



THE HORIZONS OF DATA VISUALISATION IN BUSINESS LANDSCAPE

TEXT, EYE TRACKING, COLLABORATIVE ANALYTICS AND DEEP LEARNING

Jacob L. Cybulski

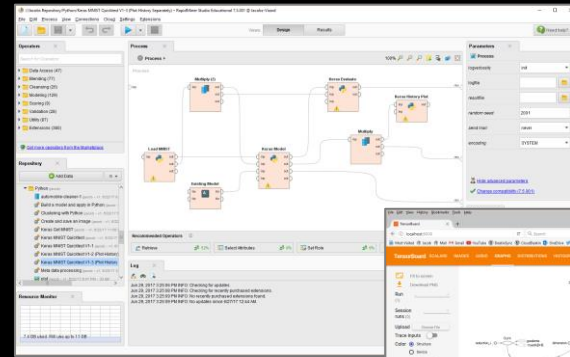
*Interactive Data Analytics Lab
Dept of Info Sys and Bus Analytics*

*Deakin Business School
Faculty of Business and Law
Deakin University*

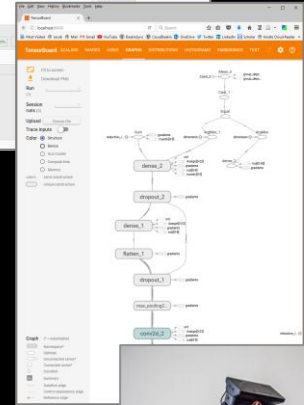
To capture the essence of information
in the moment of time

