LNUC Project 3

Visual Analytics for Engineering Smarter Systems

Prof. Andreas Kerren & Dr. Rafael Martins

CS Department, Linnaeus University

December 8, 2016

Motivation

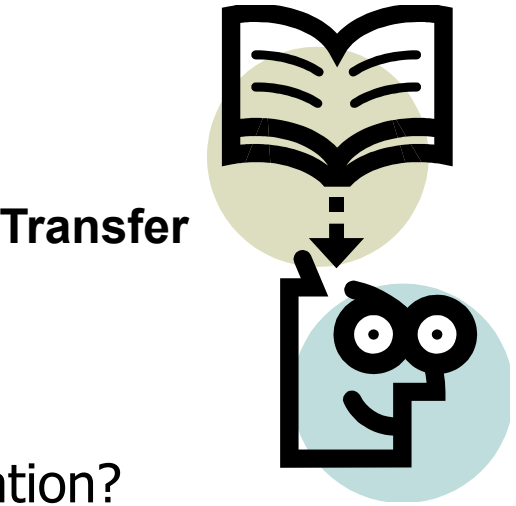
- Information Overload Problem (IOP) or *Big Data*
  - Data is stored without filtering or refinements for future use
  - Often, raw data has no value in itself
  - Time and money are wasted

Motivation

- Information Overload Problem (IOP) or *Big Data*
  - Data is stored without filtering or refinements for future use
  - Often, raw data has no value in itself
  - Time and money are wasted
- Important Questions
  - How to make use of the data?
  - Who/what defines the relevance of an information?

Motivation

- Information Overload Problem (IOP) or *Big Data*
  - Data is stored without filtering or refinements for future use
  - Often, raw data has no value in itself
  - Time and money are wasted



- Important Questions
  - How to make use of the data?
  - Who/what defines the relevance of an information?
  - **What kind of visual representation and interaction technique can facilitate problem solving and decision making?**

Motivation

- In many cases, an automatic analysis is *not* or *only partly* possible!

Anscombe's Quartet: Raw Data

	I		II		III		IV	
	x	y	x	y	x	y	x	y
	10.0	8.04	10.0	9.14	10.0	7.46	8.0	6.58
	8.0	6.95	8.0	8.14	8.0	6.77	8.0	5.76
	13.0	7.58	13.0	8.74	13.0	12.74	8.0	7.71
	9.0	8.81	9.0	8.77	9.0	7.11	8.0	8.84
	11.0	8.33	11.0	9.26	11.0	7.81	8.0	8.47
	14.0	9.96	14.0	8.10	14.0	8.84	8.0	7.04
	6.0	7.24	6.0	6.13	6.0	6.08	8.0	5.25
	4.0	4.26	4.0	3.10	4.0	5.39	19.0	12.50
	12.0	10.84	12.0	9.13	12.0	8.15	8.0	5.56
	7.0	4.82	7.0	7.26	7.0	6.42	8.0	7.91
	5.0	5.68	5.0	4.74	5.0	5.73	8.0	6.89
mean	9.0	7.5	9.0	7.5	9.0	7.5	9.0	7.5
var.	10.0	3.75	10.0	3.75	10.0	3.75	10.0	3.75
corr.	0.816		0.816		0.816		0.816	

