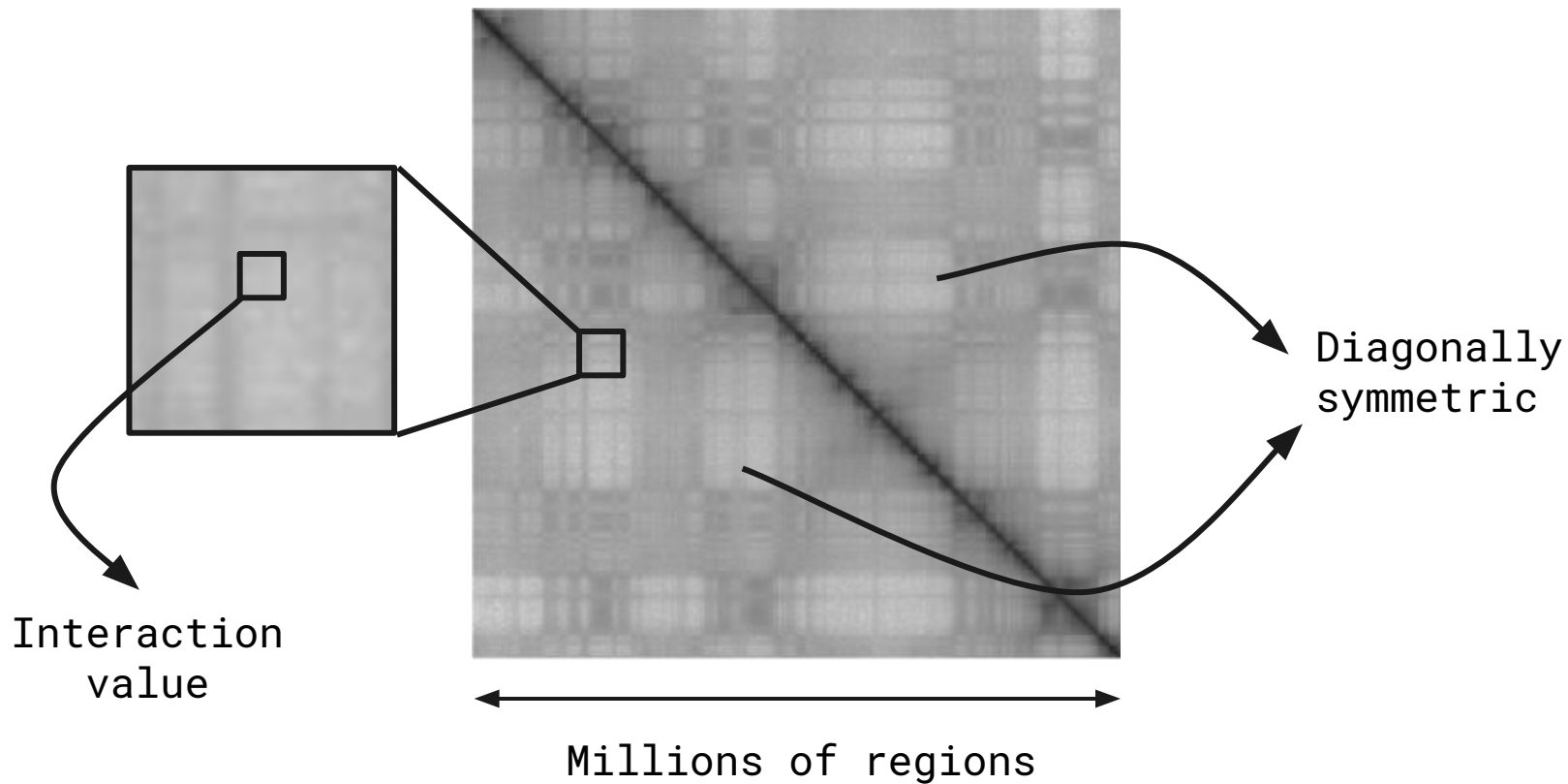
Matrices!