INTERACTIVE DATA ANALYTICS / JACOB'S PROJECTS

Interactivity
Collaboration &
Affordances



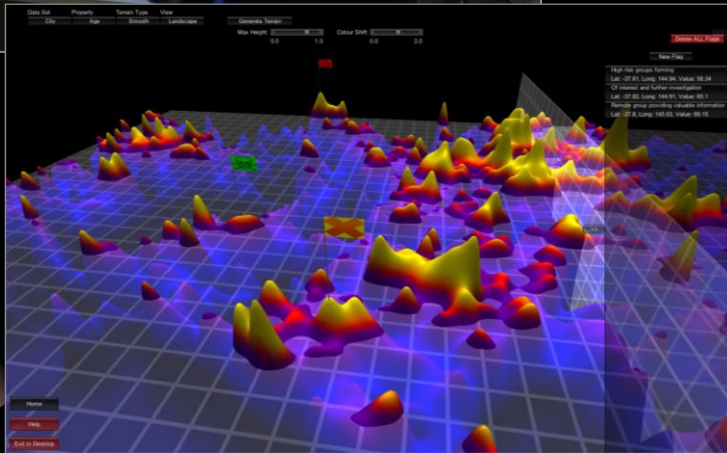
Statistics
Data Mining
Deep Learning



Business
Applications



Data
Analytics &
Sensemaking



Visualisation, Immersion & Metaphors

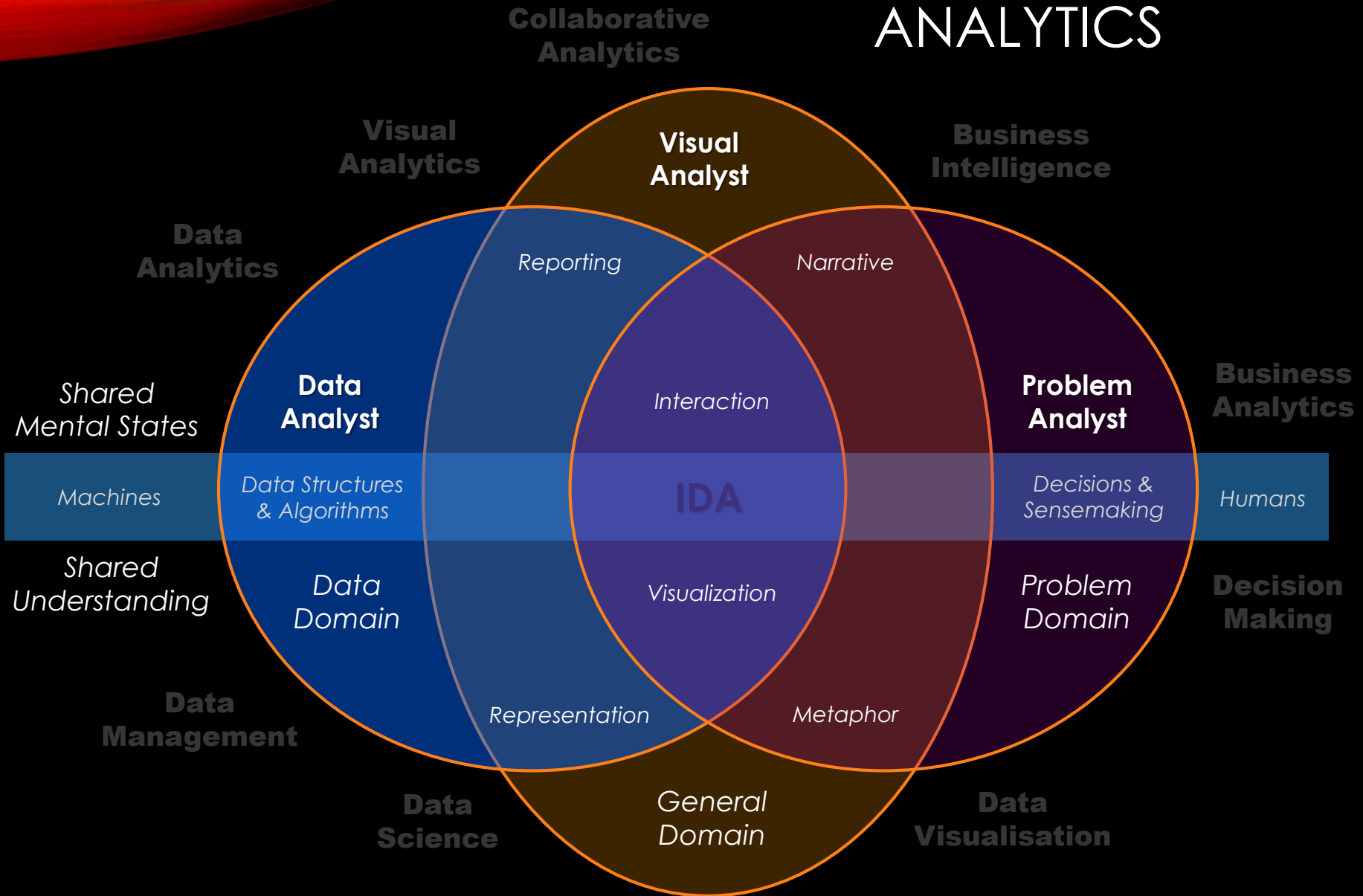


Education



Devices

INTERACTIVE DATA ANALYTICS



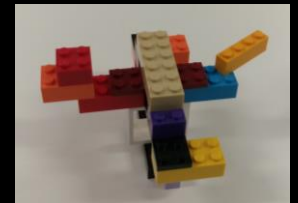
IDA serves different audiences



ANU RESEARCH SCHOOL OF BUSINESS COLLEGE OF BUSINESS & ECONOMICS TEXT ANALYTICS



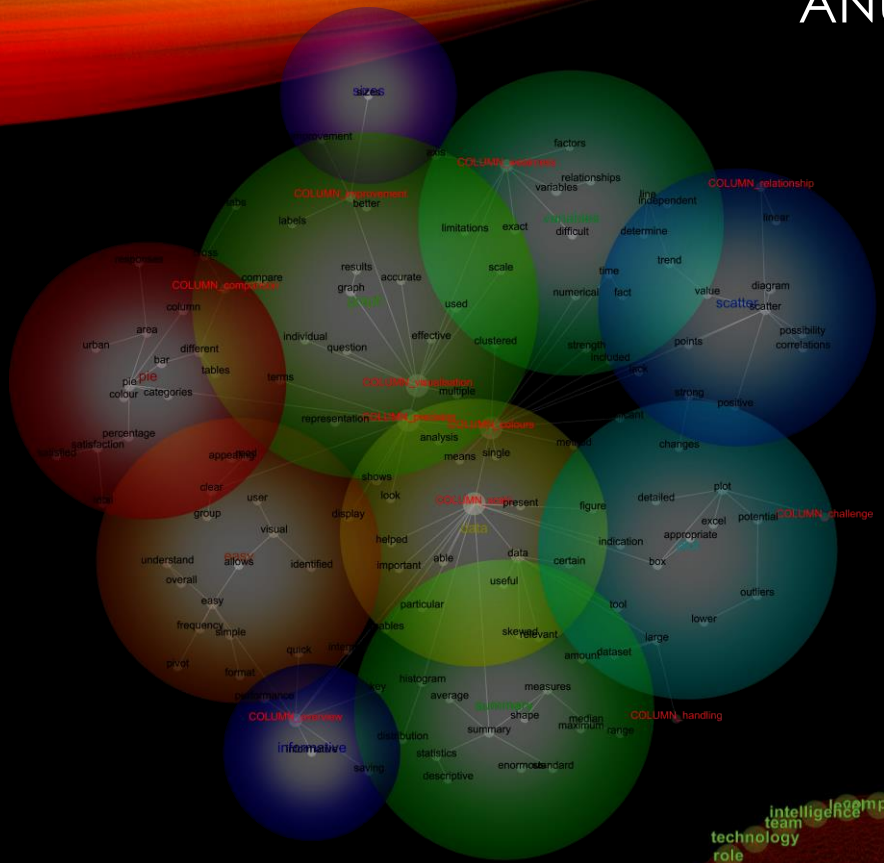
ANU INNOVATION HUB COLLEGE OF BUSINESS & ECONOMICS EYE TRACKING



Collaborators:

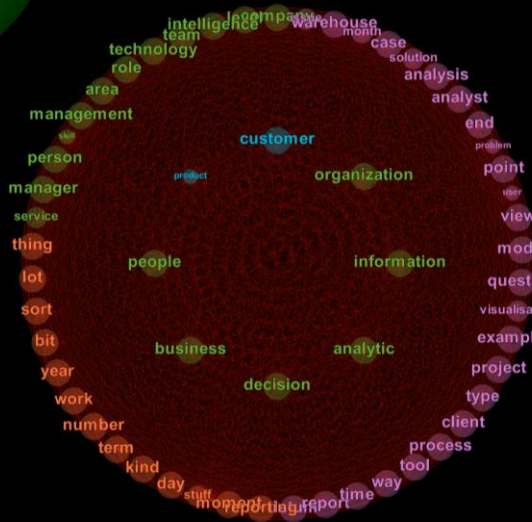
- Shirley Gregor
- Amir Riaz

ANU RESEARCH SCHOOL OF BUSINESS TEXT ANALYTICS

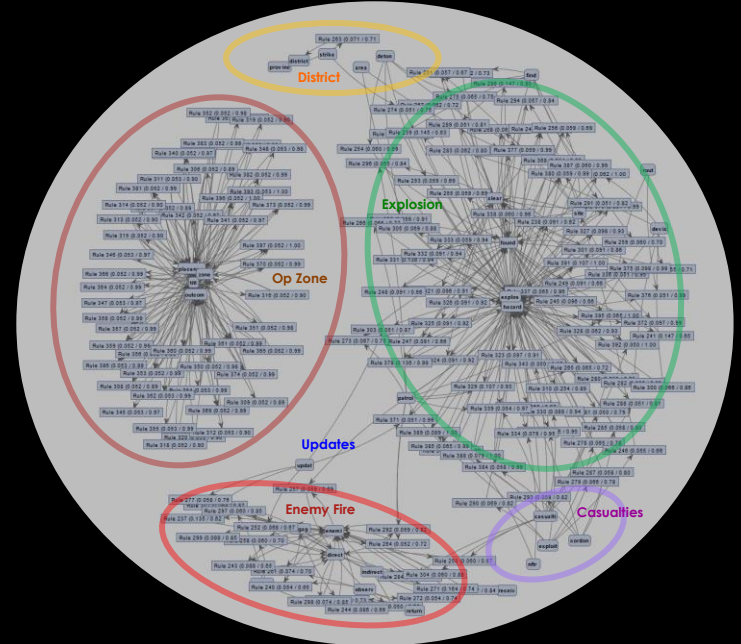


Analysis of student assignments to establish learnt concepts / Leximancer

Past Projects in Text Analytics

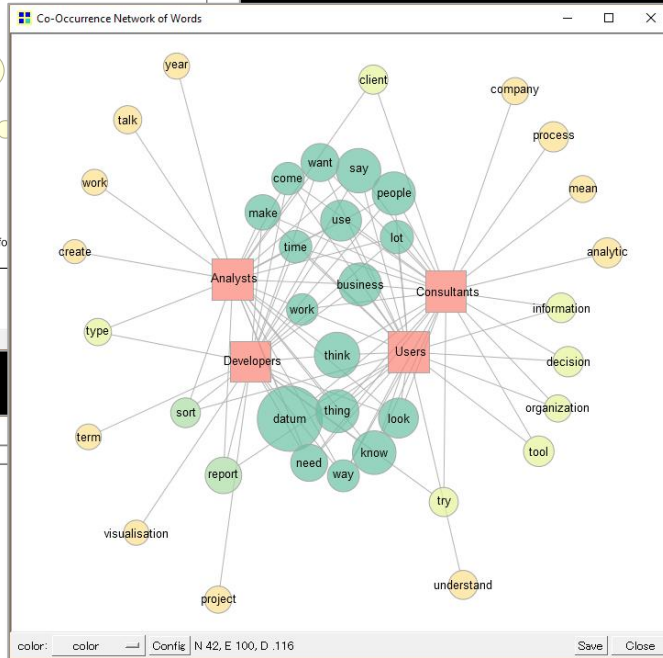
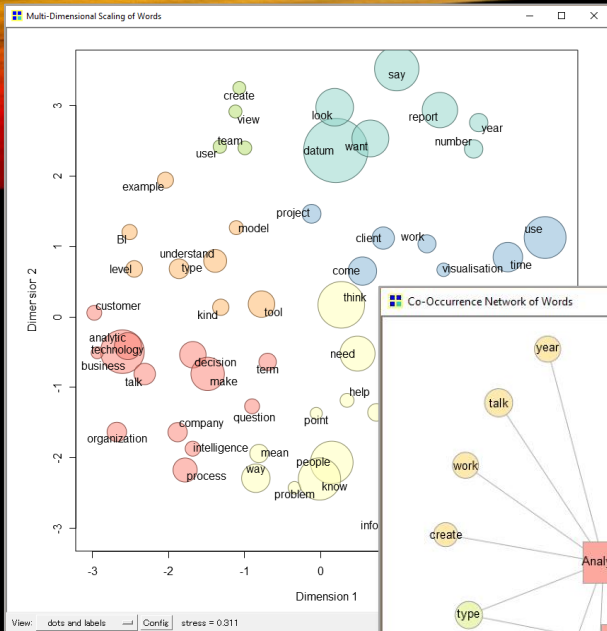


Analysis of interviews with BI analysts / Gephi

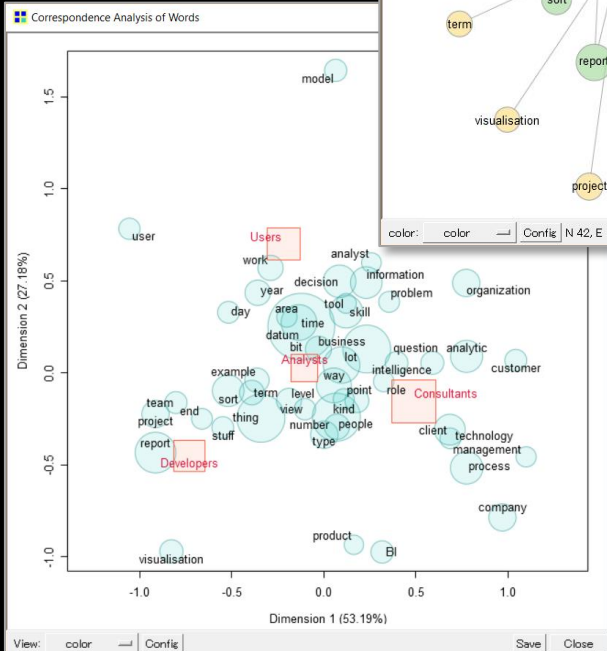


Analysis of Wikileaks Files on Afgan Wars / RapidMiner Studio

ANU RESEARCH SCHOOL OF BUSINESS TEXT VISUALISATION



Term co-occurrence in a sentence or paragraph, provides insights into their relationship. The chart is produced by using a method of force directed graph drawing, which defines attraction and repelling forces between data points, and iterative repositioning the points to minimize the network energy.



KH Coder for quantitative analysis of text

Multidimensional scaling aims to reduce dimensionality of geometric spaces. The method uses an algorithm that tries to re-arrange all points in multi-dimensional space with an aim to fit (squeeze) them into a required number of dimensions.

Correspondence analysis relies on cross-tabulation of documents vs. term frequency to calculate Chi-Squared distance between terms, which can be used in singular value decomposition to identify planar coordinates of terms and documents. Spatial proximity of terms in 2D provides an intuition of their similarity.

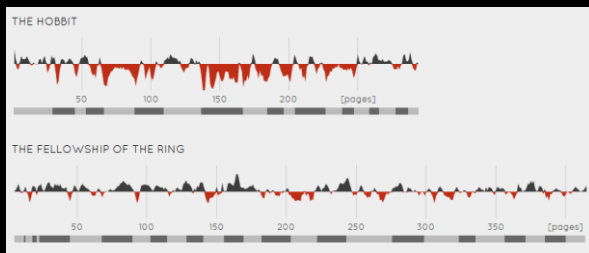
INTERESTING PROJECTS IN TEXT VISUALISATION

J.R.R. Tolkien

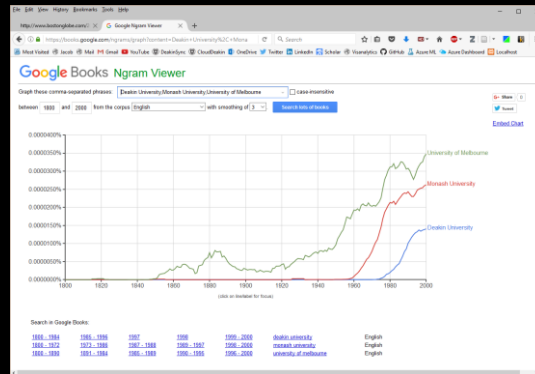
The author of many popular fantasy books, such as *The Hobbit* and *The Lord of the Rings*.

His many volumes of writing can be quickly “grasped” via a web site which allows visualization of various aspects of his work.