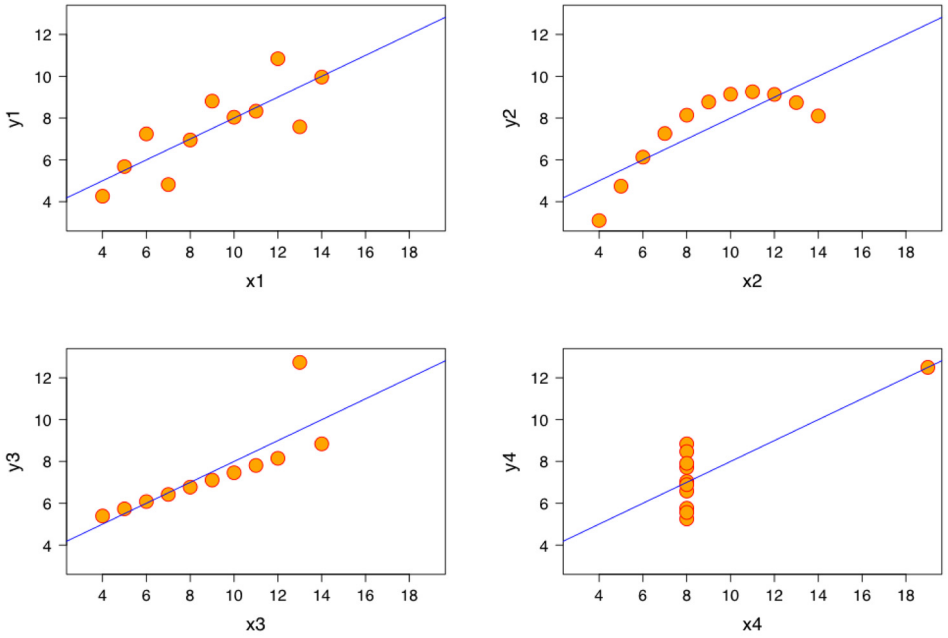
Motivation

- In many cases, an automatic analysis is *not* or *only partly* possible!

Anscombe's Quartet: Raw Data

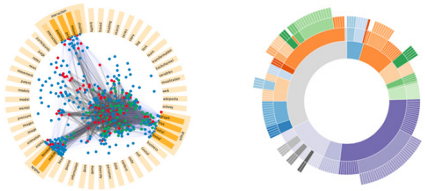
	I		II		III		IV	
	x	y	x	y	x	y	x	y
	10.0	8.04	10.0	9.14	10.0	7.46	8.0	6.58
	8.0	6.95	8.0	8.14	8.0	6.77	8.0	5.76
	13.0	7.58	13.0	8.74	13.0	12.74	8.0	7.71
	9.0	8.81	9.0	8.77	9.0	7.11	8.0	8.84
	11.0	8.33	11.0	9.26	11.0	7.81	8.0	8.47
	14.0	9.96	14.0	8.10	14.0	8.84	8.0	7.04
	6.0	7.24	6.0	6.13	6.0	6.08	8.0	5.25
	4.0	4.26	4.0	3.10	4.0	5.39	19.0	12.50
	12.0	10.84	12.0	9.13	12.0	8.15	8.0	5.56
	7.0	4.82	7.0	7.26	7.0	6.42	8.0	7.91
	5.0	5.68	5.0	4.74	5.0	5.73	8.0	6.89
mean	9.0	7.5	9.0	7.5	9.0	7.5	9.0	7.5
var.	10.0	3.75	10.0	3.75	10.0	3.75	10.0	3.75
corr.	0.816		0.816		0.816		0.816	

