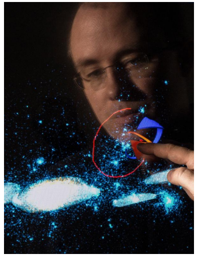
Aviz
Master internships

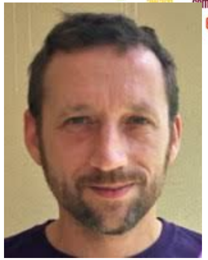
2020/2021

Introduction to Aviz

Aviz
Visual Analytics Project

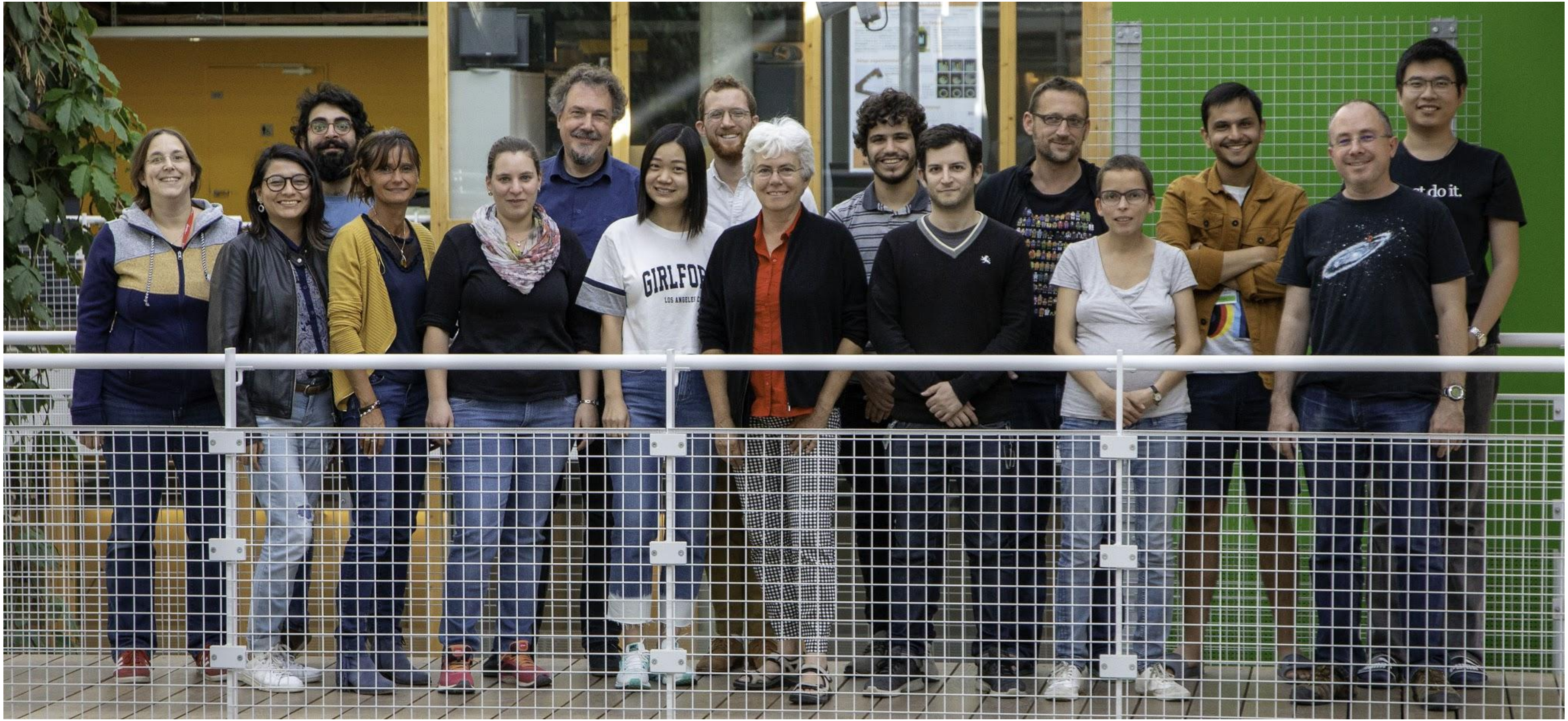
Inria informatics mathematics





Aviz
Visual Analytics Project

informatics mathematics
Inria



Scientific Context and Overall Objectives

- Data is now abundant. We need more effective methods, techniques, and tools to **understand** it and **make decisions** with it

- Aviz is a **multidisciplinary project** that
  - seeks to improve **data exploration** methods, techniques, and tools
  - based on **interactive visualization**

- **Visualization** is [Card et al. 99]:

  > “the use of computer-supported, interactive, visual representations of data to amplify cognition.”

Research Program

Aviz focuses on four research themes:

1. Progressive Data Analysis
2. Physicality in Input and Output
3. Perception, Cognition, and Decision Making
4. Methodologies for Visualization Research

1

Scalability

What have we done before?

Like everyone:

- Use faster languages, systems
- Avoid interacting, pre-compute
- Use more pixels

Scalability of Exploratory Data Analysis in the Big Data Era:

- When data grows or analyses become complex, operations take more time to complete
- Humans cannot explore data when latencies exceed ~10s
- How can we bound the latency without limiting the data size and analysis complexity?

1

Progressive Data Analysis

Progressive Data Analysis

- Generate results as a sequence of estimates improving over time
  - Estimates are generated with bounded latency
- Allow "Interactive Steering" to adjust parameters while computing
- Provide quality and progress measures

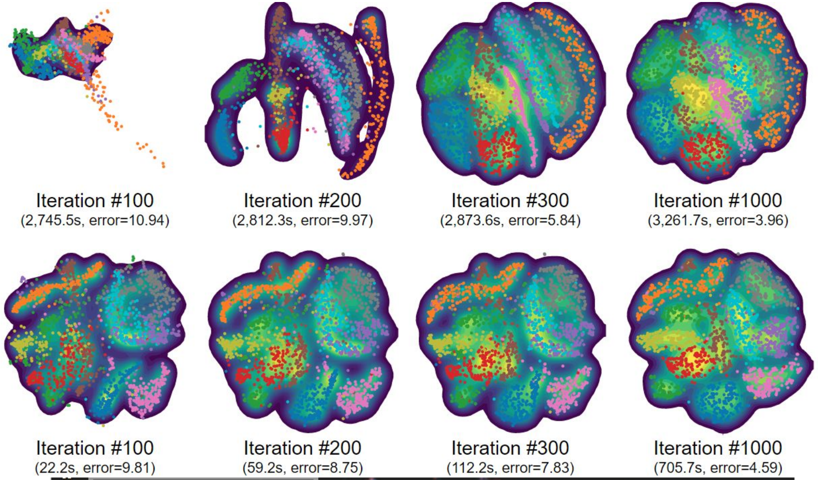
The few progressive systems described in the literature are ad-hoc

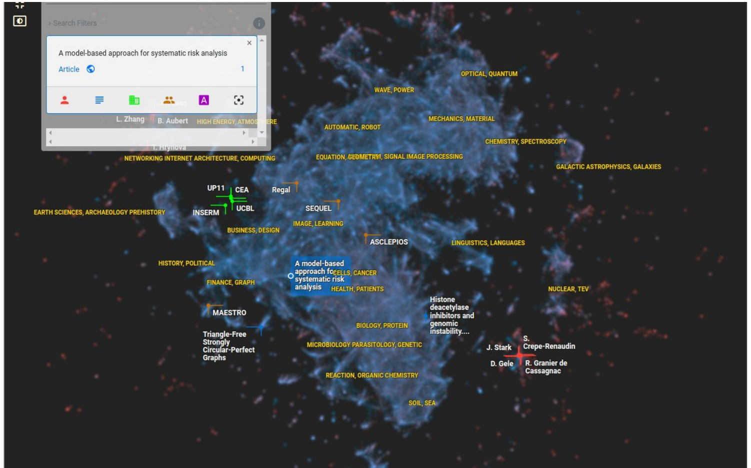
Aviz wants to design a language for analysis and visualization that is **natively progressive,** raising multiple research challenges

1

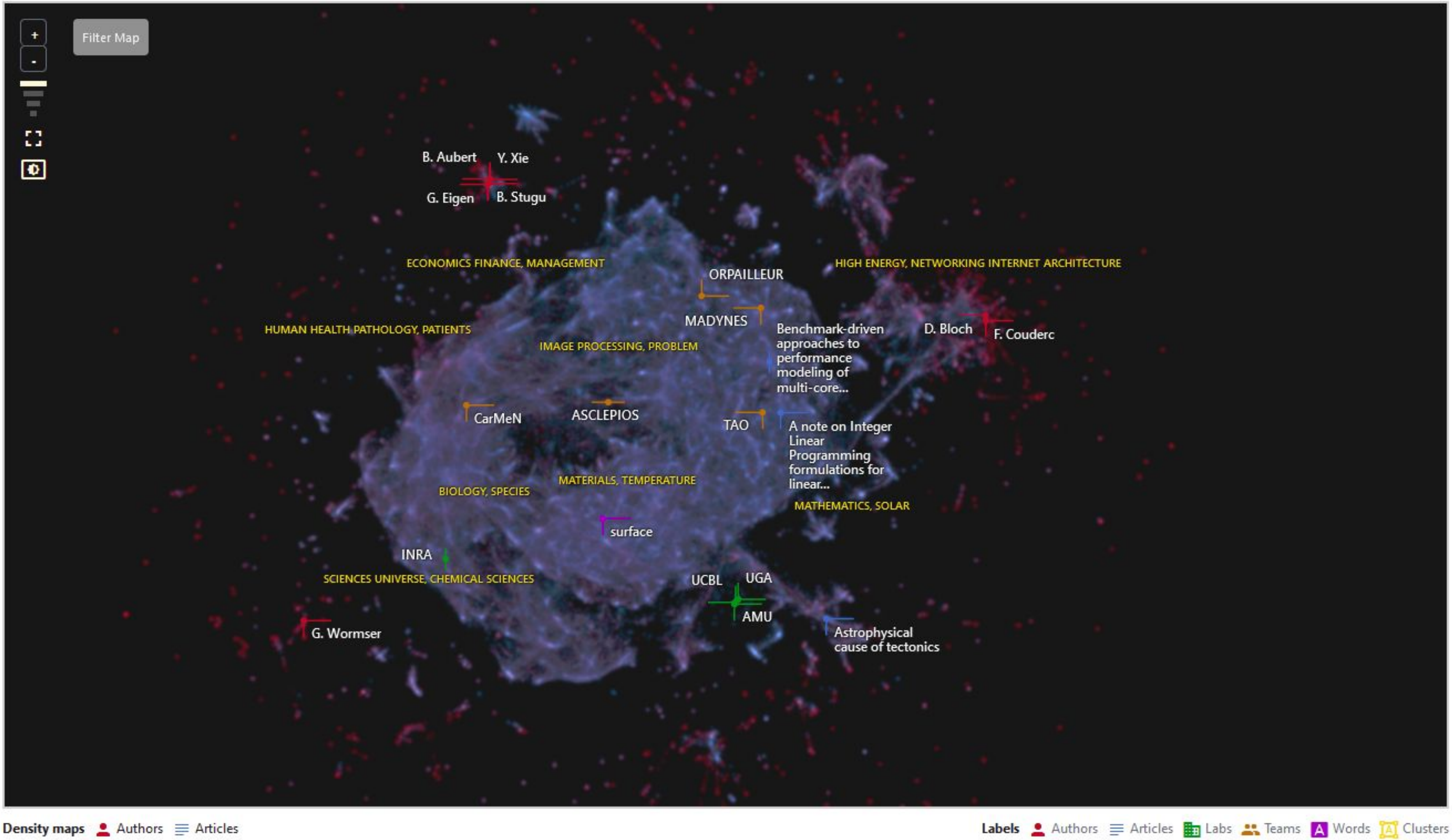
Progressive Data Analysis

- Progressive data analysis
  - Language for progressive computation
  - Visualizations, UI and requirements for progressive data analysis
  - Progressive algorithms e.g. kNN, projections, k-means, ML algorithms
- Scalable visualizations
  - Large multidimensional projections
  - Exploring BitCoin/Blockchain
  - Health data from CNAM





