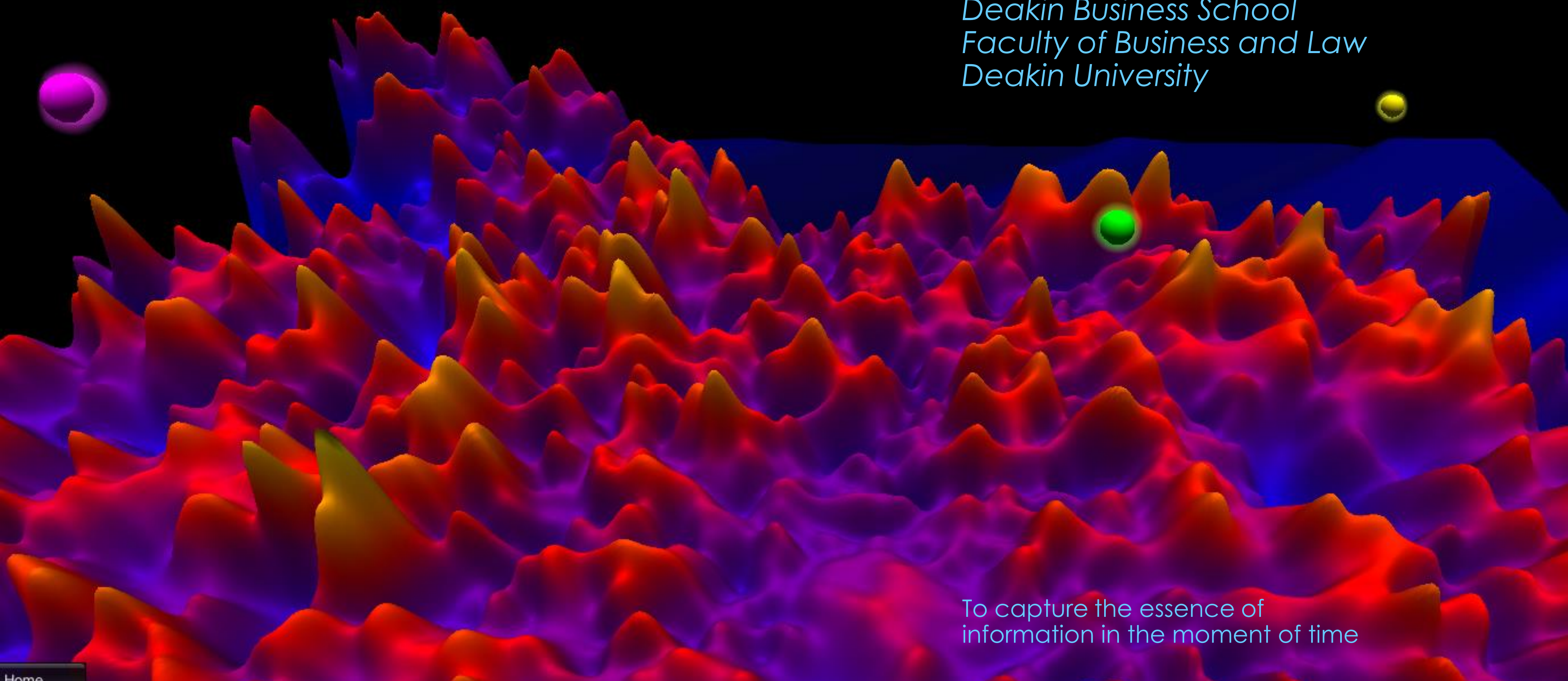
~~INTERACTIVE~~ Collaborative VISUAL ANALYTICS