Among many insights that can be gained from the web site charts we can see that *The Hobbit* (about Bilbo) has a largely negative mood, whereas *The Fellowship of the Ring* (about Frodo) seems a more cheerful book.



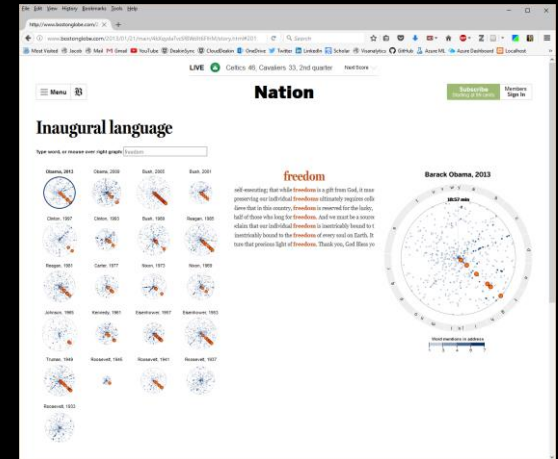
Text Visualisation *Requires* Text Analytics



Google Ngrams

With access to millions of documents Google is able to visually present statistics about any concept mentioned on the web.

Here is the comparison of the number of in-text references to three well-known Australian universities.



BostonGlobe

Analysis of language used by different USA presidents in their inaugural statement.

“Freedom” and “country” seem to be the most important words used in their speeches.

“Peace” used to be important.

“Jobs” and “happiness” were the invention of Obama.



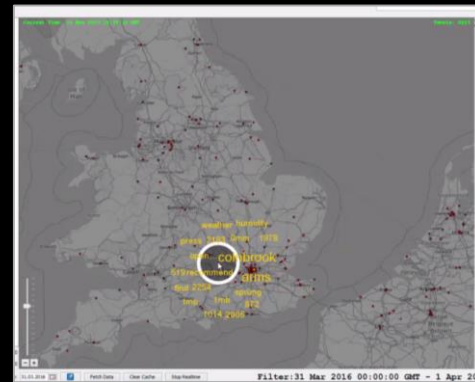
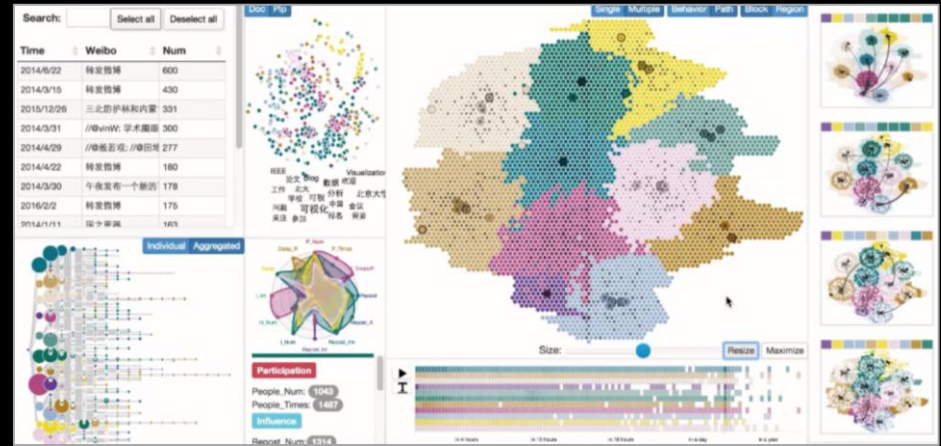
VIS'2016 (BALTIMORE, USA) FUTURE TRENDS

Text streaming and flowing (e.g. Twitter)
by Shixia Liu and students, Tsinghua University



Analysis of dynamic text in
complex social contexts

Information diffusion and propagation
by Siming Chen et al, Peking University



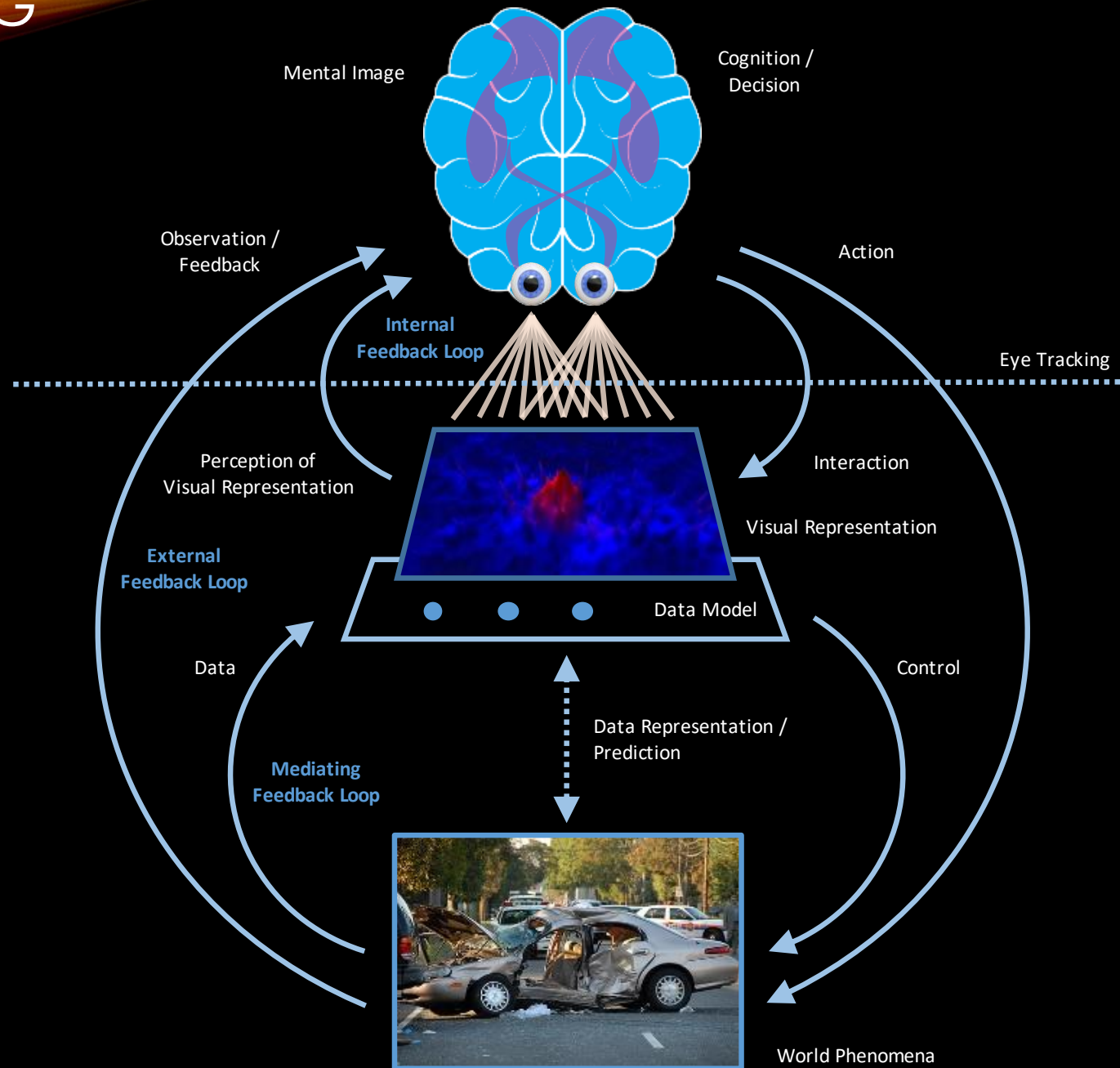
Documents Compass
by Florian Heimerl, Markus John, Qi Han and Steffen Koch
Universität Stuttgart