Projection of entities extracted from HAL

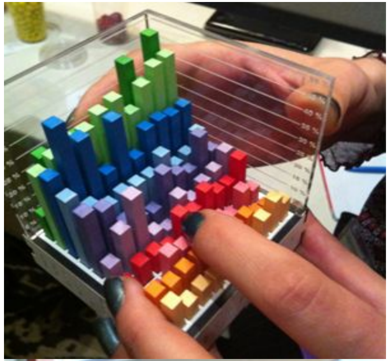


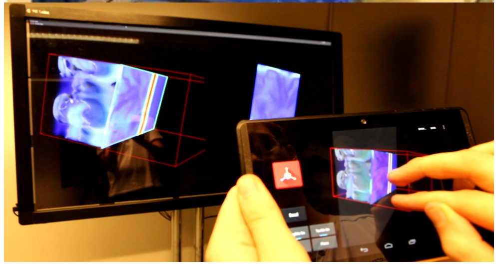
2 Physicality in Input and Output

A. Scientific Foundations for Data Physicalization

B. Situated Data Representations for Personal Analytics

C. Visualization using Augmented Reality Devices



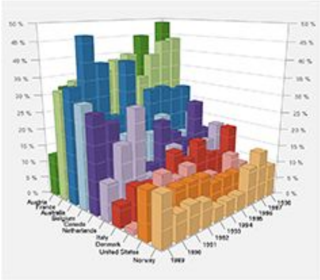



Zooids:
Building Blocks for Swarm User Interfaces

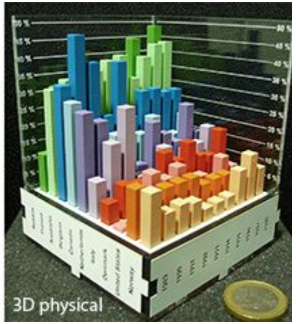
2A Scientific Foundations for Data Physicalization

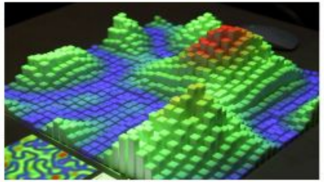
- What we did:
  - First empirical evaluation of physical visualizations
  - Robots for dynamic physical visualization (zooids)
  - Research agenda

- What we will do:
  - Study the underlying mechanisms, and the situations where physicality may be beneficial
  - Focus on benefits of physical data manipulation



3D on-screen





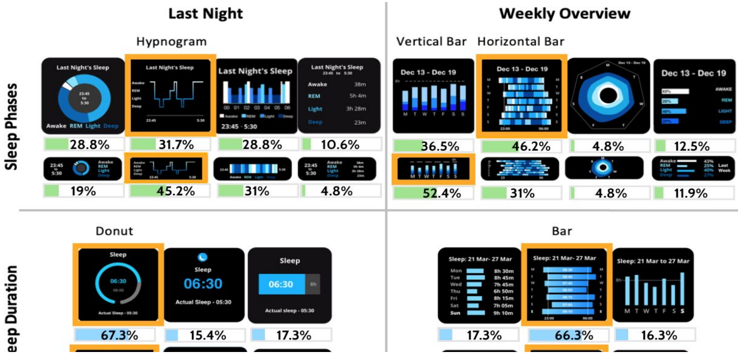
2B Situated Data Representations for Personal Analytics

What we did:

- Defined the concept of situated data representations
- Background research on fitness trackers / noise capturing

What we will do:

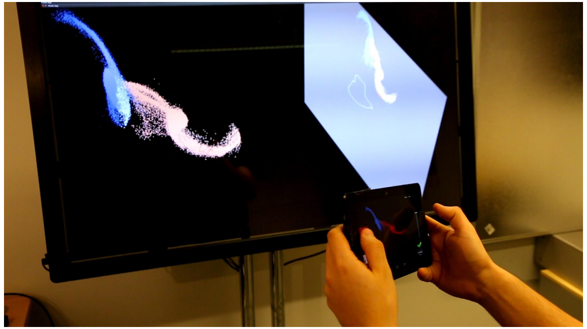
- Explore applications in personal analytics (EMBER ANR project)
- Explore micro visualizations for situated visualization (MicroVis ANR project)



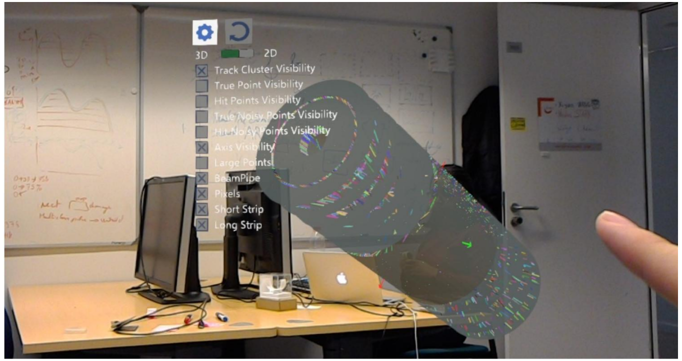
2c

Visualization using Augmented Reality Devices

- visualizations that provide stereoscopic vision or control, without removing the real world; **hybrid interaction** scenarios



- exploration of 3D navigation, precise control, combinations of traditional interfaces with AR representations



CHI 2020

Towards an Understanding of Augmented Reality Extensions for Existing 3D Data Analysis Tools

Xiyao Wang[1], Lonni Besançon[2], David Rousseau[3],
Mickael Sereno[1], Mehdi Ammi[4], and Tobias Isenberg[1]

1. Université Paris-Saclay, CNRS, Inria, LRI
2. Linköping University
3. Université Paris-Saclay, CNRS, IJCLab
4. University of Paris 8

3 Perception, Cognition, and Decision Making

A. Perceiving Statistical Information from Visualizations
   - Don't inspect each visualized item individually
   - How can people better understand “the gist”?

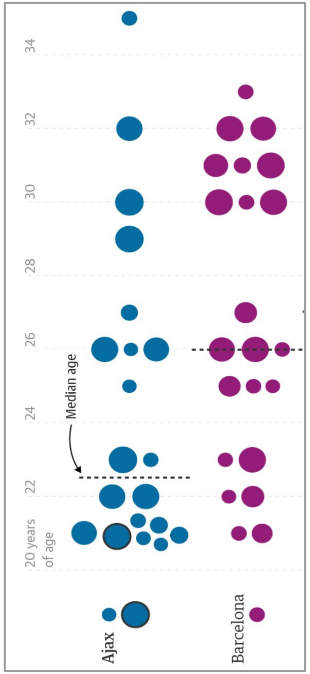
B. Decision making
   - Decision making and cognitive biases are important research topics in psychology, economics and marketing
   - How can visualizations support decision making?
   - Explainable machine learning (organized workshops)

C. Illustrative Visualization
   - Takes inspiration from illustrators’ experience
   - How can we support these styles with visualization?

3A Perceiving Statistical Information from Visualizations

- What we did:
  - Measured single-item comparison with higher precision
  - Found accuracy and bias difficulties with set comparison

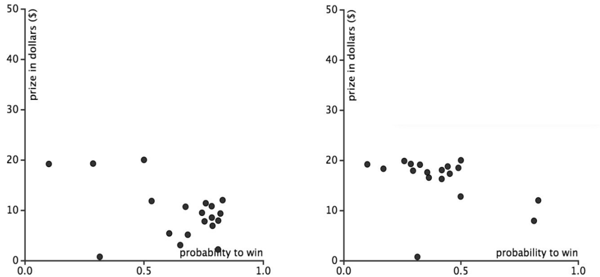
- What we will do:
  - Understand causes of difficulties when comparing sets
  - Predict when and why people will have these difficulties when using graphs

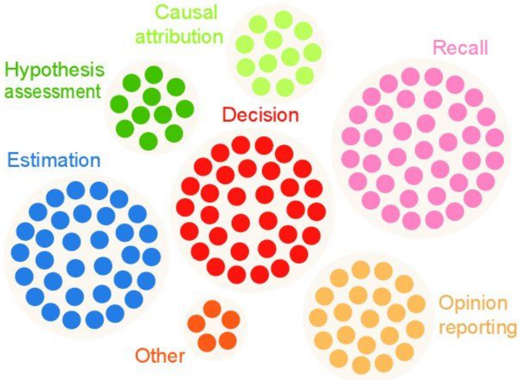


3B Visualization-supported decision making

- What we did:
  - First empirical demonstration of a cognitive bias in visualization
  - Taxonomy of cognitive biases for visualization research

- What we will do:
  - Study group decision-making & social choice with visualizations
  - Decision making in the presence of uncertainty and incomplete information
  - Visualizations to promote empathy and prosocial behavior (collab with Brazil)

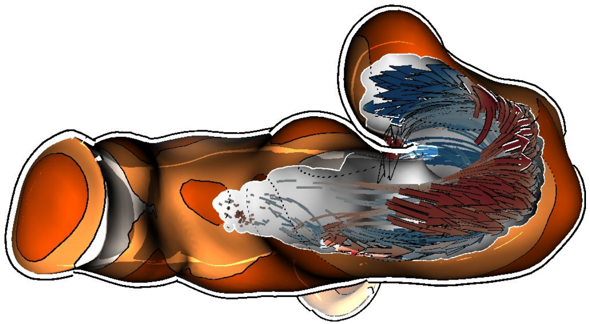


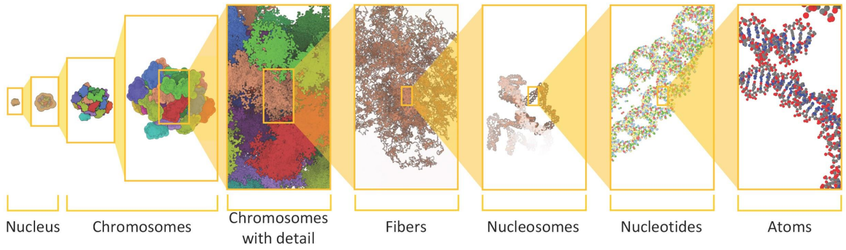


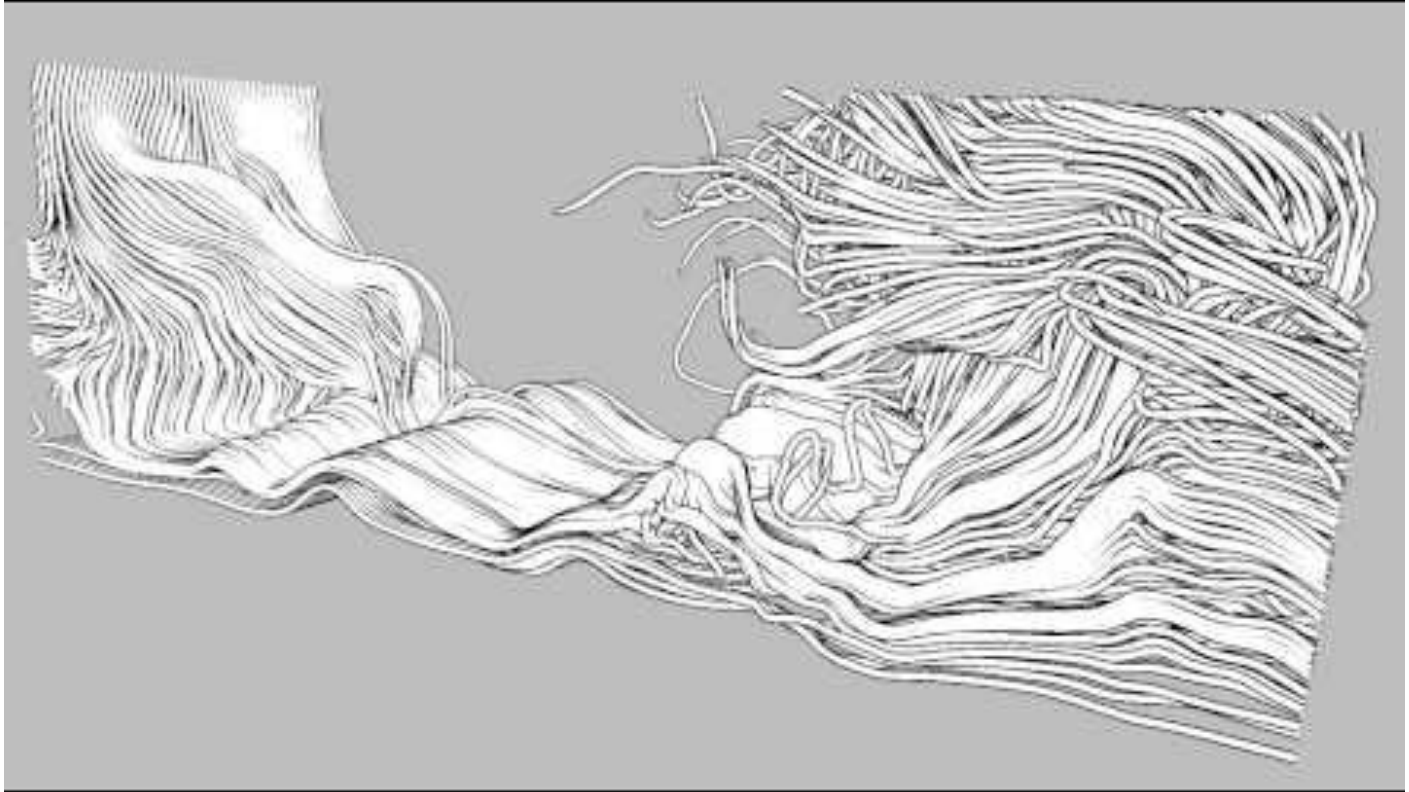
3c

Illustrative Visualization

- learn from existing illustrations & professional illustrators
- often for spatial representations of 3D data, but also some for abstract data
- examples: medical and genome visualizations
- insights on use of abstraction in visualization







4 Methodologies for Visualization Research

A. Promoting and Following Open Research Practices
   - Aviz researchers engaged in open research practices
   - Active at teaching and steering our field towards more transparent practices
   - Collecting and reporting information on prominence of open research practices

B. The Communication of Statistics Results
   - Aviz researching inaccurate statistical results in existing visualization literature
   - Many problems related to inappropriate reporting of statistical results in science, e.g., "dichotomous inference"
   - Aviz will conduct research on how to design statistical charts expressing nuances and uncertainty

C. Shaping the Visualization Community
   - Aviz researchers are heavily involved in the organization structure of visualization conferences and journals
   - Helping newly created committees to restructure the community based on data.