+ identical linear regression

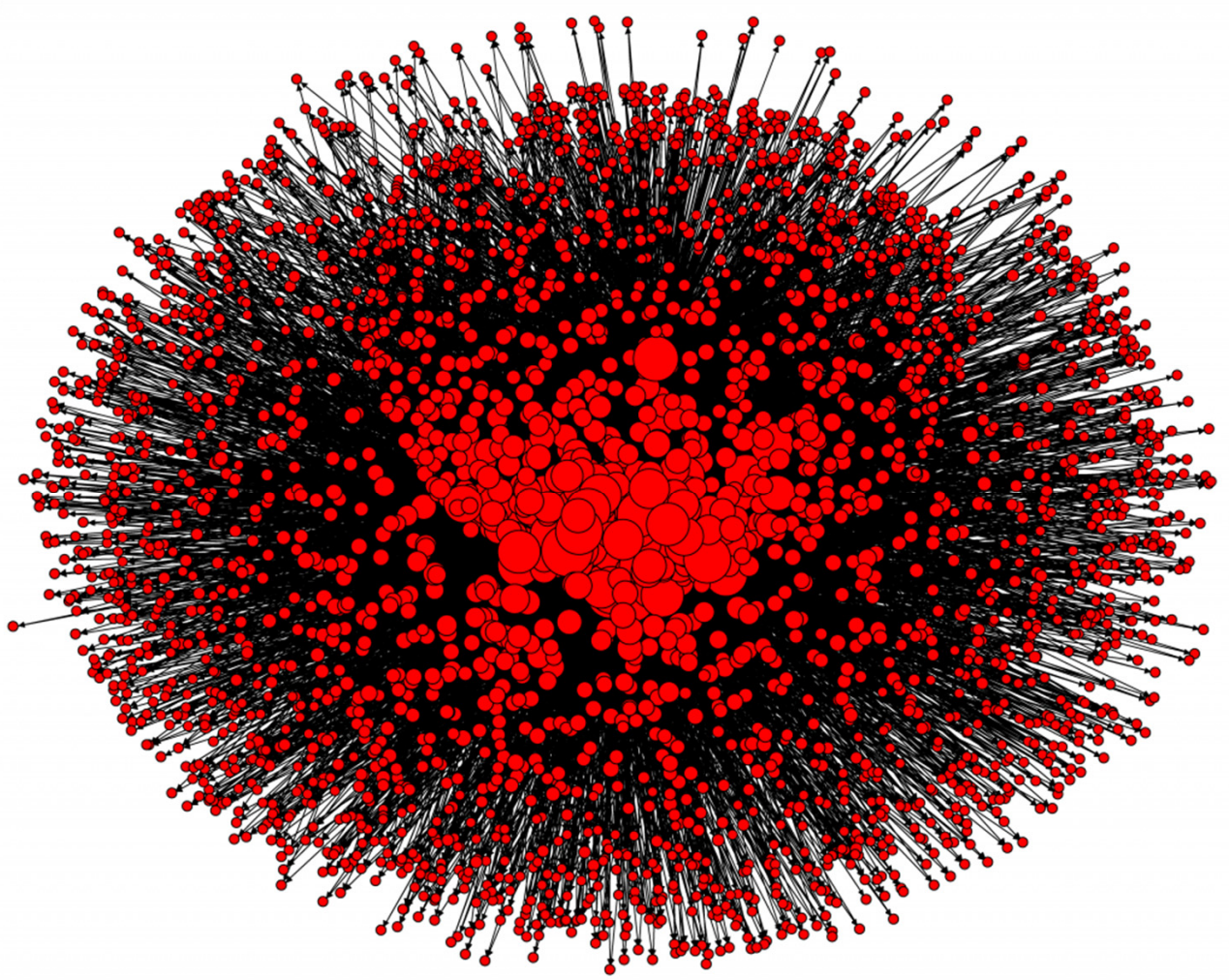


InfoVis

- Information Visualization (InfoVis)
  - Pure interactive visualization methods do not work for billions of abstract data records
  - Not enough pixels, too much clutter, …