Hi-C data

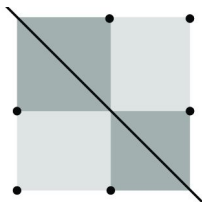
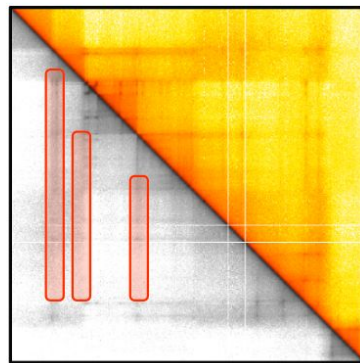
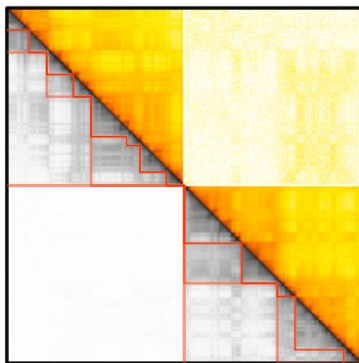
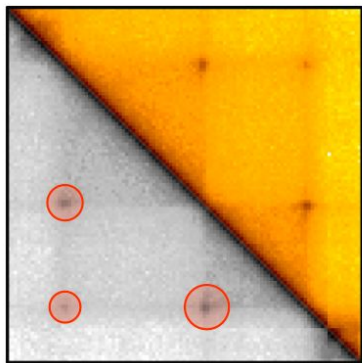


Edited Fig. 3 from HiPiler paper

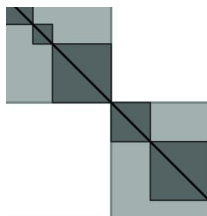
Interaction Matrix



Why: Regions of Interest (ROI)



Loops



Domains



Flames

Interaction Matrix

Task

- View ROI's
- Examine pattern of ROI
- Compare patterns of different ROI's
- Preserve global context

Challenge

- Too many ROI's
- Need to zoom
- Often far away

How: HiPiler



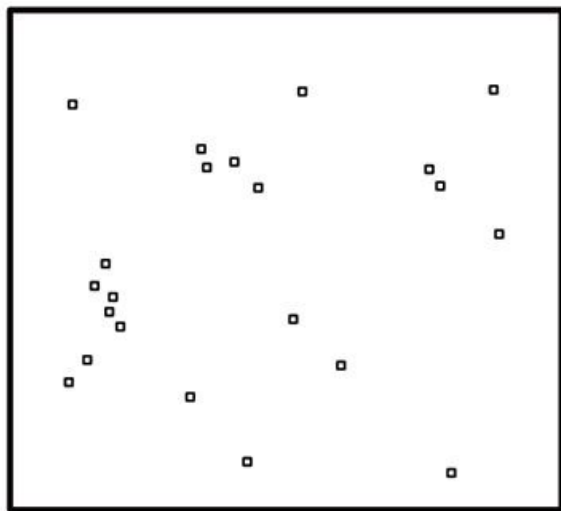
HiPiler

Task abstraction (abridged)

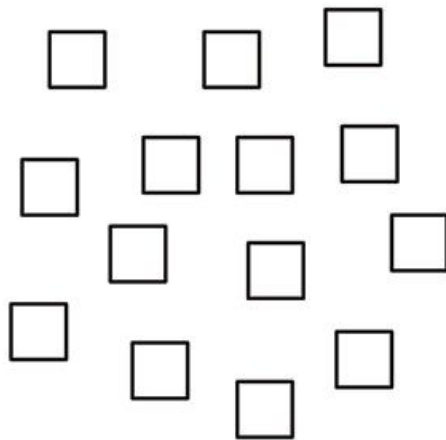
1. Search for known patterns
2. Examine one instance
3. Compare instances