Prix internationaux

- Jean-Daniel Fekete
  - IEEE VGTC Technical Achievement Award
  - IEEE VGTC Visualization Academy
  - ACM SIGCHI Academy
- Catherine Plaisant
  - IEEE VGTC Visualization Career Award
  - IEEE VGTC Visualization Academy
  - ACM SIGCHI Lifetime Service Award
  - ACM SIGMM in Test of Time Awards 2020 (for year 2004)
- Tobias Isenberg
  - Associate Editor of the year (2019), Elsevier Computers & Graphics
- Petra Isenberg
  - Top 10 most cited scholars from the top venues of this field in the past 10 years (#7)
- Steve Haroz
  - ACM SIGCHI 2020 best paper award
- AVI 2020 best paper honorable mention award
- 2 IEEE VIS 2020 poster design awards
- VAST Challenge: Outstanding Comprehensive Mini-Challenge 1 Solution

Collaborations in 2020

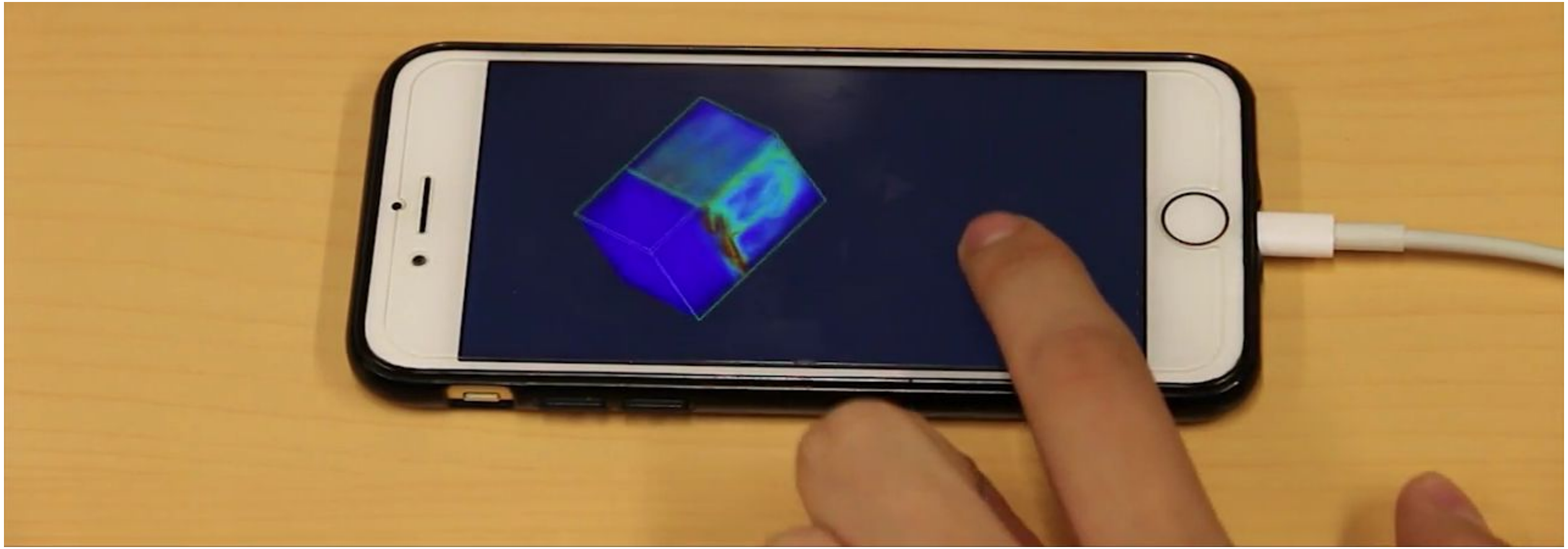
- Leading industries and universities
  - Microsoft Research
  - *North American U's:* U. of Calgary, NYU, U. of Maryland, Northwestern, UofT, Georgia Tech, Ohio State U.
  - *European U's:* TU Wien, U. of Stuttgart, U. of Potsdam, KU Leuven, Linköping U., U. of Vienna, U. of Nottingham
  - CERN, Monash U. (Australia), U. of Canterbury in Christchurch (NZ), KAUST (Saudi Arabia)
  - China: University of Hong Kong, University of Qingdao, Xi'an Jiaotong-Liverpool U.
- Other INRIA Project-Teams
  - Ilda, EX-SITU, TAU, Hybrid
- French Organizations and Companies
  - Telecom Paris: DataIA HistorIA
  - EHESS: Exploration of large dynamic social networks
  - IRT-SystemX: Visualizing the Blockchain(s) and Bitcoin
  - EDF: Interaction in 3D, Visual Analytics, and Progressive Data Analysis
  - InraE-MaIAGE: Interactive visualization of cell lineage (Naviscope project)

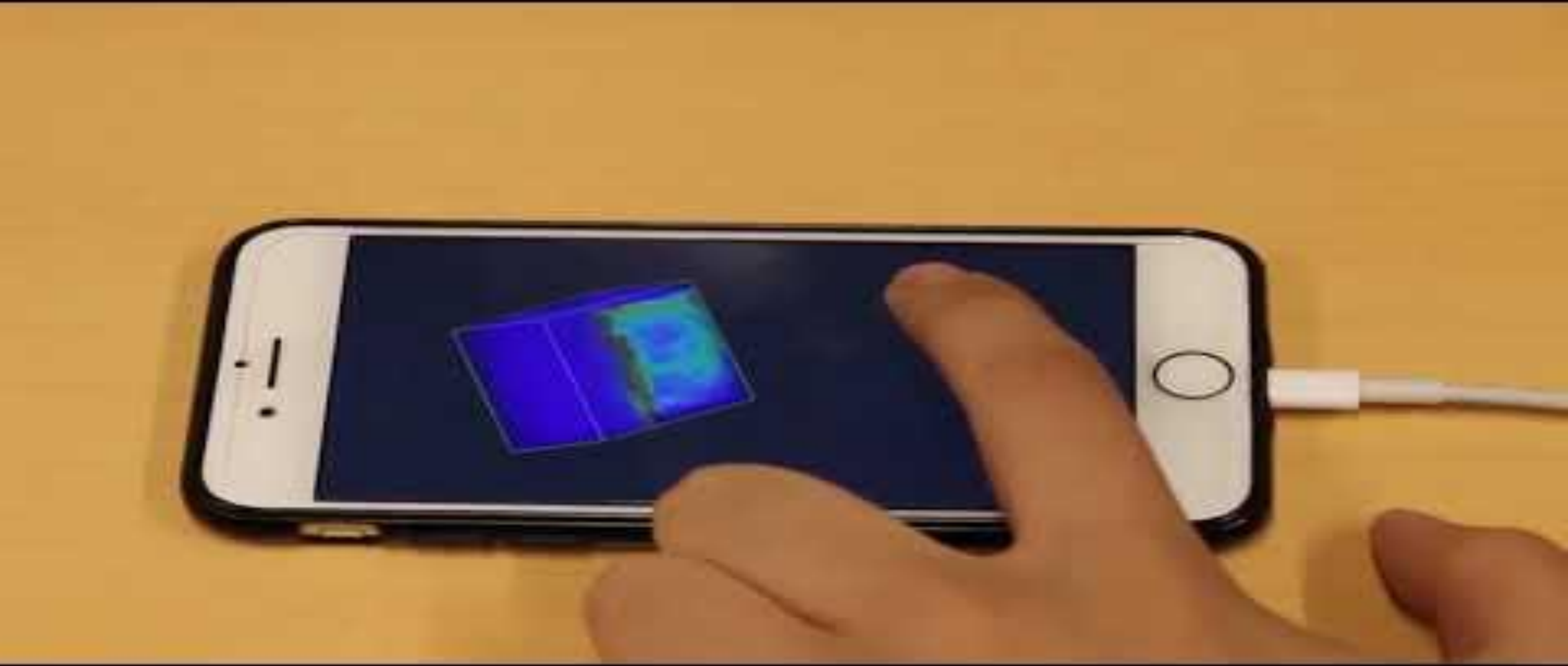
Publications

- 42 new articles/papers published/accepted 2020–2021
  - 18 journal articles (IEEE TVCG, CGF, ...)
  - 8 conference papers
  - 2 book chapters
  - 14 other

Past Internship: Pressure-based 3D Navigation on Phones

- Xiyao Wang, now PhD student (about to defend)
- problem: 3D navigation on phones without occluding much of the small display
- need 7 DOF: translation in x, y, z; rotation around x, y, z; and uniform scaling





Master Thesis Defense

Visualization for Academic Hiring

HONG Yumin
09/09/2019

Supervisors
Pierre Dragicevic, Petra Isenberg

Motivation

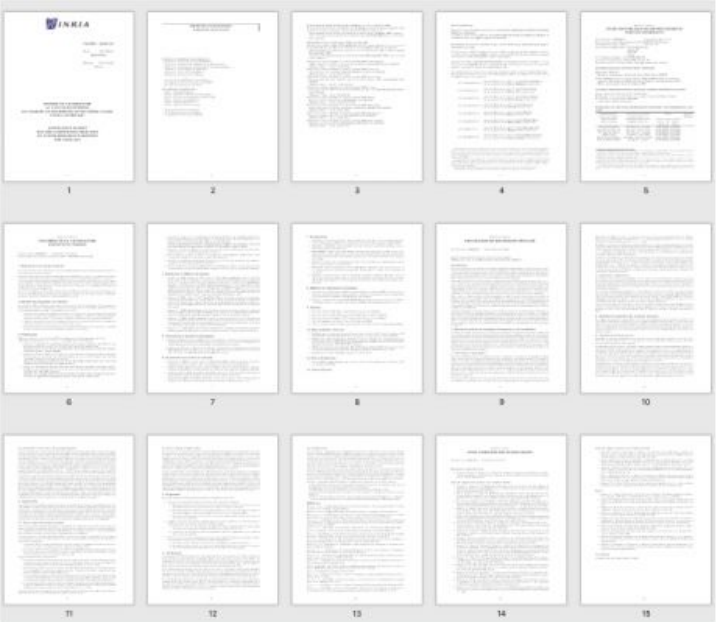


Figure. Application Document Example from an INRIA applicant

1

Rich, complex and unstructured information

2

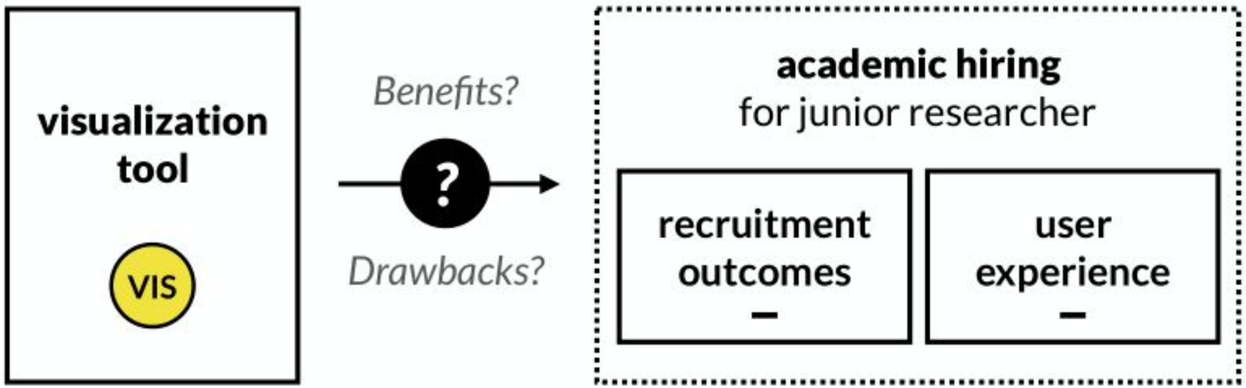
Lots of efforts and time required when reviewing

3

No research or visual tool designed for helping with academic hiring

Research Objective

Whether a visualization tool can improve
not only **recruitment outcomes**
but also **user experience**