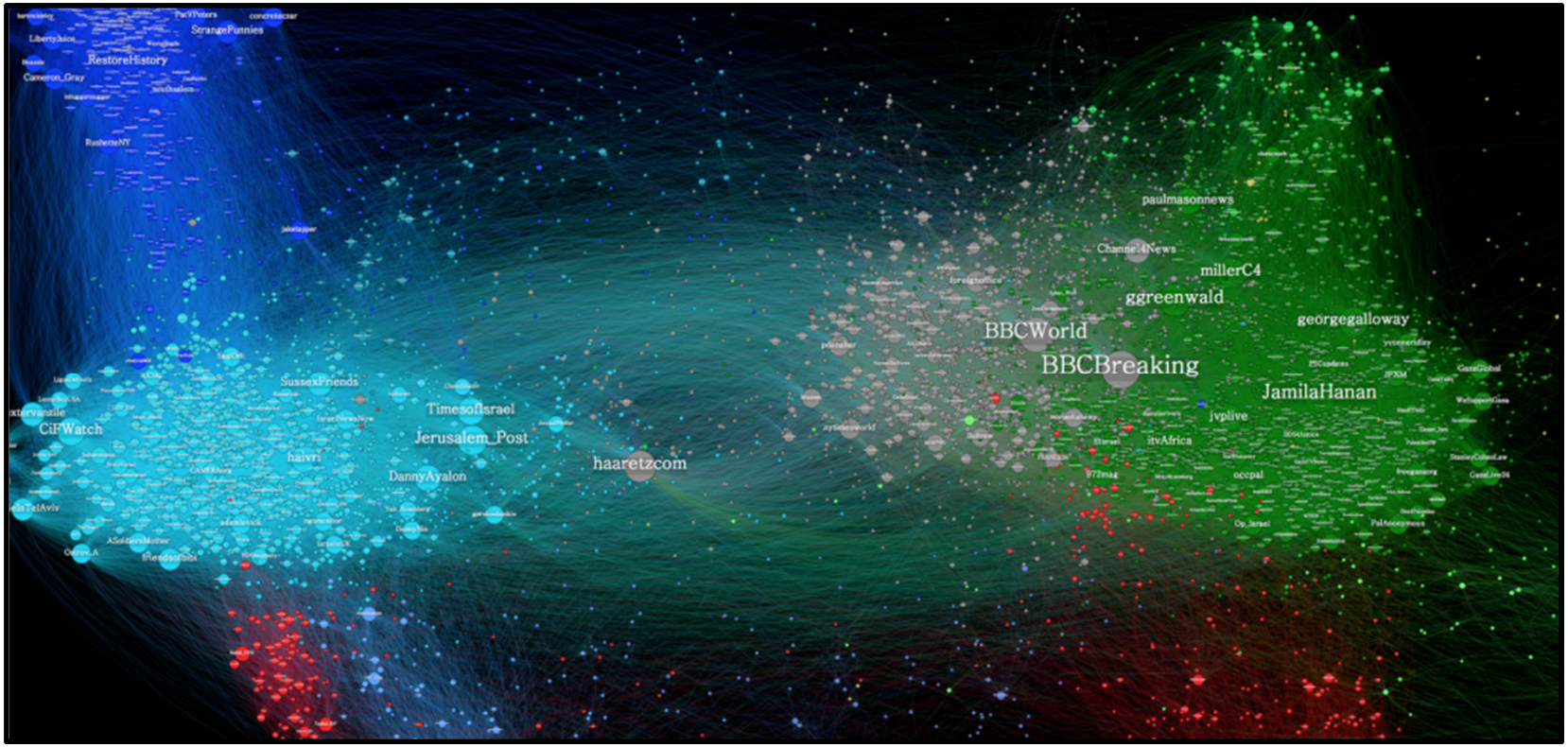


InfoVis



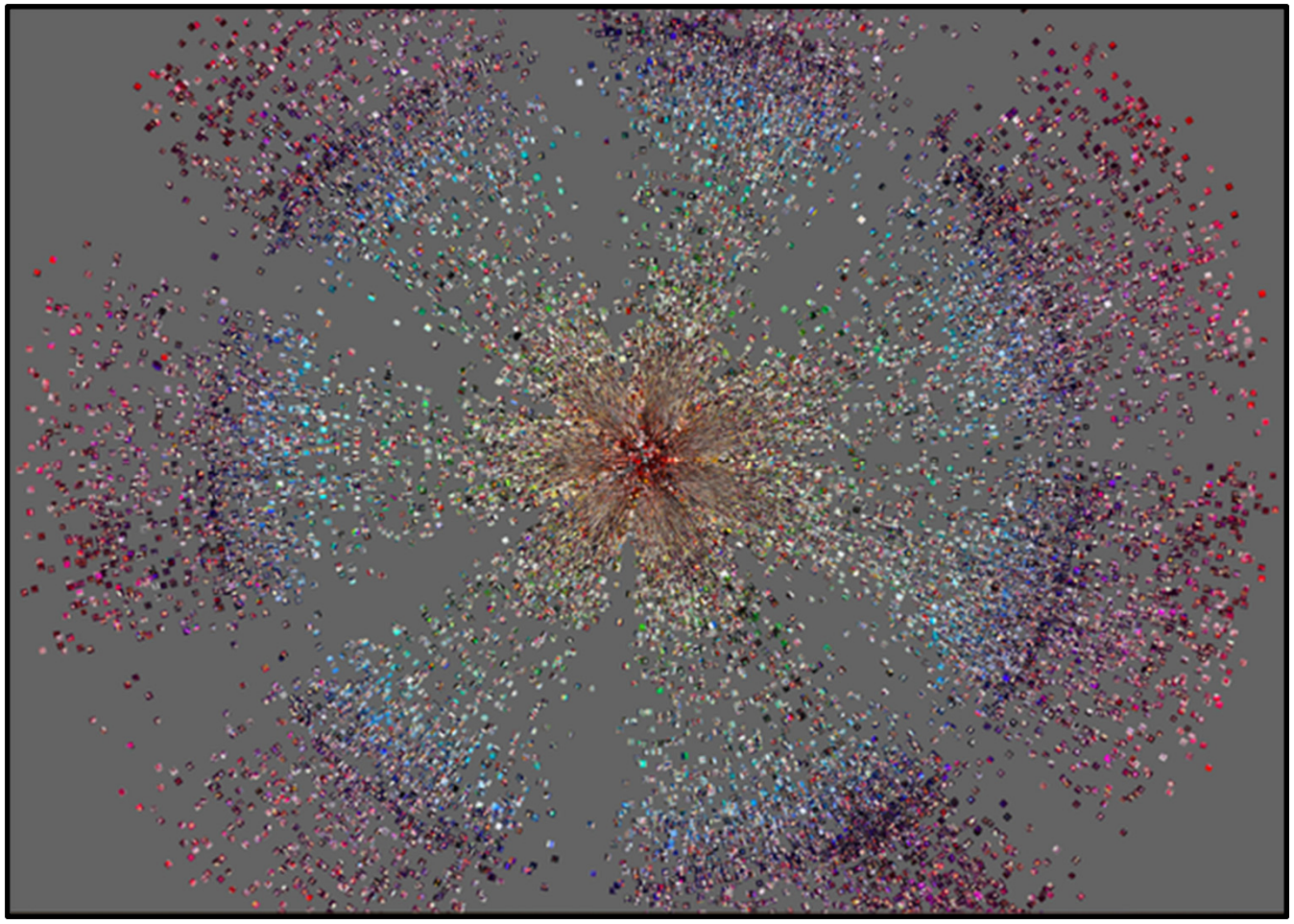
https://www.flickr.com/photos/porternovelli/3102296497

InfoVis



http://www.timesofisrael.com/gorgeous-graphics-tell-big-data-stories-on-israeli-site/

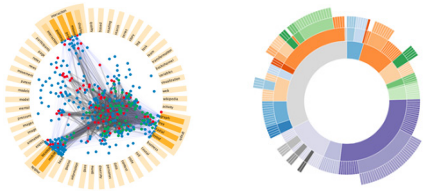
InfoVis



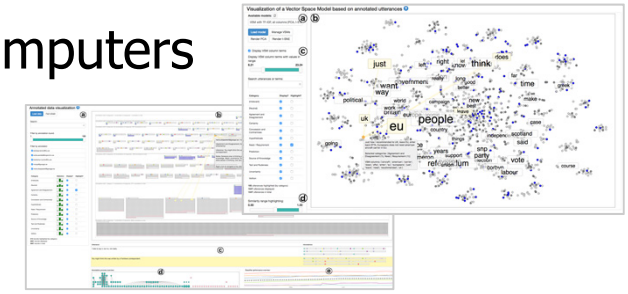
http://firstmonday.org/ojs/index.php/fm/article/view/4711/3698

Visual Analytics & InfoVis

- Information Visualization (InfoVis)
  - Pure interactive visualization methods do not work for billions of abstract data records
  - Not enough pixels, too much clutter, …



- Visual Analytics (VA)
  - Science of analytical reasoning facilitated by interactive visual interfaces
  - Combines the strengths of humans and computers
    - Information and Scientific Visualization, HCI
    - Data Mining and Data Management (ML, SP, …)



Visual Analytics & InfoVis

- VA is used to
  - synthesize information and derive insight from massive, dynamic, ambiguous, and often conflicting data
  - "detect the expected and discover the unexpected" TM
  - provide timely, defensible, and understandable assessments
  - communicate assessments effectively for action, i.e., *decision making*