FOR BUSINESS SENSEMAKING

Jacob L. Cybulski

SAS Visual Analytics Collaboratory
Dept of Info Sys and Bus Analytics

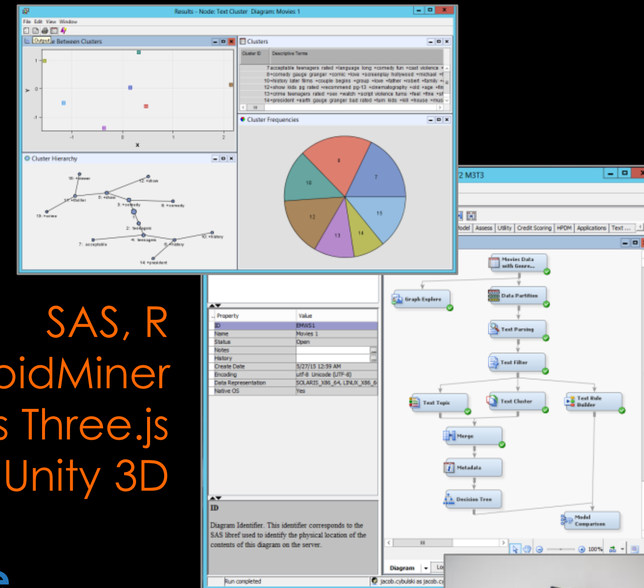
Deakin Business School
Faculty of Business and Law
Deakin University