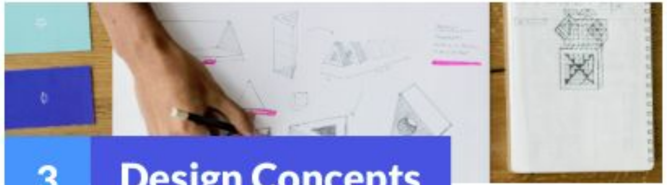
Key Contributions

1 Literature Review

A review of current candidate evaluation methods and visualization approaches in the context of academic hiring

2 User Analysis

An analysis of user scenarios and problems in the academic hiring process



3 Design Concepts

Several visualization design concepts meant to improve the quality of academic hiring

4 Discussion

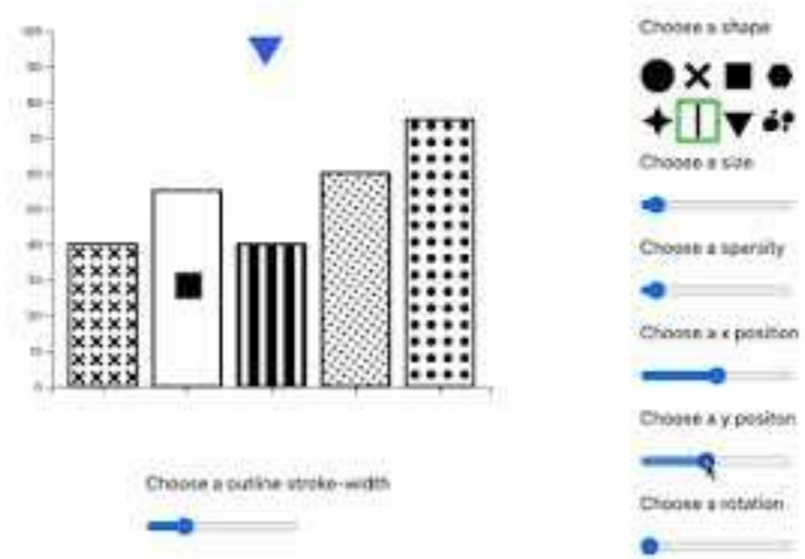
Discussion about possible benefits of visualization in academic hiring

Final Design

Studying black and white textures for visualisation on e-ink displays

Supervisor: Jamy Li (Univeristy of Twente)
Advisor: Petra Isenberg (INRIA)

Yuanyang Zhong 25/09/2020


Choose a shape
Choose a size
Choose a sparsity
Choose a x position
Choose a y positon
Choose a rotation
Choose a outline stroke-width

Black-and-White Textures for Visualization on E-ink Displays

Yuanyang Zhong*
Université Paris-Saclay, CNRS, Inria, LRI, France

Tobias Isenberg†
Université Paris-Saclay, CNRS, Inria, LRI, France

Petra Isenberg†
Université Paris-Saclay, CNRS, Inria, LRI, France

Abstract

We introduce a design space for black and white textures and the prototype of a design tool that supports designers to explore different black-and-white textures for visualization on e-ink displays. We are currently working on an exploratory online study in which experts will design four different categorical encodings using different types of black-and-white textures.

Index Terms: Human-centered computing—Visualization—Visualization systems and tools

1 Introduction and Background

E-ink displays are now widely used on devices such as smartwatches, electronic shelf labels, and e-readers. E-ink displays are popular for their low power consumption, relatively low cost, and reduced eye strain when reading. As a tradeoff, they have only a few colors and low refresh rates. While research and development for e-ink displays has increased the number of displayable colors, most e-ink displays are still black-and-white or have only a few additional colors that could be used for visualization. Yet, many types of charts require color hue or saturation to encode quantitative or categorical data. With our research we want to understand how we can effectively visualize data in the absence of color. We explored several strategies and, inspired by historical practice, focused on texture encodings.

We first explored alternatives to the use of color hue or saturation in data graphics, then derived a design space of black-and-white texture dimensions, and finally designed a tool to explore different types of textures to encode categories on bar and pie charts. Our ultimate goal is to study the features of black-and-white textures to come up with guidelines that help designers build beautiful and effective data visualizations specifically for e-ink displays.

Throughout our discussion we use the term *texture*, following definitions in computer graphics where a texture is a 2D image applied to (the surface of) a graphical object that is potentially repeated multiple times through some form of tiling. Specifically, here we refer to a texture as a tiling pattern that is characterized by the type of shapes we use as part of the pattern (lines, dots, …), the density of the shapes, their size, orientation, etc.

The challenge of a limited number of displayable colors has been present in the use of virtually all printing technology as well as the early days of computer screens. Many historical examples exist that showcase beautiful and stunning data graphics, drawn without color. Brinton [2, p.32], e. g., shows an over 100 year old example of a heatmap (Fig. 1) that uses black-and-white textures. Bertin's Semiology of Graphics [1] discusses textures extensively, for points, lines, and areas and with various graphical properties. Bertin uses the term *texture*, however, to refer to differences in density applied to a shape[1] (e. g., vs.) and the term *pattern* to instead refer to differences in shapes applied to a mark (e. g., vs.). Here, we use a slightly different terminology as outlined above but Bertin's advice on the use of his visual variables *texture (density), orientation, and shape* was valuable for our work. Both Tufte [7] and Bertin [1] warn that textures may produce Moiré vibrations and Tufte [7] even goes so far to say that they should be avoided altogether.

Some examples of using Tufte's and Bertin's advice for black-and-white visual representations exist, such as Perin et al.'s Bertifier [6]. In computer graphics, black-and-white textures are also used, usually in the sub-field of non-photorealistic rendering such as for stippling [3,4] or hatching [5]. Here the goal is, however, the replication of traditional drawing techniques that represent shading and material properties, as opposed to data representation like in our case.

Figure 1: Rank of US states in ten education features, 1910. Taken from [2, p.32], trimmed to 12 states. Picture in the public domain (©).

2 Texture Features

In data visualization, textures can be applied to both the stroke and the fill of a data-representing mark. We discuss textures as repeated (tiled) shapes, such as lines or dots. Shapes can be applied both to the fill of a mark or along its outline. We detail the texture properties we found useful to configure black-and-white textures with examples in Table 1. In particular, we considered:

Density: how closely packed shapes inside the texture are while maintaining the same black-to-white ratio (value). That is, if more black shapes are to cover a certain mark, they will be reduced in size to maintain the same black/white ratio.

Shape Size: the size of the shape inside the texture. Increasing the size while leaving density the same will result in a change in value (black/white ratio). Changing size and density together allows keeping the black/white ratio (value) the same.

Orientation: orientation/angle of shapes inside the texture

Shape type: shape type used inside the texture (line, dot, square, …)

Grouping: This texture property is not a visual variable listed by Bertin. It describes how shapes are placed on a mark. For example, in a regular grid, with an offset, randomly etc. Changes in grouping result in a visually different texture but do not change overall value, and—depending on the grouping—also not the perception of density.

From the descriptions of texture parameters it becomes clear that choosing correct parameters for visualization is not a trivial task. Which we further explored for categorical data.

3 Visualizing Categorical Data for Black and White Textures

Past work in visualization [1] and in particular computer graphics [4, 5] includes advice on how to generate textures with a given value encoding that could be used for the visualization of ordered

*e-mail: ajax17@126.com
†e-mail: {tobias.isenberg petra.isenberg}@inria.fr

[1] To be precise: To Bertin, changes in texture only affect the number of separable marks in an area and not the overall value.

Best poster design award IEEE VIS 2020

Interactive Version

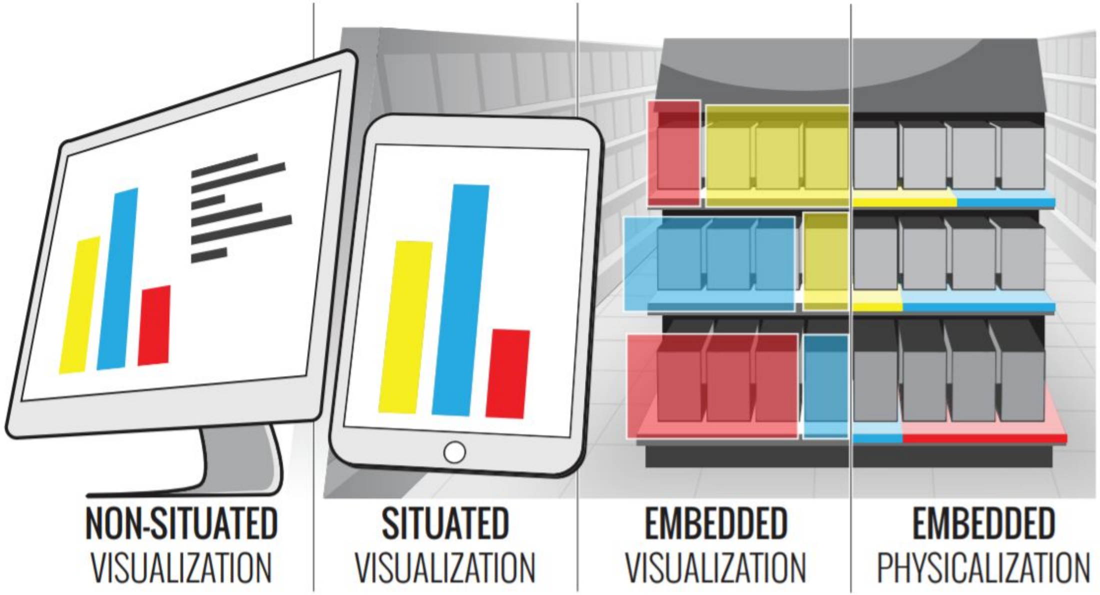
Visualization in Motion

02/03/2020 - 31/08/2020

Student: Lijie Yao (Now Ph.D. Student)
Supervisors: Petra Isenberg (INRIA)
Anastasia Bezerianos (LRI)

Background Introduction

- Non Situated Visualizations
- Situated Visualization
- Embedded Visualization
- Embedded Physicalization

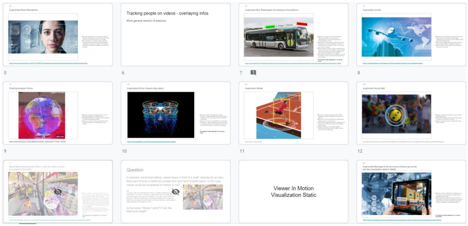


Embedded data representation (Wesley Willett, 2016)

Specific Topic

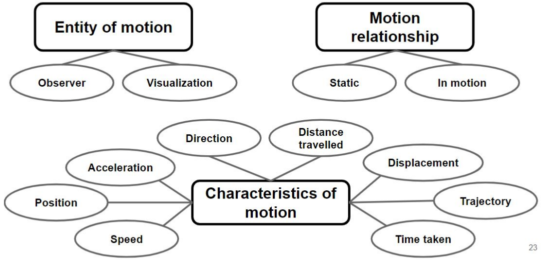
- Understanding characteristics and constraints of visualization under movement