Visual Analytics & InfoVis

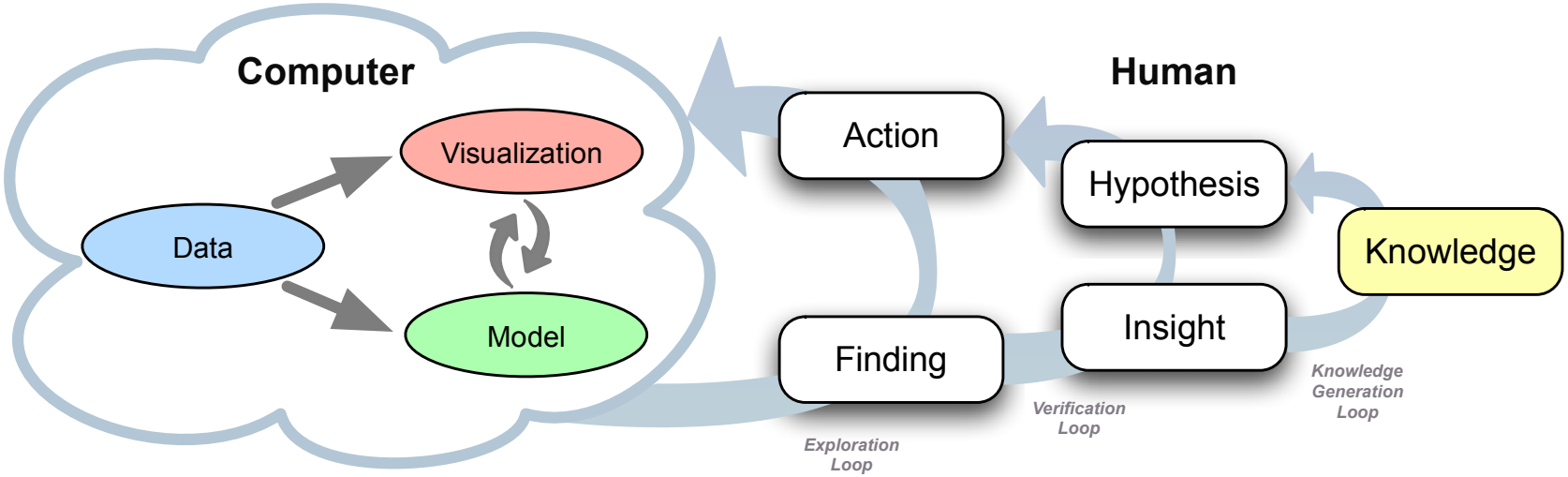
- VA is used to
  - synthesize information and derive insight from massive, dynamic, ambiguous, and often conflicting data
  - "detect the expected and discover the unexpected" TM
  - provide timely, defensible, and understandable assessments
  - communicate assessments effectively for action, i.e., *decision making*
- VA mantra
  - "Analyze first, show the important, zoom, filter and analyze further, details on demand"

Project 3 - VAESS

- Overall aims of the project
  - Generate new knowledge to better understand and engineer complex cyber-physical systems
  - Increase performance, guarantee reliability, better prediction, control/monitoring of their behavior, …

Project 3 - VAESS

- Knowledge generation cycle is based on data collection and **visual analytics** *



* https://dx.doi.org/10.1109/TVCG.2014.2346481

Project 3 - VAESS

- Multidimensional and streaming data comes
  - from the systems themselves and their environments,
  - from the computational models which are used for analysis,
  - and from resulting models that describe the systems and their behavior
- Research goal
  - We will develop foundational visual analytics principles, techniques and tools to reach the aim of better understanding and engineering cyber-physical systems

Project 3 - VAESS

- System domains and data providers
  - still open (e.g., power grids, vehicle fleets)
  - depends on scientific challenges, tasks to solve, domain experts available, …
- Kickoff meeting in January 2017
  - Identification of project members and collaborators
  - Definition of the concrete challenges on the research side (VA, ML, IoT, …) and on the application side

Project 3 - VAESS

- If your analysis problem fits into this scope and you're interested in to get involved

andreas.kerren@lnu.se

- Contact us even if your data, analysis tasks, etc. don't perfectly fit into this VA project
  - Could be the basis of future common proposals for external funding around InfoVis and VA