To capture the essence of
information in the moment of time

SAS VISUAL ANALYTICS COLLABORATORY @ DEAKIN

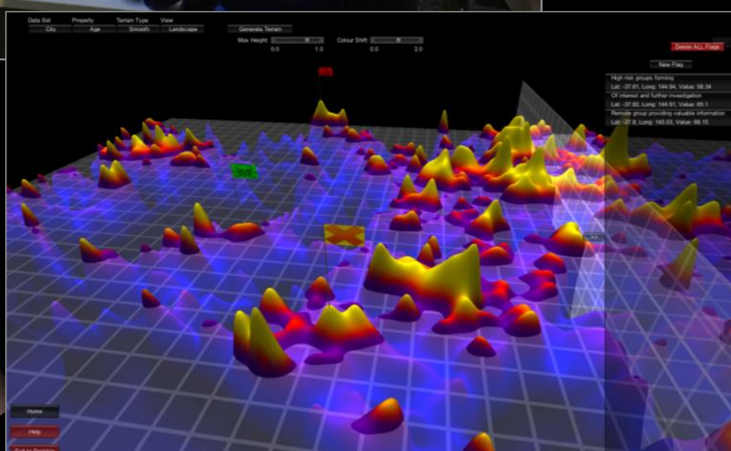
Research

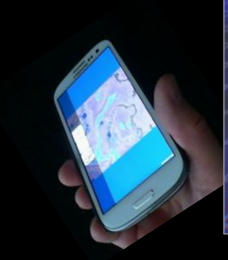


SAS, R
RapidMiner
D3.js Three.js
Unity 3D

Immersive Visual Analytics

Exploration Interactivity & Collaboration





Collaborative & Interactive 3D Visual Analytics

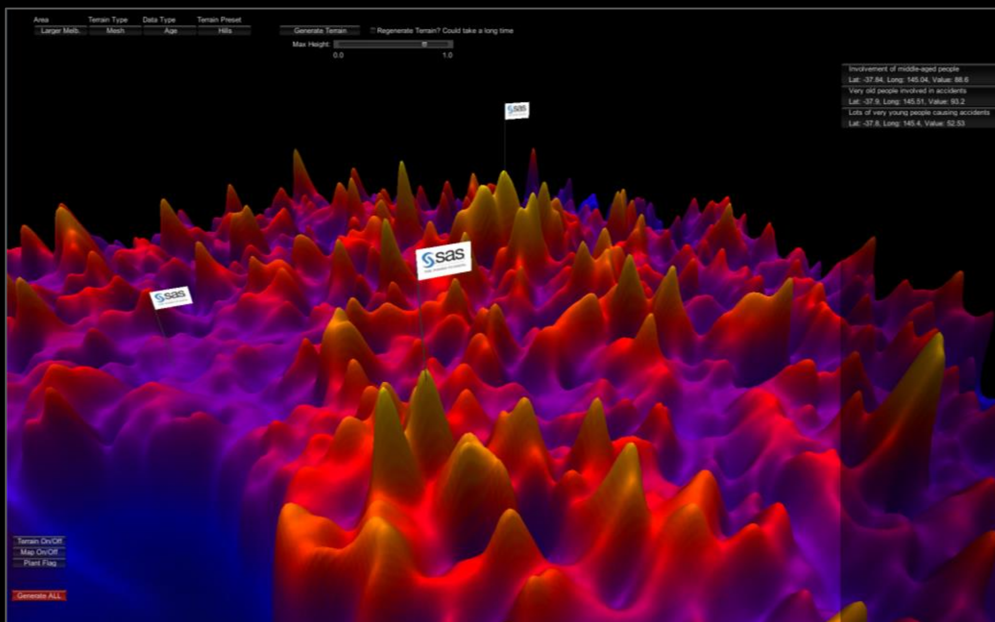