Research Approach

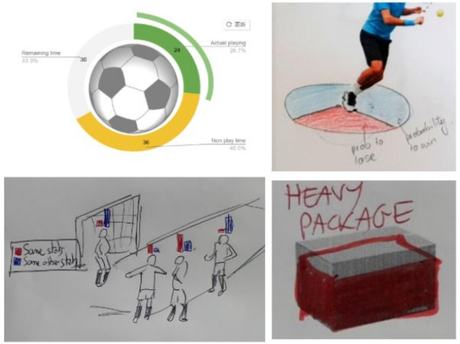


1 - Sample Collection

2 - Literature Review



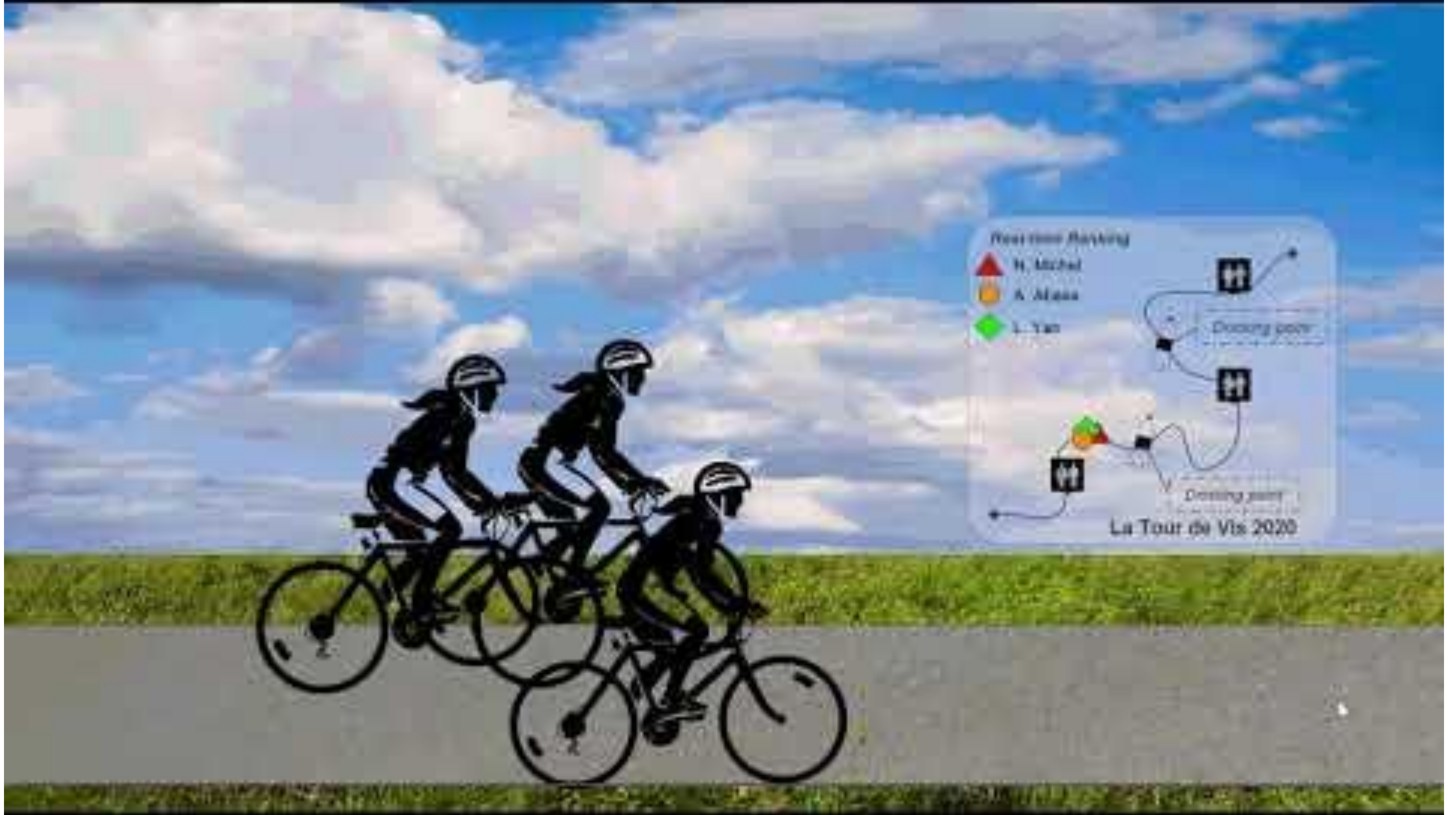
3 - Design Space



4 - Online Workshop

5 - Exploration Tool

6 - Crowdsourcing Experiment


N. Michel
La Tour de Vis 2020

Situated Visualization in Motion

Lijie Yao*
Université Paris-Saclay, CNRS,
Inria, LRI, 91405 Orsay, France.

Anastasia Bezerianos†
Université Paris-Saclay, CNRS,
Inria, LRI, 91405 Orsay, France.

Petra Isenberg*
Université Paris-Saclay, CNRS,
Inria, LRI, 91405 Orsay, France.

Figure 1: Three examples of visualizations viewed under motion. Left: Visualization of rotor speed and battery level attached to a flying drone. Middle: An athlete running past a sign showing the elevation levels of the trail ahead. Right: A player with an augmented soccer ball

Abstract

We contribute a first design space on *visualizations in motion* and the design of a pilot study we plan to run in the fall. Visualizations can be useful in contexts where either the observation is in motion or the whole visualization is moving at various speeds. Imagine, for example, displays attached to an athlete or animal that show data about the wearer – for example, captured from a fitness tracking band; or a visualization attached to a moving object such as a vehicle or a soccer ball. The ultimate goal of our research is to inform the design of visualizations under motion.

Index Terms: Human-centered computing—Visualization—Visualization systems and tools

1 Introduction

Our research contributes to the area of situated data visualization [7]. In a situated data visualization, the data is directly visualized near the physical space, object, or person it refers to (=the data referent). Situated visualizations can often be found in contexts in which the data referent or observer does not remain static but is moving at various speeds. Imagine, for example, an athlete running with a fitness band on which he/she checks current progress or a visualization overlayed on a soccer or basketball to show game statistics. In these mobile and dynamic use cases, situated visualizations have to overcome new challenges if we want the data to be readable in real-time by the human observer. That is, we have to find effective and visually stable situated data encodings that are readable under various types, directions, and speeds of movement. In the first part of our research we are developing a design space of visualization movement types while considering the relationship to the human observer. In the continuation of our research we will explore different types of display modalities including wearable devices, augmented reality, and physical visualizations to realize concrete prototypes of situated mobile visualizations. Our next goal is to contribute empirical studies to assess how different design factors and types of movement influence the perception of situated visualizations.

*e-mail: {lijie.yao | petra.isenberg}@inria.fr
†e-mail: anastasia.bezerianos@lri.fr

2 Related Work

A large body of literature in the HCI community looked at how to engineer wearable devices and studied their use–sometimes under locomotion. Much of this work has focused on usability aspects and the possibility to *interact*, rather than perceive information accurately. For example, Schneegass and Voit [6] looked at the use of smart watches while in motion, analyzing how participants interacted with a GestureSleeve while running. Their usage scenario was the starting, pausing, and stopping of a fitness tracking app.

We are not aware of studies in the Visualization community that offer any design advice for data-viewing contexts that involve motion of the whole visualization. However, there have been results for contexts in which the display itself was static but viewers were able to move. A study on the perception of visual variables on a large high-resolution wall-sized display [2], for example, tested a condition in which participants were allowed to move to see information under a chosen viewing angle and distance. No improvement in accuracy of the task was found but a steep increase in task time due to movement. Another related research area is proxemic interaction in which interfaces are controlled—at least partially—by movement in front of the display (e. g., [1, 5]). In summary, there is currently a lack of perceptual guidance on the efficacy of different types of visualizations and their use under motion and locomotion. This allows us to offer new research results on this emerging usage context for situated visualizations.

3 A Design Space for Visualizations and Motion

To describe the research space more concretely we started by deriving important dimensions to consider when we discuss visualizations that are either moving themselves or are read by a moving observer. Note, that here we do not consider visualizations where only specific marks are in motion—as is the case with most visualizations that contain animated transitions, e. g., animated scatterplots [3].

The current dimensions of our design space are:

Entity in motion: This dimension includes two values that capture which entity exhibits physical motion, either the *viewer* or the *visualization*. For example in Fig. 1 the drone is in motion while a viewer controlling the drone on the ground would

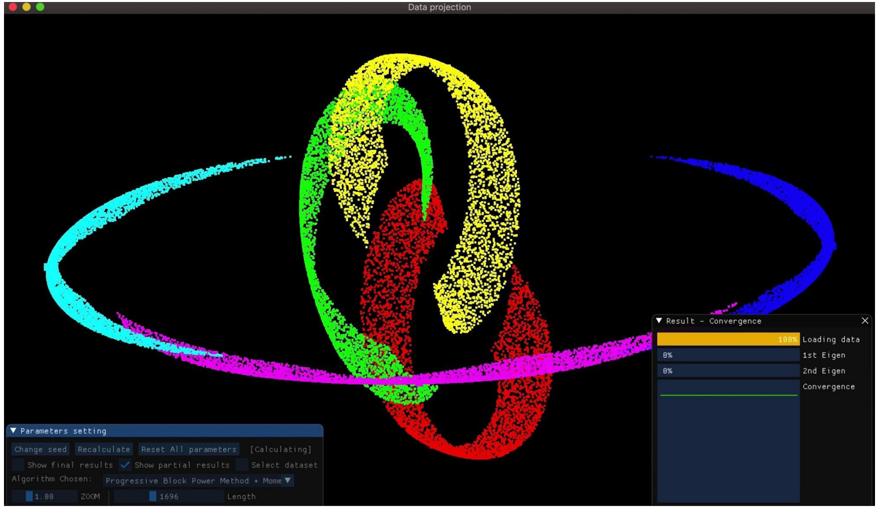
IEEE VIS 2020 Poster

[Interactive Version]