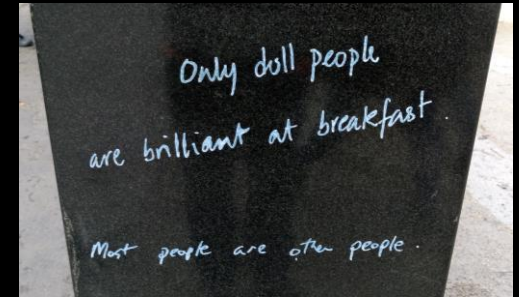
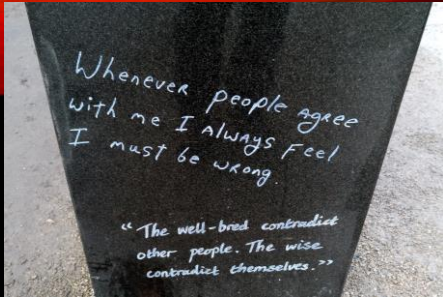
EYE TRACKING

This project explores the use of eye tracking to study the effectiveness of data visualisation in problem-solving and decision-making.

Great add on to any data visualisation research, text or 3D, or VR/AR environments.

A range of inexpensive devices are now available, e.g. Pupil Labs eye tracking binoculars for HTC Hive and MS HoloLens.





MONASH ADAPTIVE VISUALISATION LAB COLLABORATIVE VISUAL ANALYTICS

MONASH IMMERSIVE VISUALISATION PLATFORM IMMERSIVE DATA VISUALISATION



Collaborators:

- Kim Marriott
- Tim Dwyer
- Maxime Cordeil



MONASH VISUALISATION LABS



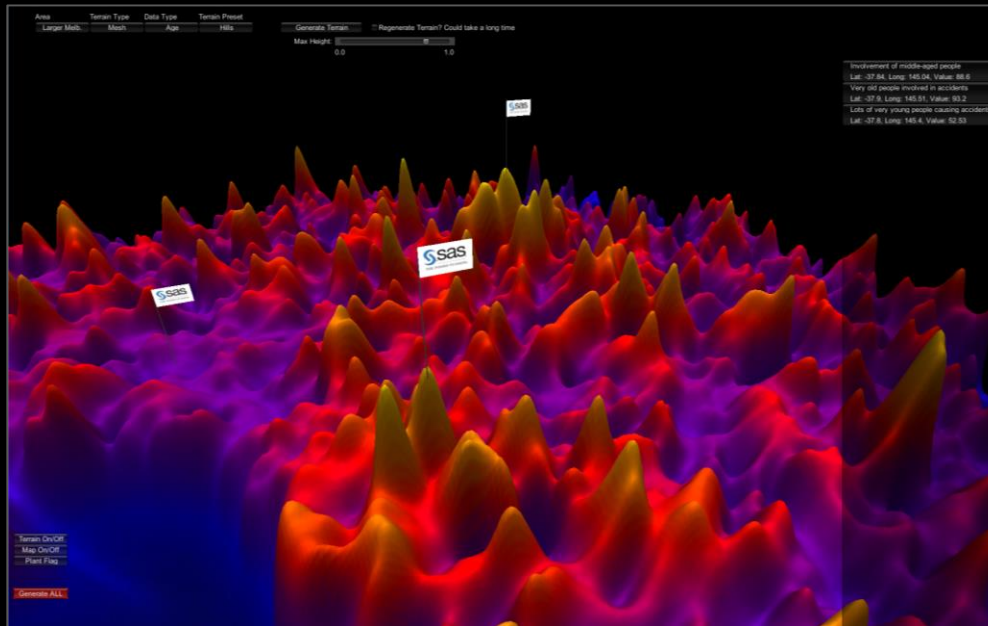
- **CAVE 2:** a world-leading capability for the display and interactive exploration of rich and large scientific and engineering datasets. It is needed by research scientists and engineers to visualise, manipulate and comprehend data such as engineering models.
- **Collaborative Visual Analytics:** research towards collaborative methods and technologies to enable a group of humans to work in a shared virtual or augmented space in order to enable groups of users to explore complex data in a collaborative virtual system.



FUNDAMENTAL PREMISE OF COLLABORATIVE VISUAL ANALYTICS

CVA = data analysis by means of interactive manipulation of visual data representation in teams