Devices

Education

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 data represents, then turning them into consensual decisions

Relies on human innate abilities of perception, cognition as well as team dynamics

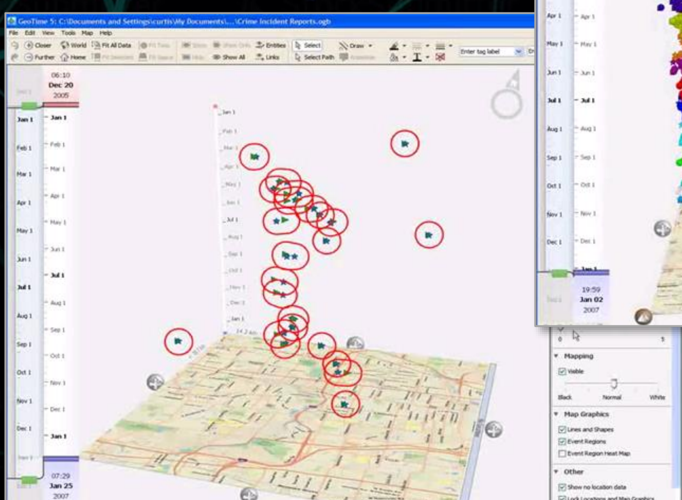
BUSINESS/SOCIAL IVA
REALITY NEEDS TO BE FOUND

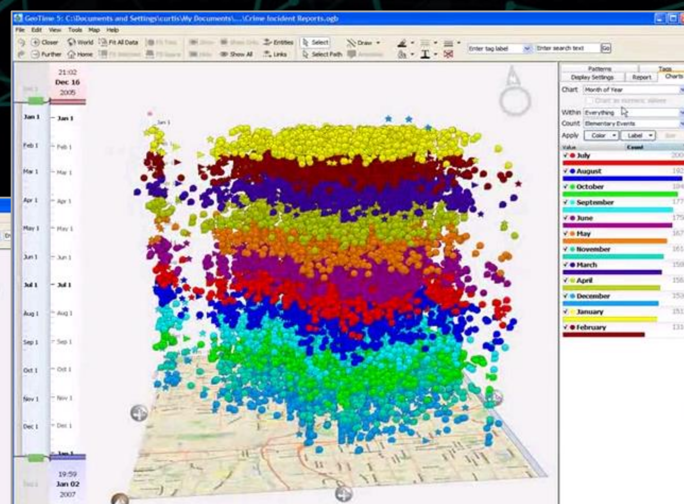
Social and business data is often abstract, e.g.