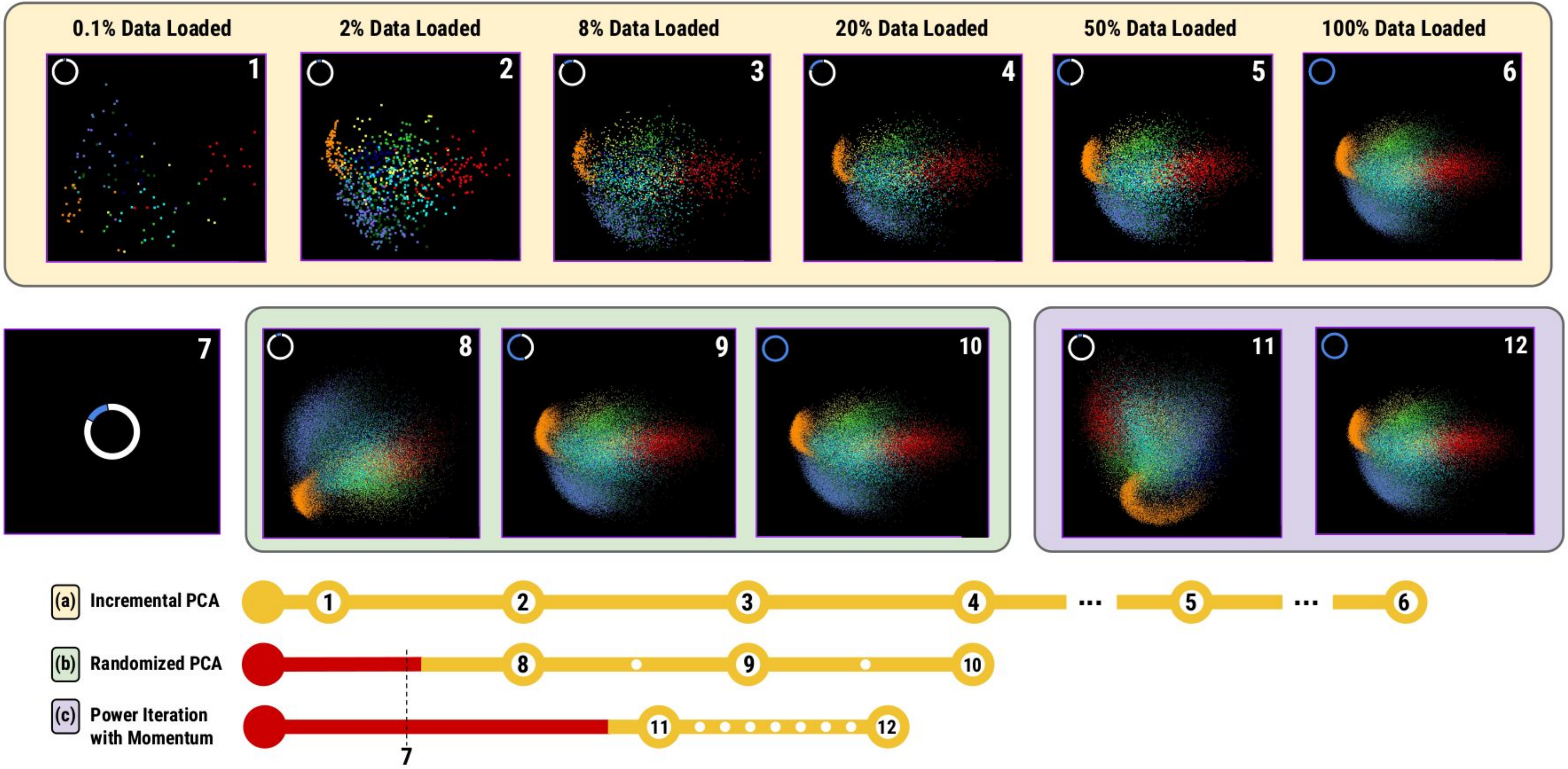
Progressive PCA for Massive Time-Series

Yuhen Feng, Alejandro Ribes, Jean-Daniel Fekete

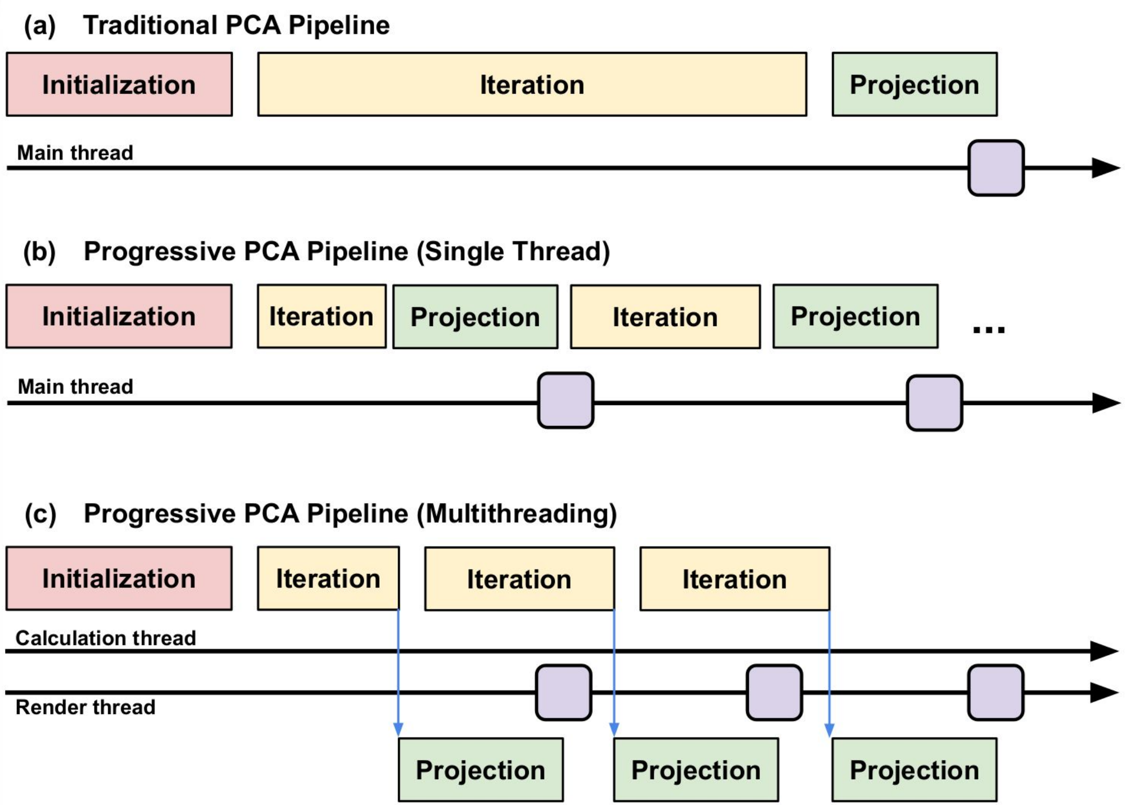
- EDF needs to forecast the behavior of complex hydraulic systems over time
- Ensemble simulations generate multiple time-series
  - Thousands of time-series
  - Tens of thousands of dimensions
- Exploring them with existing systems is not interactive



Progressive PCA: Several Trade Offs



Evolution of the Visualization Pipeline



Progressive PCA

Three options:

- IncrementalPCA
  - Loads data incrementally and can compute PCA progressively on very large datasets
  - Results start immediately
  - Mediocre quality
  - Overall slow
- Randomized PCA
  - Needs all the data first
  - Starts a bit slower
  - Good quality
  - Converges faster
- Power Iteration with Momentum
  - Starts slowly
  - Fast, excellent quality

Use a hybrid solution:

- Start the Incremental PCA
- When the data is loaded...
- In parallel, start Power Iteration or Randomized PCA
- Show Incremental PCA until the other catches-up

Understanding Micro Visualizations for Fitness Trackers

Ranjini Aravind

Supervisors:

Petra Isenberg, Tanja Blascheck

Introduction

111 million fitness trackers (smartwatches, wristbands, and sports watches) sold worldwide in 2018

- Step count
- Heart rate
- Calories burned
- Sleep duration

Our Motivation

What we know

Current use of fitness trackers

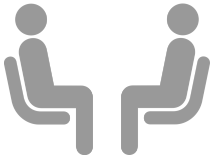
Perception of standard charts on smartwatches

What we wanted to find out

How well could fitness tracker wearers read **visualizations** on small screens

Methodology

Understand how wearers use fitness tracker visualizations



INTERVIEW

Understand how wearers perceive sleep visualizations

QUESTIONNAIRES

Compare readability of sleep visualizations on trackers

PERCEPTUAL EXPERIMENT

A Survey on Sleep Visualizations for Fitness Trackers

Ranjini Aravind[1,2], Tanja Blascheck[1,3] and Petra Isenberg[1]
[1]Inria, France
[2]Université Paris-Sud, France & Université Paris-Saclay, France
[3] University of Stuttgart, Germany

Abstract

We contribute the results of an exploratory study and a survey on visualizations for fitness trackers. Fitness trackers are becoming ubiquitous trackers of personal data. They often come with small attached displays that show micro visualizations of data such as heart rate, step counts, sleep duration, or number of floors climbed. Unfortunately, little is known about how wearers of fitness trackers use and perceive these micro visualizations. To collect data on the use of fitness visualizations, we conducted ten personal interviews with regular wearers of fitness trackers. Inspired by frequent responses regarding sleep tracking, we deployed an online questionnaire specifically on sleep visualizations for fitness trackers. Our results show that most participants were interested particularly in seeing previous night's sleep data on their fitness trackers and preferred visualizations that were easy to read like the hypnogram and bar as well as donut charts for sleep phases and duration.

CCS Concepts

• ***Human-centered computing*** → *Empirical studies in visualization; Visualization design and evaluation methods;*

1. Introduction and Related Work

In 2018, 111 million fitness trackers (smartwatches, wristbands and sports watches) were sold worldwide and sales are expected to double by 2022 [Gar18]. Most fitness trackers show micro visualizations representing tracked data such as heart rate, step count, sleep duration, or number of floors climbed. The challenges with fitness tracker visualizations are the small screens and the public nature of the interfaces [GPK*16]. A recent study on smartwatch visualizations [BBB*19] found that bar and donut charts allow quicker data comparisons than radial bar charts. However, there is still a lack of general advice on how to design visualizations for fitness trackers.

Our work takes inspiration from previous research that studied visual data exploration of fitness data on larger form factors, like desktops and smartphones. A study on self reflection [CLZ*17] identified two exploration patterns. First, visual cues, like peaks or extremes, prompt wearers to recall their past behaviors. Second, wearers recall their past behaviors to come up with a question and explore their data. These two patterns include tasks like recalling detail, looking for detail, comparison of time segmented values, identifying trends, making value judgments, finding a distribution, correlation, outlier, or summary, and making a prediction [CLZ*17]. The tasks that wearers perform depend on the design of the visualization and choices in visual variables (e. g. position, color), mapping variables (e. g. extremes, averages) and computational variables (how aggregate data was computed) [ACG14]. A crowd sourcing study [BLIC19] on a mobile phone compared two layouts (linear and radial) representing sleep and temperature data. The authors found that linear layouts allow quicker value reading, value comparison, and range comparison compared to radial charts. Chen [Che17] proposed a smartwatch visualization design for large time series which supports finding specific time-dependent patterns and trends. In contrast to this past work, we contribute the results from two studies specifically about the use of visualizations on fitness trackers.

2. Interviews and Card Sorting

To gain first insights into the space of visualization use on fitness trackers, we conducted an exploratory study consisting of personal interviews and a card sorting exercise, which shed light on the priorities of the wearers when they look at fitness visualizations.

Procedure. We conducted ten interviews with regular wearers of fitness trackers recruited via email in our institution. Each interview, on average, lasted around 30 mins. We asked participants details about their fitness tracker, which visualizations they remembered, and how their use of the associated phone or web app differed from their use of the fitness tracker. For the card sorting exercise [WR88], we gave participants 33 different smartwatch-sized data visualization cards (see OSF). We asked participants to organize these cards into self-defined groups without using visualization type or data type as a criterion. One author transcribed all interviews, categorized the answers to our questions, generated a similarity matrix and a hierarchical clustering dendrogram from the card sorting groups.

Results. The most common data that participants mentioned checking on their fitness tracker included step count (10/10 ×), distance (7/10 ×), and heart rate (4/10 ×). On the associated phone or web

EuroVis 2019 Poster
Publication coming soon

How do users of fitness trackers perceive visualizations on their tracker?

Pic by Tim Foster on Unsplash

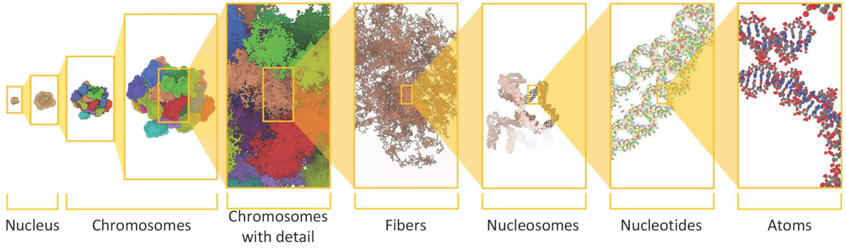
Internship Topics

List of internship topics

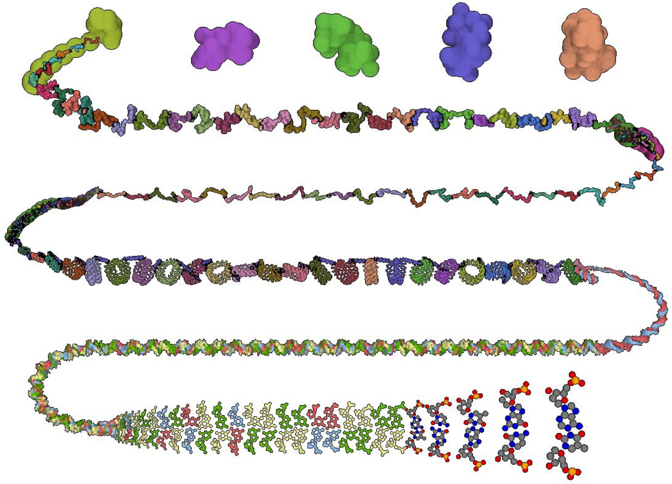
- Tobias: Hybrid 3D Genome Visualization
- Jean-Daniel: Visualization of Multimodal Brain Connectivity
- Jean-Daniel: Advanced Network Visualization in the Jupyter Notebook
- Jean-Daniel: Scatterplots to Infinity [and Beyond]
- Natkamon: Characterization of Entities in the Bitcoin Blockchain
- Natkamon: Analysis and Visualization of the Bitcoin Forum
- Alaul: Data Representations for Smartwatch Screens

Hybrid 3D Genome Visualization

supervised by Tobias Isenberg
tobias.isenberg@inria.fr



- existing implementations for both temporal and spatial control of scale for 3D genome data
- **internship goal:** combined implementation that allows one to switch from one to the other, essentially to “pivot” between both representations
- **requirements:** knowledge of 3D graphics coding, C++
- based on existing software framework Marion
- collaboration with TU Vienna & KAUST
- scientific publication planned
- adjusted topics in this context are possible



ScaleTrotter: Illustrative Visual Travels Across Negative Scales

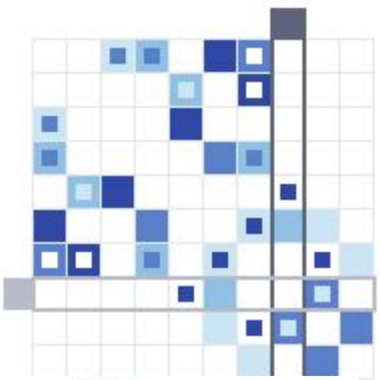
Sarkis Halladjian, Haichao Miao, David Kouřil, M. Eduard Gröller, Ivan Viola, Tobias Isenberg

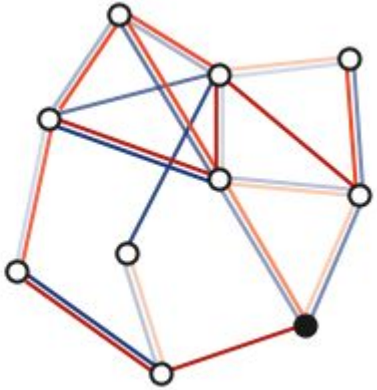
Visualization of Multimodal Brain Connectivity

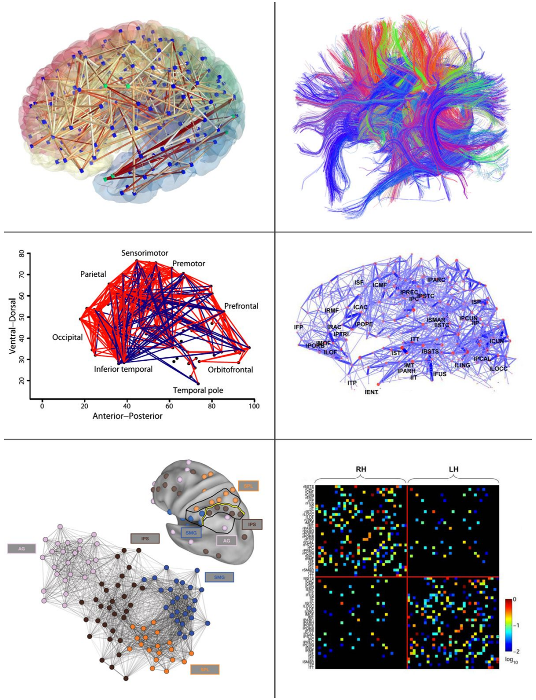
supervised by Jean-Daniel Fekete
Jean-Daniel.Fekete@inria.fr