The How

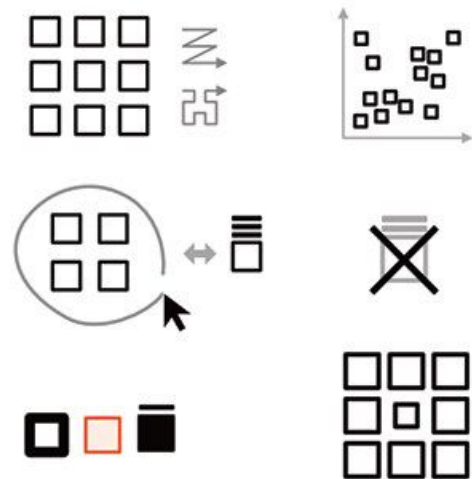
Matrix



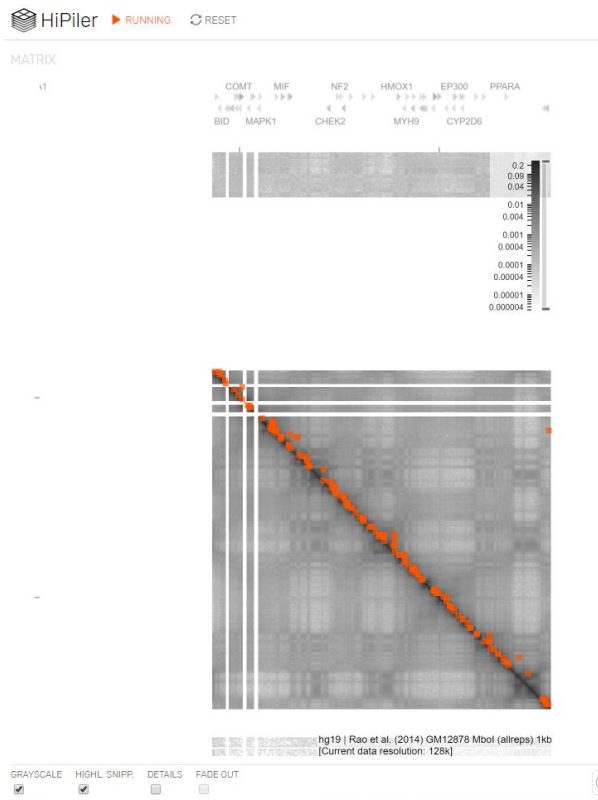
Snippets



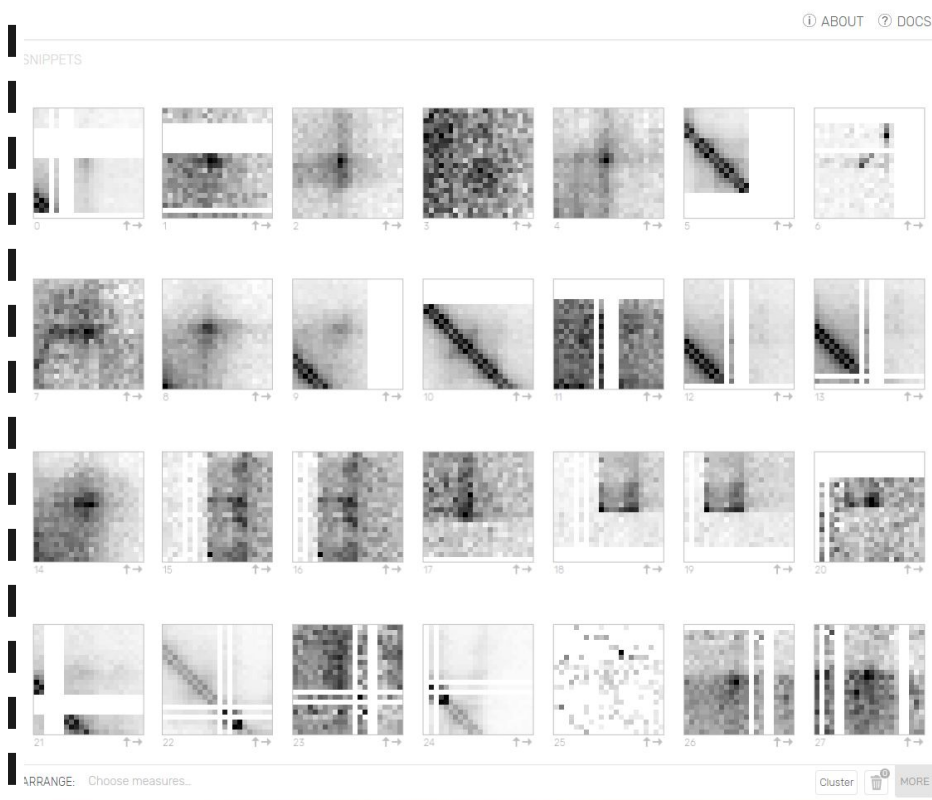
Interaction



The How

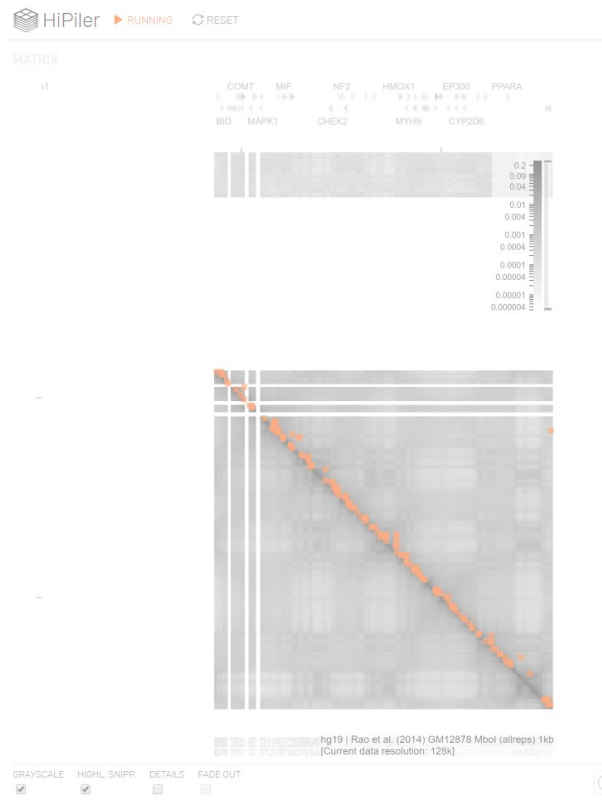


Context

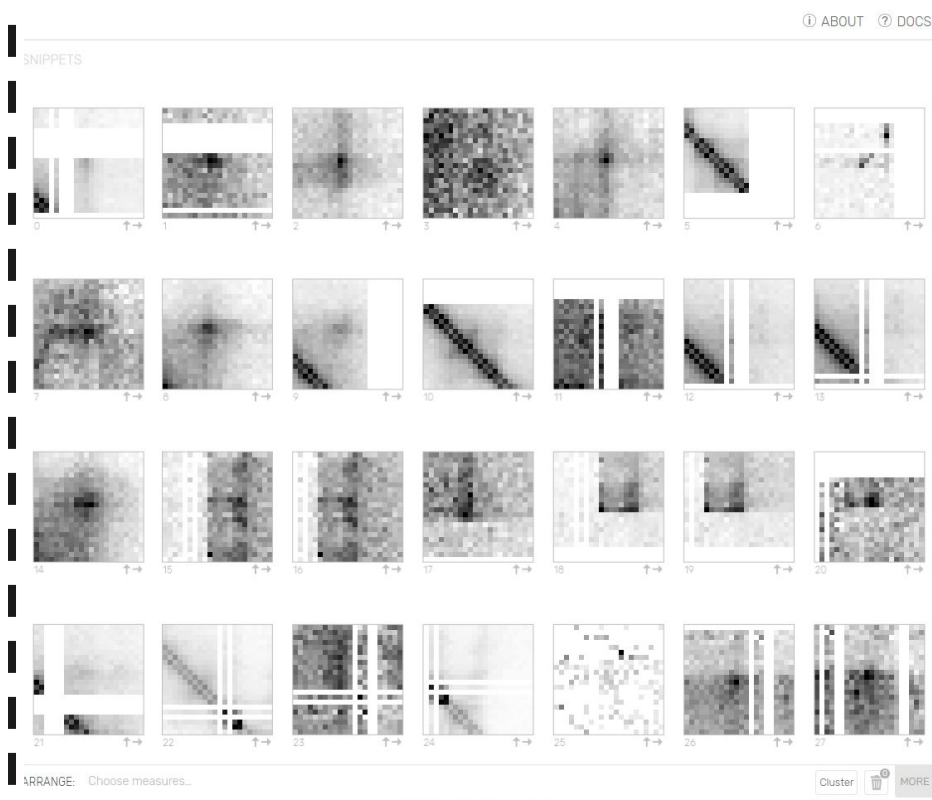


Snippets layout

The How



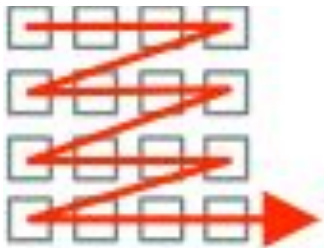
Context (mainly HiGlass)



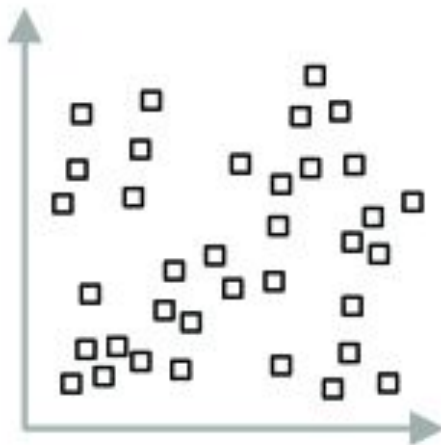
Snippets layout

Layout (using user specified metrics/attributes)

1D
sorting



2D
scatterplot



+3D
tSNE clustering



Layout: Multi-dimensional clustering

- tSNE clustering
- Configurable



Selection, aggregation, filtering