Allows engaging instinctively with complex data

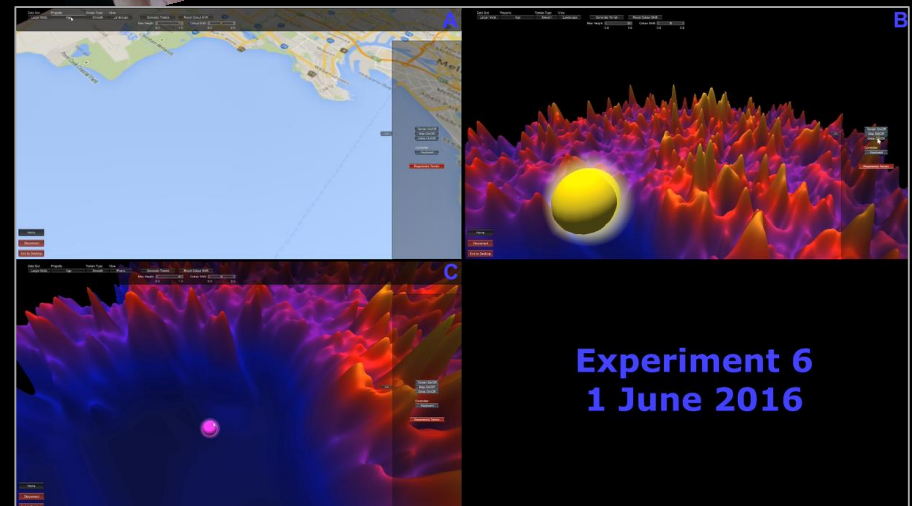
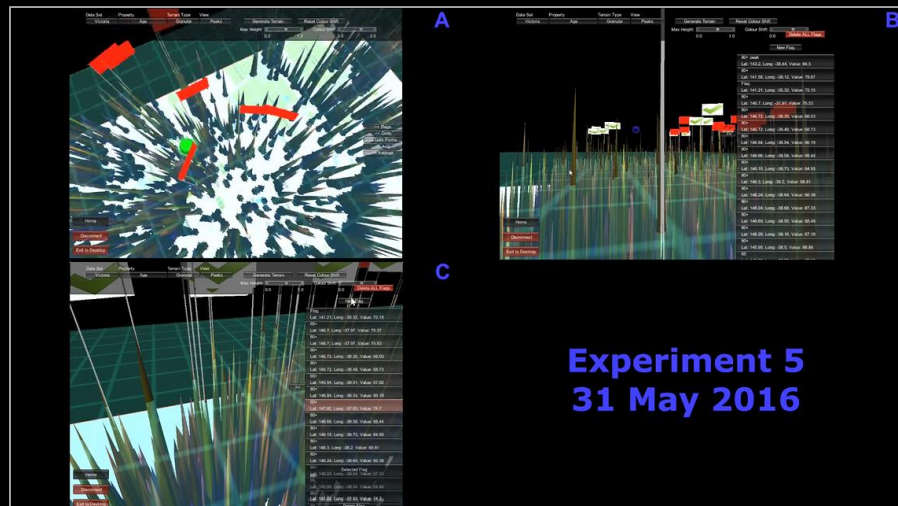
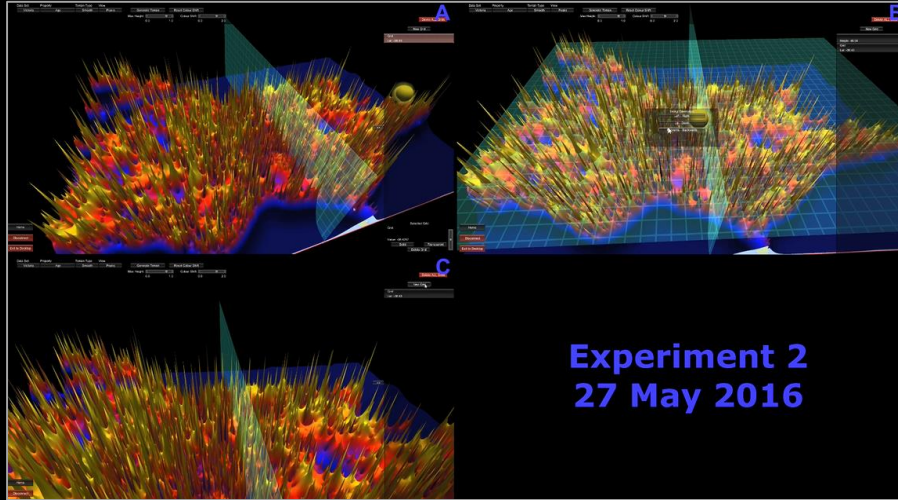


Assist gaining, communicating and sharing of insights into data and phenomena that data represents, then turning them into consensual decisions

Relies on human innate abilities of perception, cognition as well as team dynamics to act in the world