- *police effectiveness,*
- *spread of infection,*
- *investment risk*

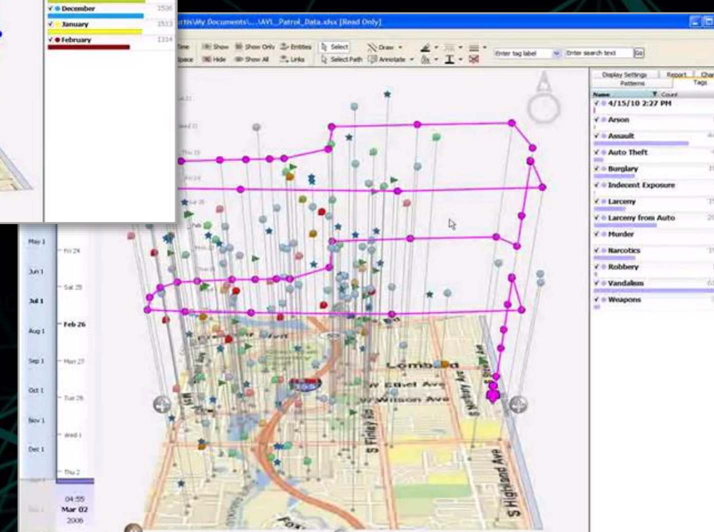
Complex data and analytics requires team effort

Unlike the scientific applications, business applications often do not provide natural real-life data form or representation





Its visual representation needs to be found in the analytic process and metaphor



GeoTime:
Effectiveness of Police Patrols

MAKING SENSE
OF TWO QUESTIONS

How can you gain insights from data and its visualization to assist decision making?

Traditionally this is the aim of Business Intelligence. Also the question posed by data analytics and visualisation researchers.

How can data and its visualization assist analytic teams in making sense of business and as a result gain insights to improve decision making?

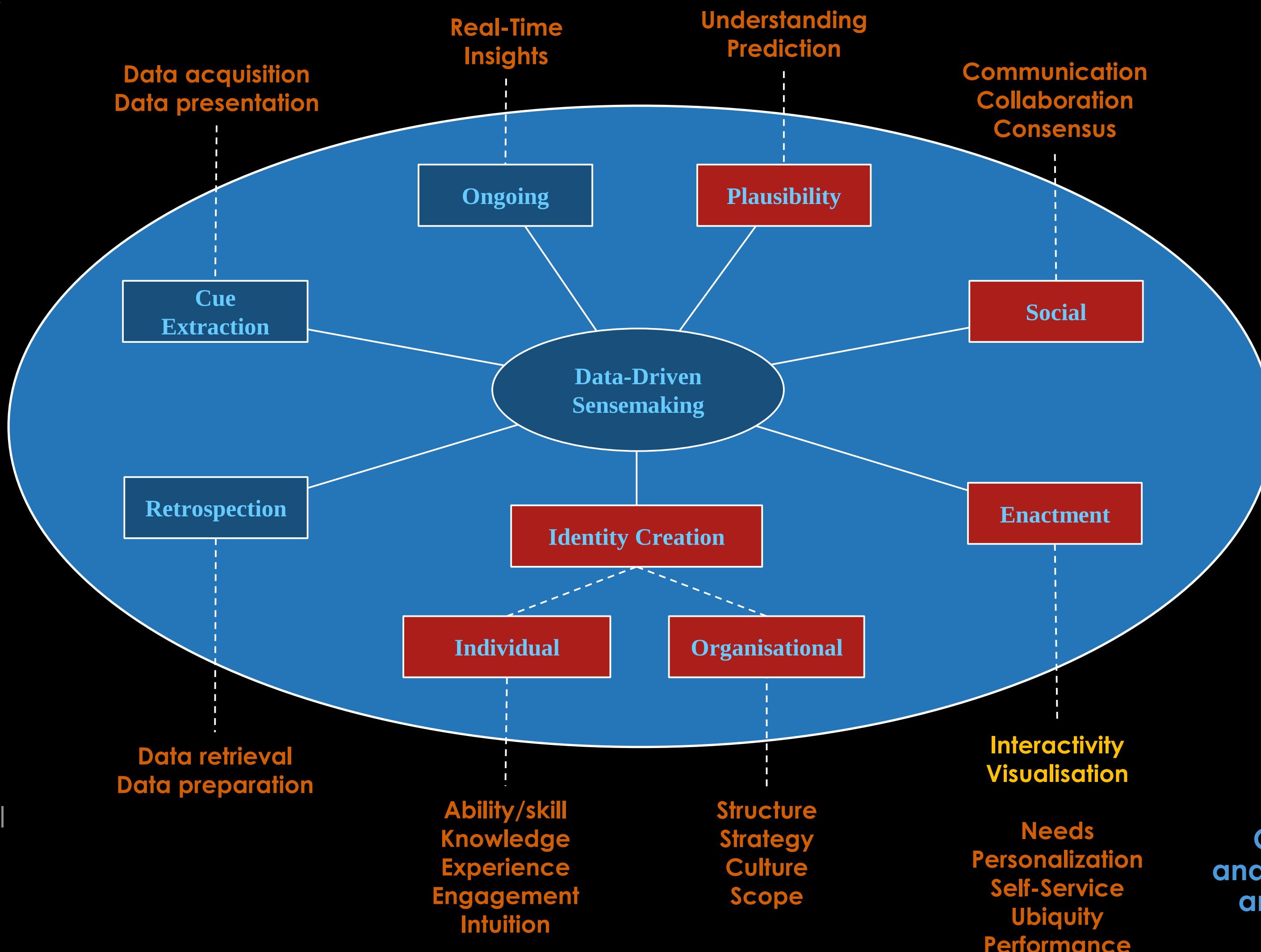
Traditionally this is the aim of business executives. Also the question posed by IS researchers, which should be the aim of BI.