- Brain connectivity can be computed through several modalities:
  - Fiber tracts (physiological connectivity)
  - fMRI (functional MRI)
  - EEG at various frequencies
- Each of these modalities provide a full matrix
  - Probability / count for fibre tracts
  - Correlation between signals for fMRI&EEG
- The matrices are thresholded to create a graph
- These graphs need to be compared to understand their differences
- These graphs vary with time
- **How can be best visualize them?**

Collaboration with Fabrizio de Vico Fallani, Inria & Brain and Spine Institute (ICM) in Paris



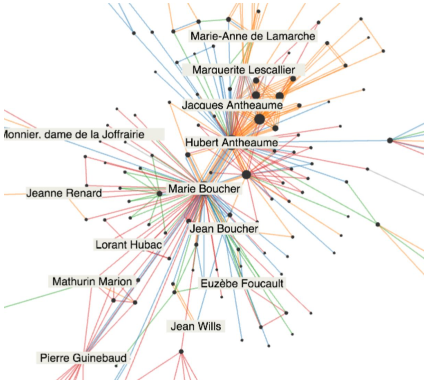


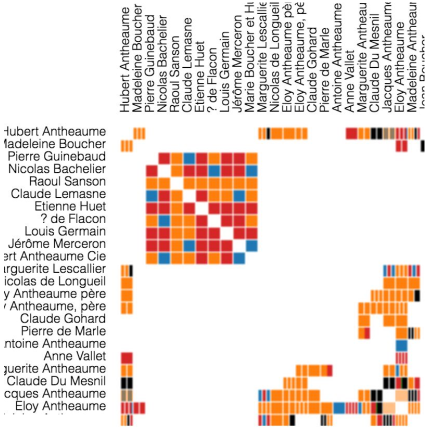


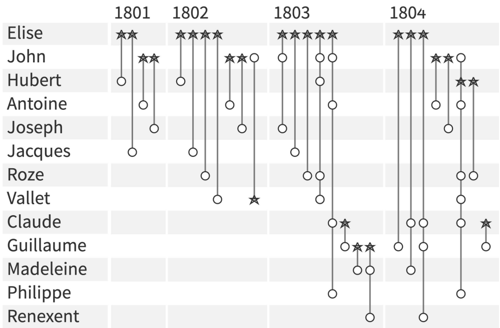
Advanced Network Visualization in the Jupyter Notebook

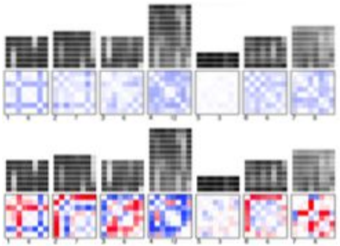
supervised by Jean-Daniel Fekete
Jean-Daniel.Fekete@inria.fr

- Aviz has designed many network visualizations
- But they live in a closed world where data scientists cannot use them from their favorite working environment
- We want to make them available in the jupyter notebook, popular in data science
  - Connected with NetworkX in Python
- What are the interaction and communication mechanisms needed to benefit from the best of both worlds: the notebook and interactive vis?
  - Coordinated views
  - Shared selection
  - Filtering on vertices and time
- Interact with the Python NetworkX community to get feedback.







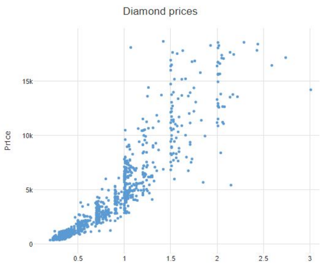


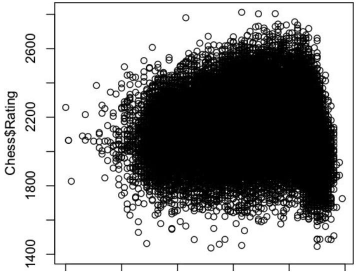
Scatterplots to Infinity [and Beyond]

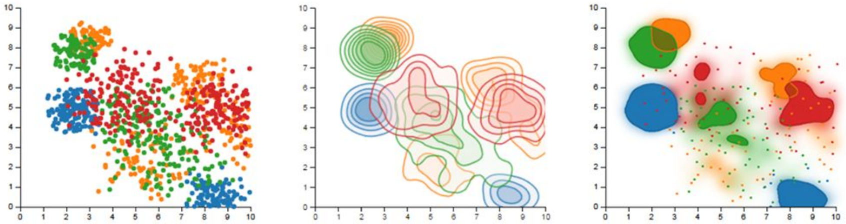
supervised by Jean-Daniel Fekete
Jean-Daniel.Fekete@inria.fr

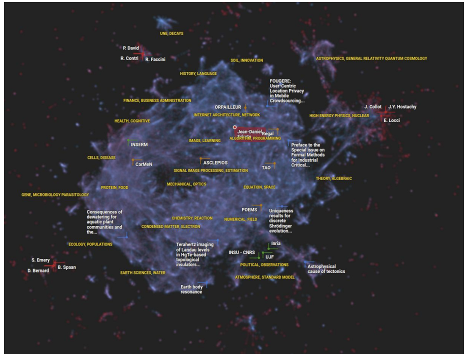
Scatterplots are used a lot in visualization, but they need special treatments to scale

- We want to be able to visualize as many points as we have, to infinity [and beyond]
- Compute a density map
  - Each pixel is a count or a density
- Use a progressive method:
  - Create the density map by loading the data by chunks, e.g. 100,000 at a time
  - Visualize intermediary results as they are computed
- Several issues:
  - The scale can change
  - Need to pan and zoom
- Python + JavaScript (browser, notebook or not?)





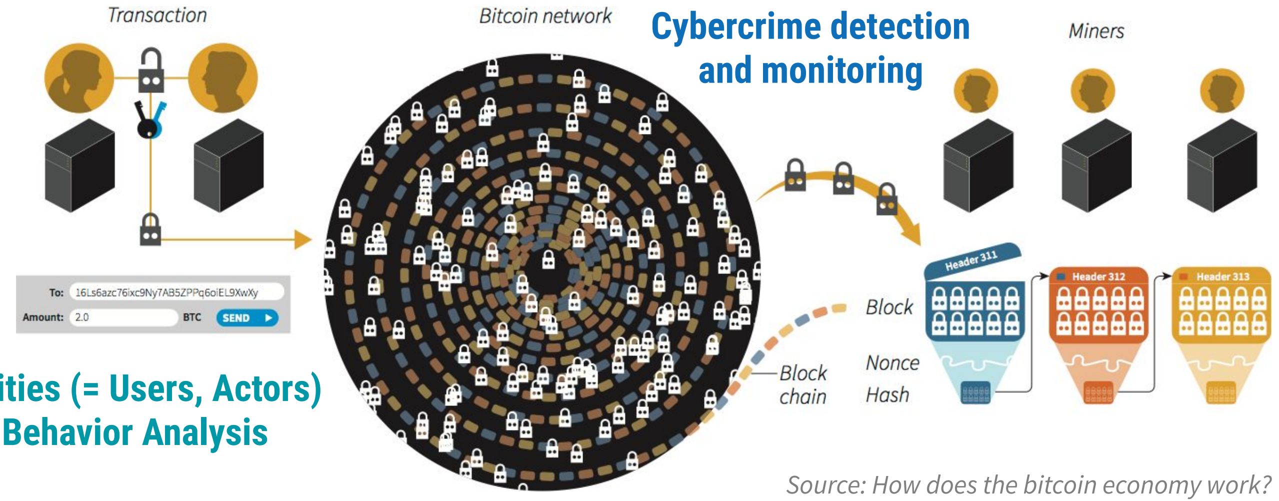




bitcoin Visual Analytics

supervised by Natkamon Tovanich
natkamon.tovanich@irt-systemx.fr
and Petra Isenberg

- Bitcoin is a cryptocurrency that stores transactions in a public blockchain.
- ~600 million transactions at this moment (>300 GB of raw data) and constantly growing.
- Transaction data is public but the owners of transaction (entity) are pseudonymous.



Cybercrime detection and monitoring

Mining Pool Analysis

Entities (= Users, Actors) Behavior Analysis

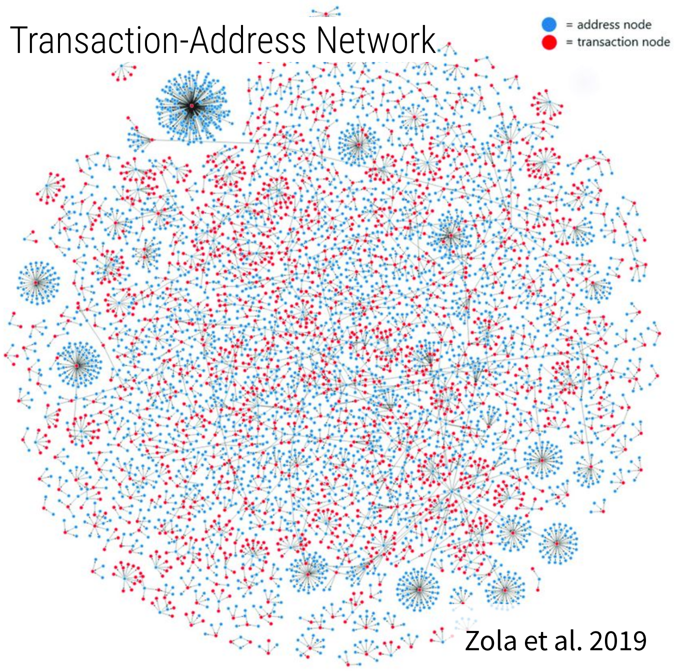
Source: How does the bitcoin economy work?
https://blogs.thomsonreuters.com/answerson/bitcoin-economy/

bitcoin Visual Analytics

supervised by Natkamon Tovanich
natkamon.tovanich@irt-systemx.fr
and Petra Isenberg

Internship Topic: Characterization of Entities (= Users, Actors) in the Bitcoin Blockchain

Transaction-Address Network



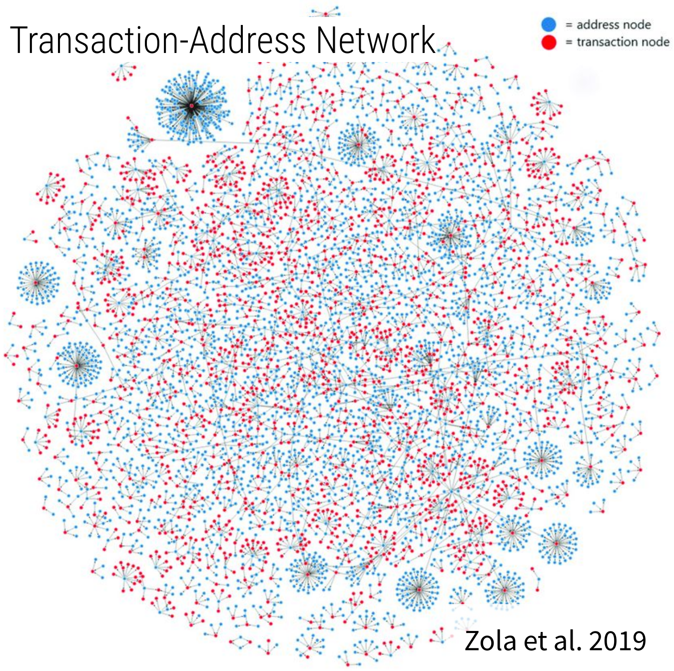
- How to identify the pseudonymous users from the large Bitcoin transaction data?
- What are the different types of Bitcoin users?
- Can we identify groups of entities based on the transaction and network activities?
- What is the general behavior of the entity types in Bitcoin?
- How has the activity of each entity type changed over the years?

bitcoin Visual Analytics

supervised by Natkamon Tovanich
natkamon.tovanich@irt-systemx.fr
and Petra Isenberg

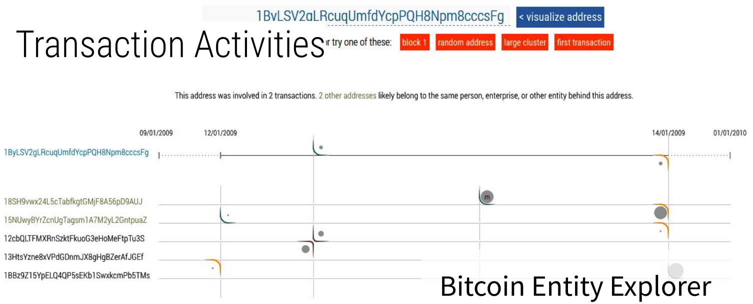
Internship Topic: Characterization of Entities (= Users, Actors) in the Bitcoin Blockchain