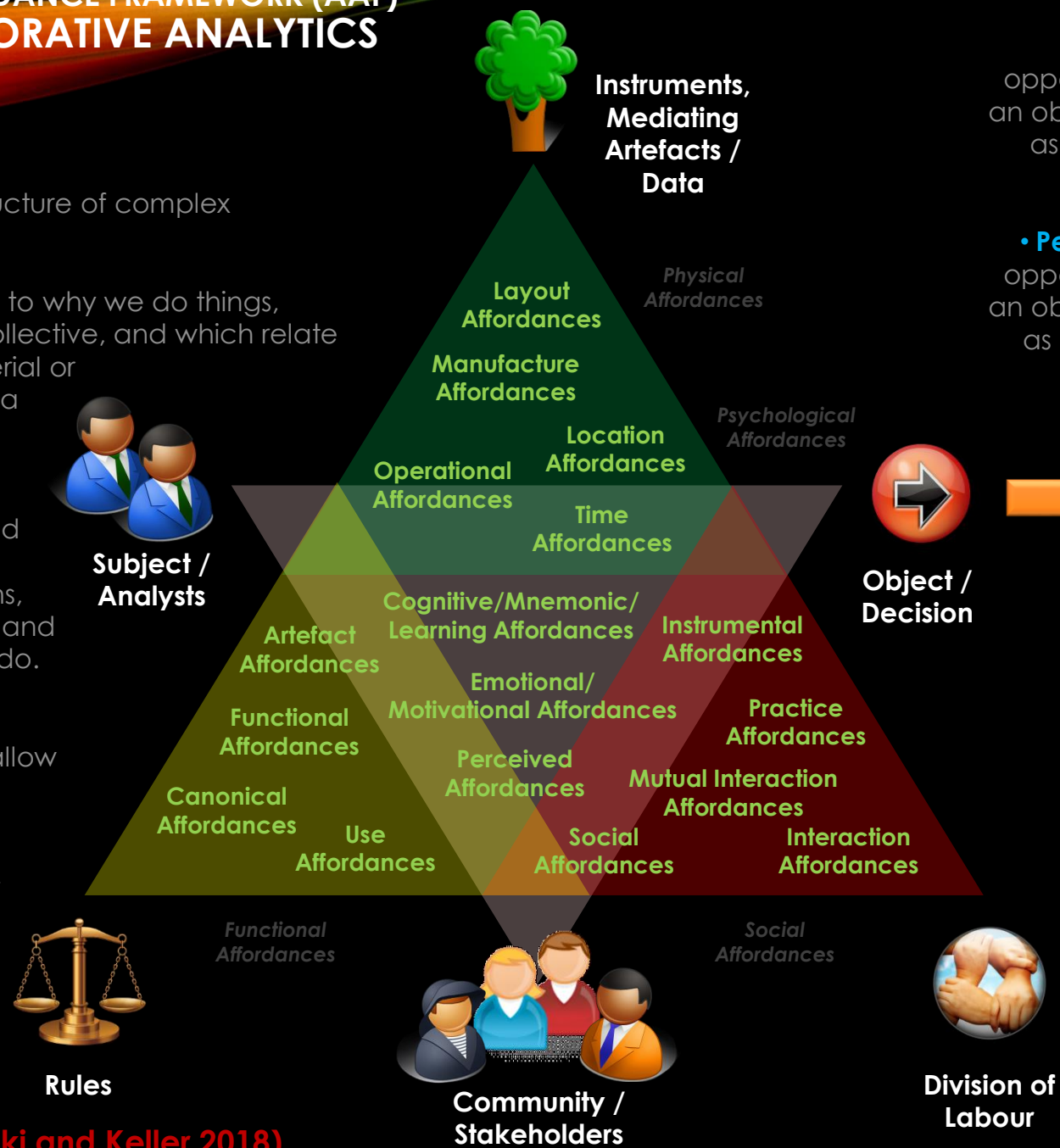


TEAM PROBLEM SOLVING COLLABORATIVE ANALYTICS



ACTIVITY/AFFORDANCE FRAMEWORK (AAF) FOR COLLABORATIVE ANALYTICS

- **Activity System**
is an organised structure of complex human activities.
- **Activities**
provide motives as to why we do things, which are often collective, and which relate to an object, material or ideal, that satisfies a personal or social need.
- **Actions**
realise activities and are defined by our conscious intentions, are goal directed, and relate to what we do.
- **Operations**
are typically unconscious and allow to execute actions according to the specific conditions in the environment and relate to how we do things.



• **Affordance**
opportunity of action on an object or environment as presented to actors (Gibson).

• **Perceived Affordance**
opportunity of action on an object or environment as perceived by actors (Norman).

Business analytics activity system should furnish affordances to support data visualization, interactivity and collaboration.



ICM VISLAB AT WARSAW UNIVERSITY VISUALISATION AND DEEP LEARNING



Collaborators:

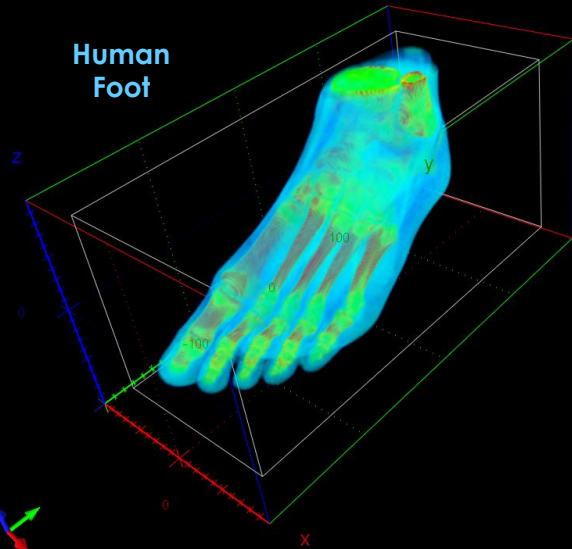
- Bartosz Borucki
- Marek Niezgódka



-22C

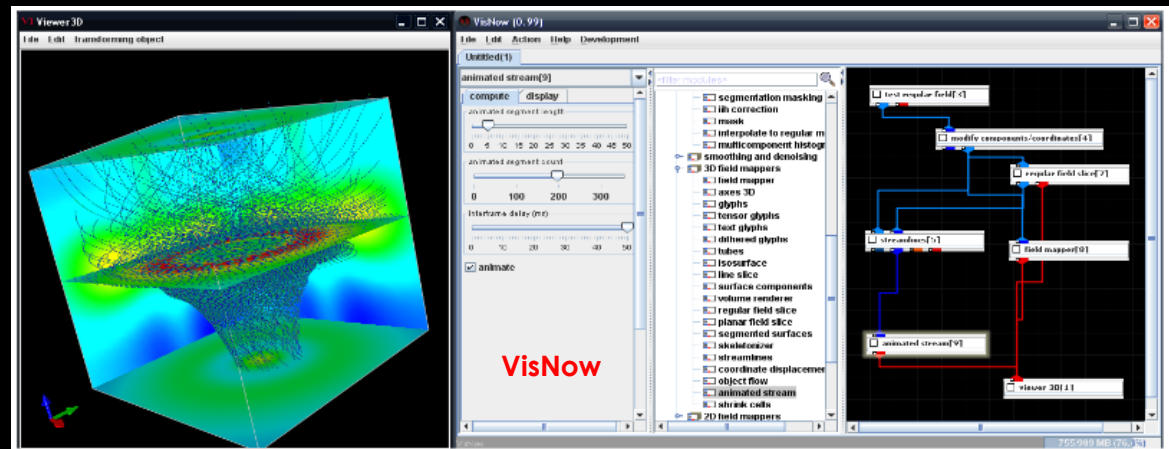
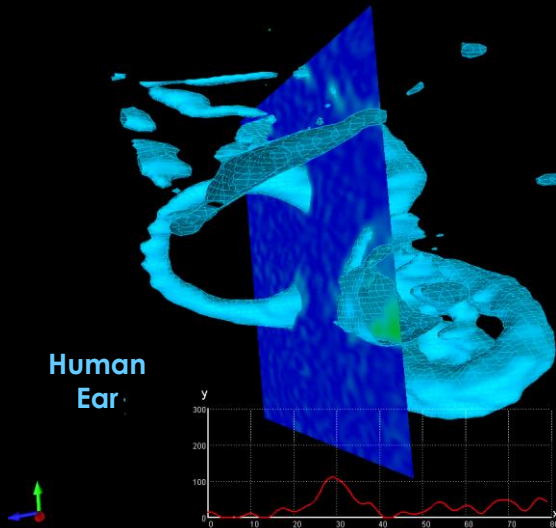
ICM VISLAB AT WARSAW UNIVERSITY