Sensemaking is the prerequisite of informed decision-making (Namvar and Cybulski 2015, 2016)

DATA-DRIVEN SENSEMAKING FRAMEWORK

Seven properties of Weick's sensemaking:

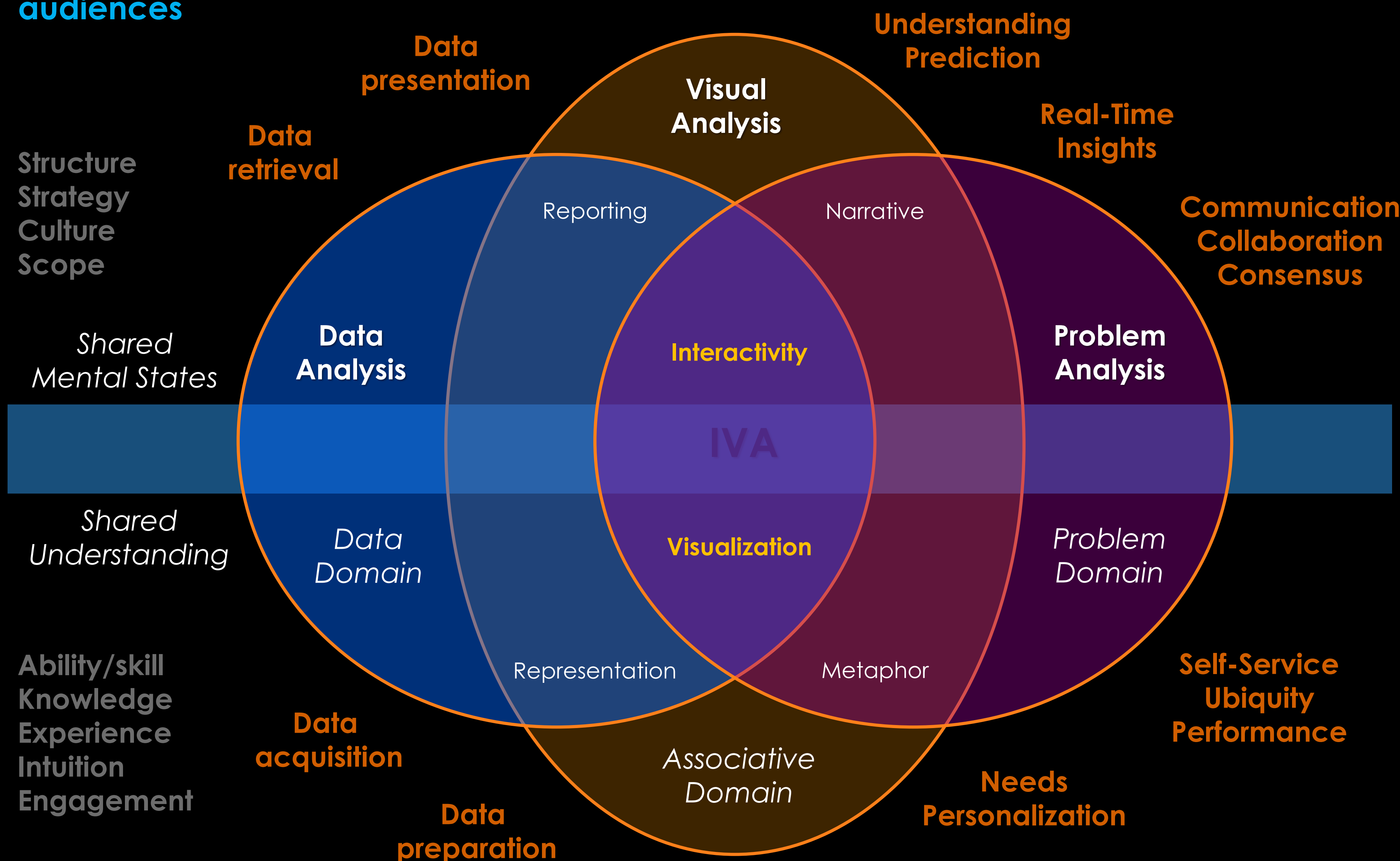
- **Retrospection** to give meaning to past events
- **Cue extraction** to detect important events
- **Ongoing process** allowing to respond to change
- **Plausibility** preferred over accuracy
- **Social nature** of sensemaking in organisational settings
- **Enactment** to actualise own ideas and identity in the environment
- **Identity creation** to develop individual and organisational capacity for sensemaking