Transaction-Address Network



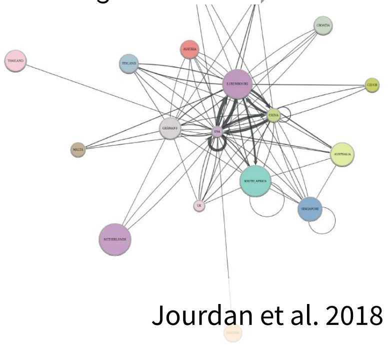
Zola et al. 2019

Transaction Activities



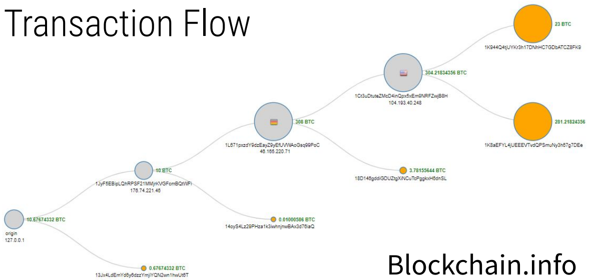
Bitcoin Entity Explorer

User Ego-Network



Jourdan et al. 2018

Transaction Flow



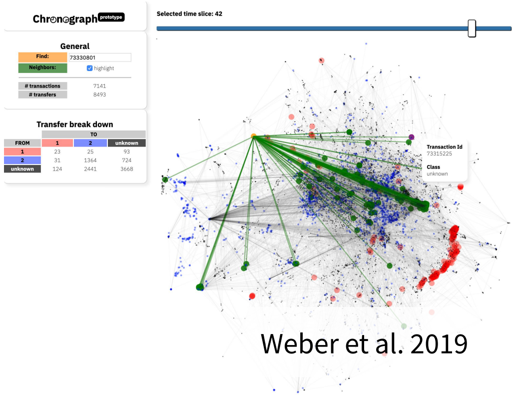
Blockchain.info

bitcoin Visual Analytics

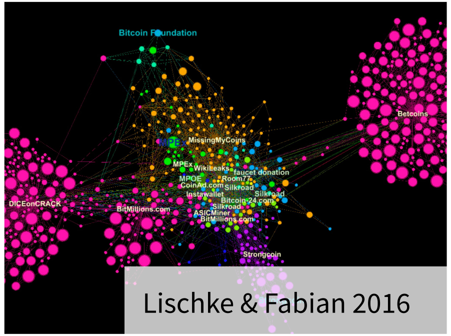
supervised by Natkamon Tovanich
natkamon.tovanich@irt-systemx.fr
and Petra Isenberg

Goals: Modeling + Visualization

1. Build a **machine learning model** to extract patterns from Bitcoin transaction/entity graph.
2. Develop a **visual analytics tool** to help explore Bitcoin data based on the model.

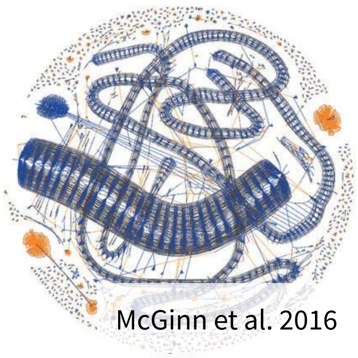


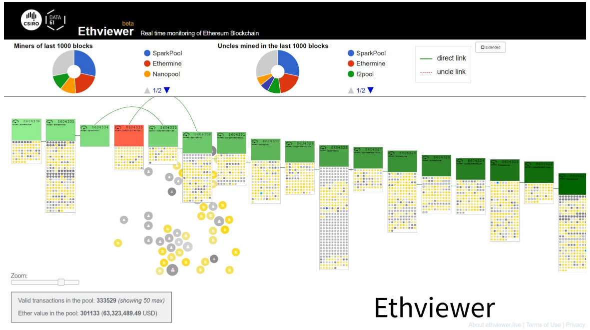
Weber et al. 2019



Lischke & Fabian 2016

We are open for any research domain that you are interested in, e.g. *characterization of user types*, *transaction network analysis*, *cybercrime detection*, etc.


McGinn et al. 2016

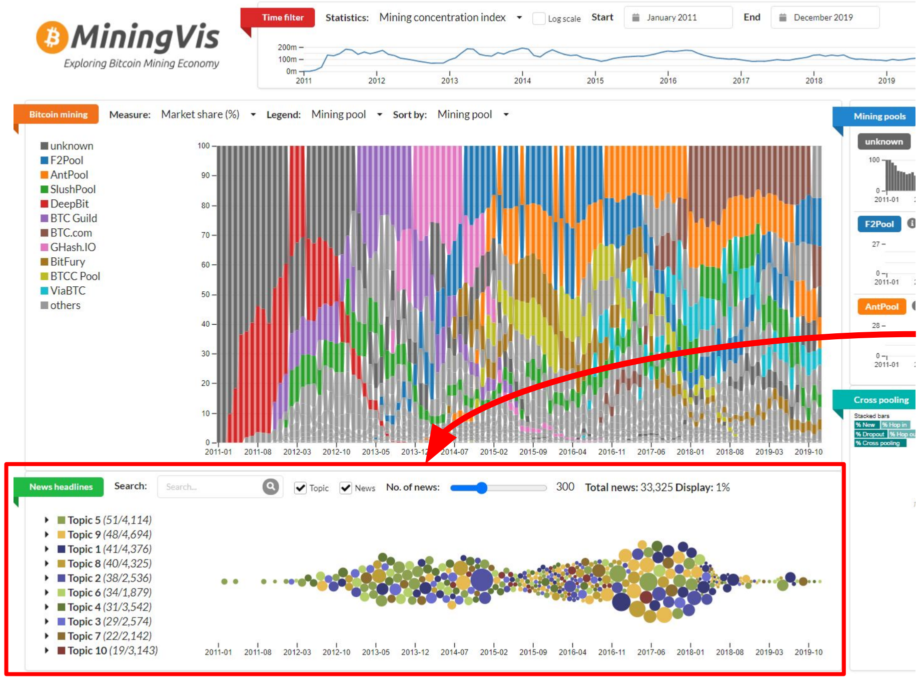


Ethviewer

bitcoin Visual Analytics

supervised by Natkamon Tovanich
natkamon.tovanich@irt-systemx.fr
and Petra Isenberg

Internship Topic: Analysis and Visualization of the Bitcoin Forum

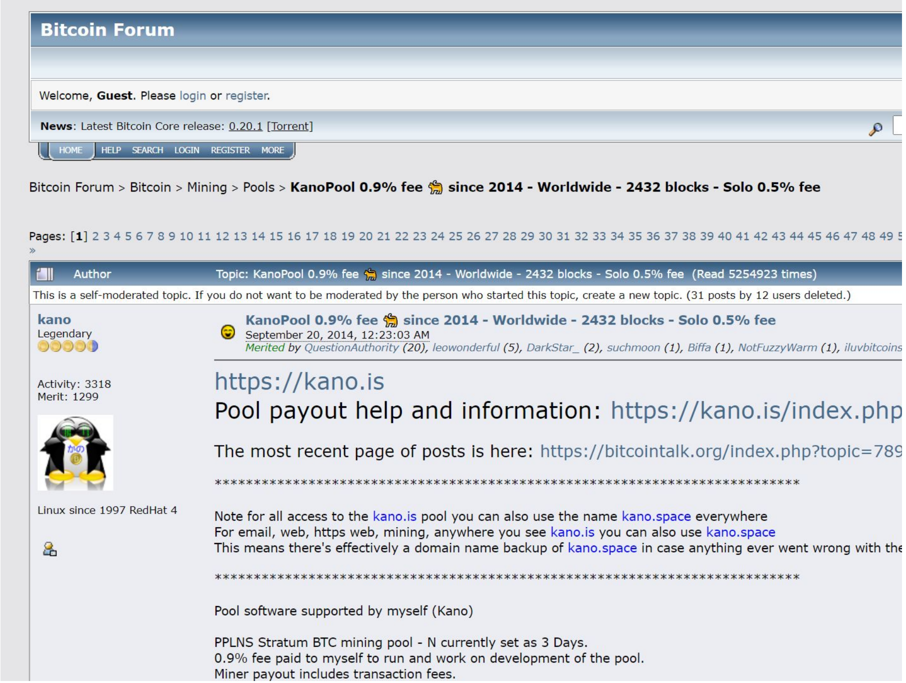


- MiningVis tool help the economist exploring and analyzing the evolution of mining pool in Bitcoin network
- We want to understand <u>the external factors</u> *(e.g. market price, news, and <u>social media</u>)* that affect the evolution of mining pools.

bitcoin Visual Analytics

supervised by Natkamon Tovanich
natkamon.tovanich@irt-systemx.fr
and Petra Isenberg

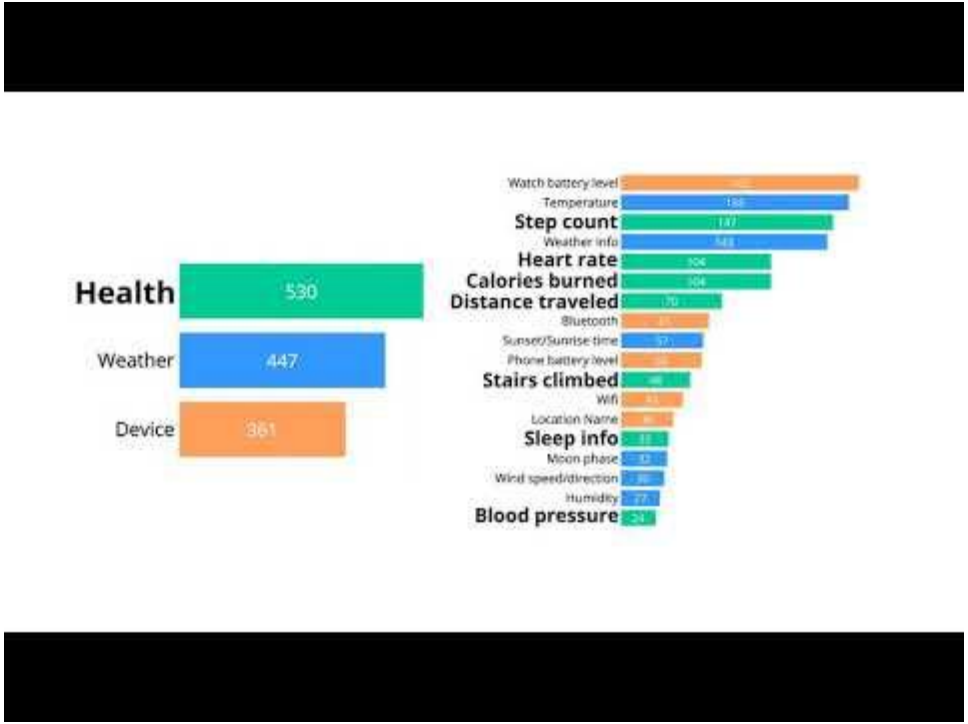
Internship Topic: Analysis and Visualization of the Bitcoin Forum



- Applying text mining techniques to analyze discussions in the forum.
  - Topic modeling, information extraction, text classification, opinion mining, etc.
- Devise a visual analytics tool to help the economist to relate the evolution of mining pool with activities in the online forum.
  - Text + time-oriented visual analytics

Data Representations for Smartwatch

supervised by Alaul Islam
mohammad-alaul.islam@inria.fr
& Petra Isenberg
petra.isenberg@inria.fr



Alaul Islam, Anastasia Bezerianos, Bongshin Lee, Tanja Blascheck, Petra Isenberg. Visualizing Information on Watch Faces: A Survey with Smartwatch Users. Short Papers of the IEEE Conference on Visualization (VIS), Oct 2020, Los Alamitos, CA, United States. https://hal.inria.fr/hal-03005319v2

Data Representations for Smartwatch

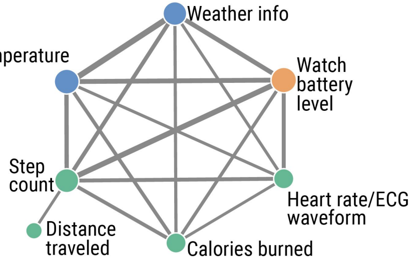
supervised by Alaul Islam
mohammad-alaul.islam@inria.fr
& Petra Isenberg
petra.isenberg@inria.fr

Smartwatches can track a wide variety of data. On average people see **FIVE** types of data.

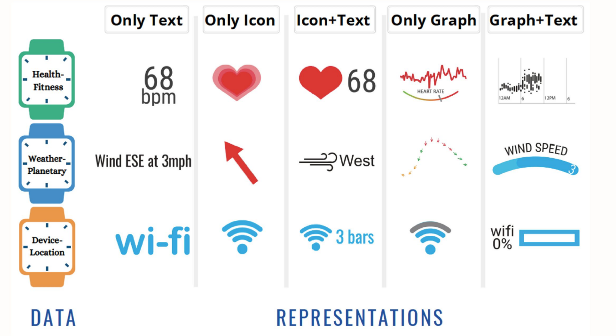
- **internship goal:**
  - Design smartwatch faces with micro-visualizations based on our past research
  - Investigate face preferences for smartwatch wearers
  - If time permits: evaluate
- **requirements:** interest & some design skills, knowledge of Design tools(e.g,Adobe)
- **a plus:** knowledge of Wear OS/Android, interest in future publication

adjusted topics in this context are possible

1st



2nd



3rd