Human
Foot



- **ICM: A well funded research facility, associated with the National Supercomputer Centre, which offers mathematical and machine modelling services to research and commercial organisations. It is well known for providing weather forecasts and visualization to European Union.**
- **Laboratory of Visual Analysis (VisLab): is the research and development group in the area of visualization and image analysis, embracing research, software development, infrastructure and applications. The major fields of interest cover scientific visualization and medical image analysis.**
- **VisNow: a dataflow driven modular visual data analysis platform with extensive data processing and scientific visualization. Supports very large datasets. Developed over the period of 10 years at VisLab, used very extensively.**

Human
Ear



FROM ICM VISLAB TO DEEP LEARNING AT DEAKIN

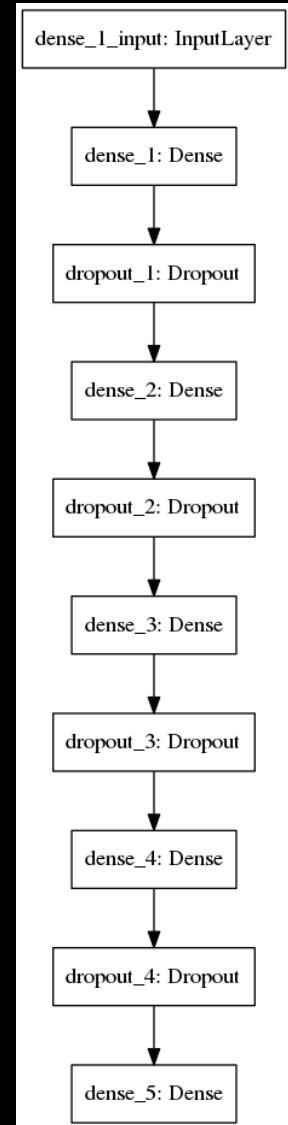
Project at ICM VisLab (2 weeks)

- A German hospital required assistance with postoperative diagnosis of Achilles tendon injuries.
- They provided VisLab staff with 2000 CAT scans (in 7 planes) with additional information of previous diagnoses.
- VisLab used this information to create a *deep learning* classifier of medical images (using UC Berkeley Caffe).
- The performance (98%) was assessed at the level exceeding that of professional diagnosticians.



Projects at Deakin

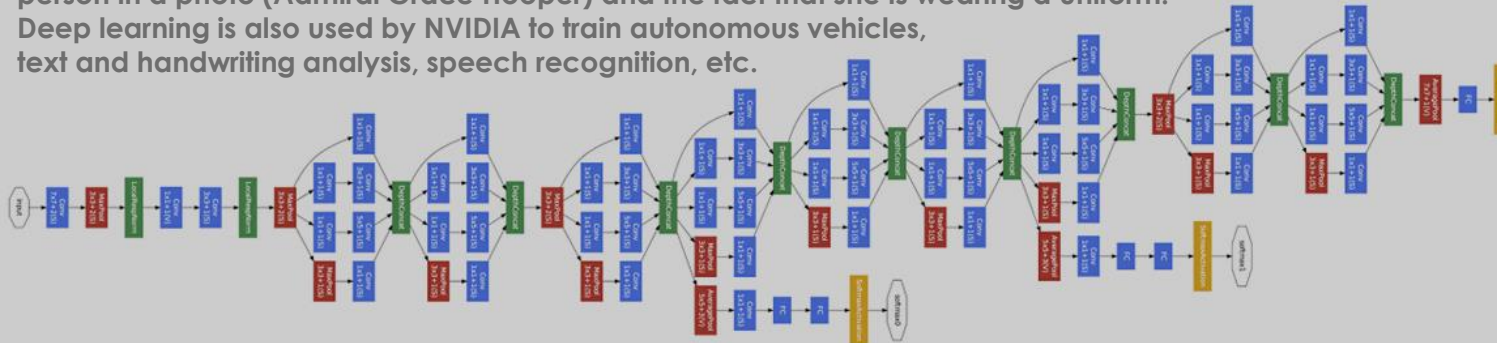
- To develop a new facility at Deakin Business School to support *deep learning* analytics to solve complex business problems.
- The facility will rely on the data analytics and visualization expertise of researchers in the Department of Information Systems and Business Analytics.
- Deep learning analytics allows processing of very large sets of data in numeric, text, image and video format. It often results in highly accurate predictions.
- So far, the majority of applications are in Science and Engineering.
- Business applications include prediction of stock market changes, real-estate prices, analysis of social media content, real-time analysis of policy announcements, fraud detection, anomaly analysis, etc.



WHAT IS DEEP LEARNING AND HOW CAN YOU ACCESS IT

- Deep learning is a class of machine learning techniques which aim at building very large neural networks (a class of predictive) models.
- Such deep neural networks consist of thousands of simpler models, called neurons, which are based on multiple regression and non-linear transformation of data.
- While the concept of neural networks has been known for many years, until recently their large-scale processing was computationally prohibitive.
- However, special techniques have been developed to train such large neural networks efficiently.
- High-performance graphics cards (GPUs), which have 1000s of processing cores, allow the creation and use of deep models.
- Deep learning programming libraries, such as Tensorflow, TFLearn, Keras, MxNet, Caffe, CNTK, H2O, can be used from popular data analytics packages, e.g. Anaconda, R / R Studio, RapidMiner, SAS, SPSS, Azure, etc.
- As their popularity gains the momentum and their power increases, in recent years, deep learning systems have won nearly all data mining competitions (e.g. Kaggle).

Google Inception network is used in image recognition. For example, it is able to identify a person in a photo (Admiral Grace Hooper) and the fact that she is wearing a uniform. Deep learning is also used by NVIDIA to train autonomous vehicles, text and handwriting analysis, speech recognition, etc.





DEAKIN BUSINESS SCHOOL VISUALISATION AND DEEP LEARNING

Start-up List:

- Jacob Cybulski
- Rens Scheepers
- Dilal Saundage
- Ali Tamaddoni
- Mory Namvar
- Bardo Fraunholz
- Lemai Nguyen

We are currently working towards the development of capacity to deploy deep learning solutions to our colleagues and business partners.

Inspired by our Polish colleagues...

The first applications will include commercial image classification, text analysis, social media sentiment analysis, and stock market predictions. Data visualization will also be used to explore the data and to interpret results.



THANK YOU