- Lasso selection
- Combine snippets into a pile
- Delete snippet (pile)

Aggregation

Pile cover



Mean



Var

In stack preview

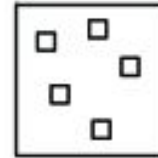


Cover is
snippet

Inspection



Pile



New view

Stated Limitations

- Only square ROI
- Fixed rows/cols order
- Scalability (beyond 2000 snippets)

Evaluation

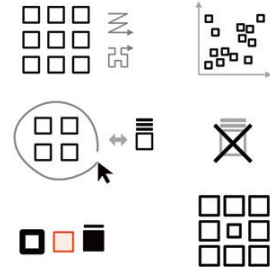
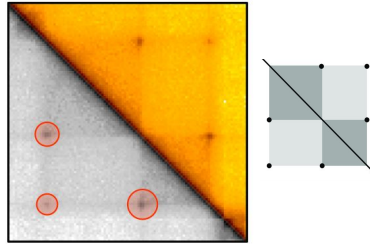
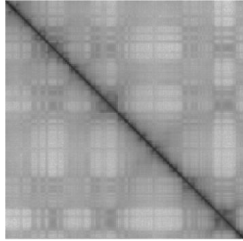
- Five domain experts/users
- 1-2 hours interview
 - Training
 - Pre-prepared data
 - Own data

Evaluation feedback

- Good snippet/matrix linking
- Validated (behaviourally) task abstractions
- Easy to learn
- Big improvement on state-of-the-art

Summary

Summary



What?

Interaction matrix

Why?

Explore and validate

How?

Snippets:

- Linked
- Laid out
- Select/Aggregate/Filter

Critique

Great!

- Tool:
 - Open source
 - Nice aesthetic
 - Does what it says!
- Paper:
 - Easy to read
 - Great online resources (demo, video, slides, docs, etc...)

Not so great :-)

- Software:
 - Selection sucks, real bad
 - Slow (too much on-the-go computation)
 - Heavy on client-side memory
 - Unintuitive settings
 - Very buggy
 - Too much installation overhead (server)

Thanks!
Questions?