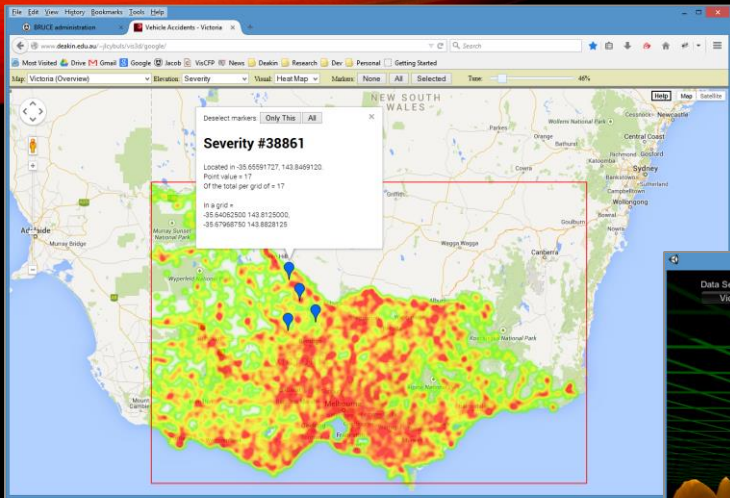
Could analytics and vis help?

COLLABORATIVE-ANALYTIC FRAMEWORK

IVA serves different purposes for different audiences

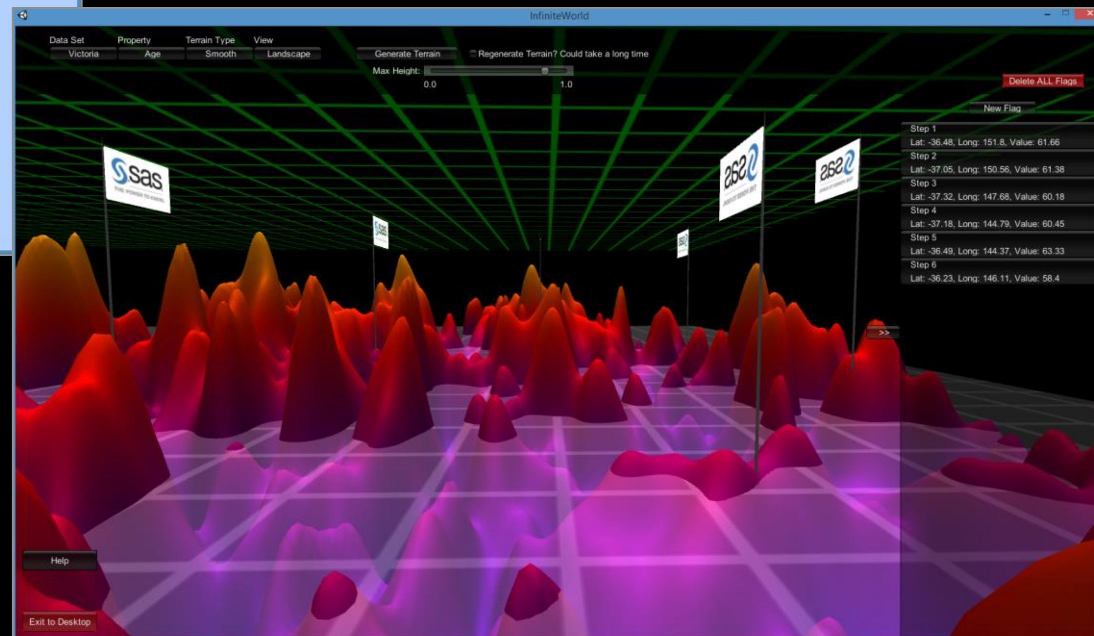


CONCEPTUALISATION
SENSEMAKING IN 2D OR 3D?



Visual Analyst 2D with Accidents Data

- What metaphors and narratives suit abstract data?
- What are the fundamental building blocks for creating metaphors and narratives?
- What are the principles of composing metaphors and narratives out of primitive elements?

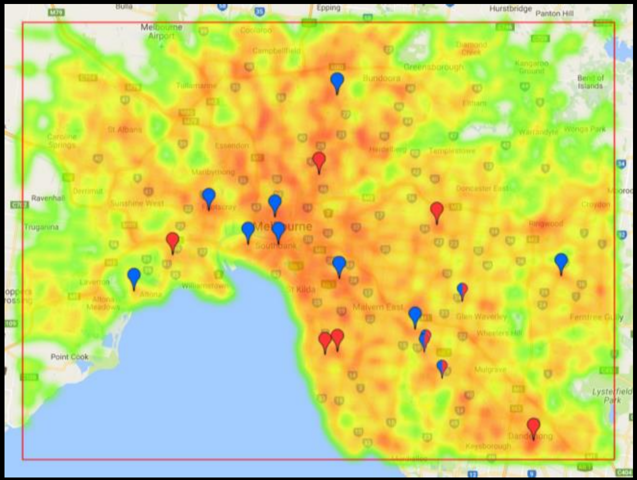


Visual Analyst 3D with Accidents Data - Age

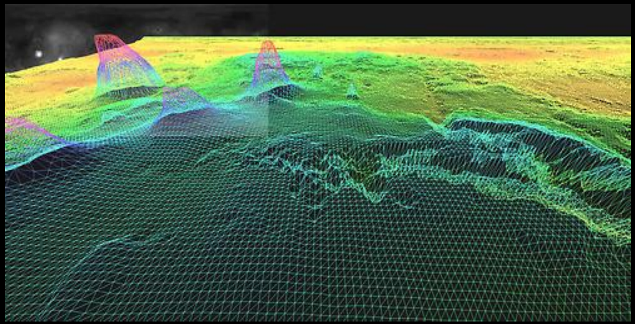
- How people interact with and make sense of 2D and 3D visual metaphors?
- What visual metaphors effectively support team work and creation of shared narratives?

TERRAIN IN DATA VISUALISATION

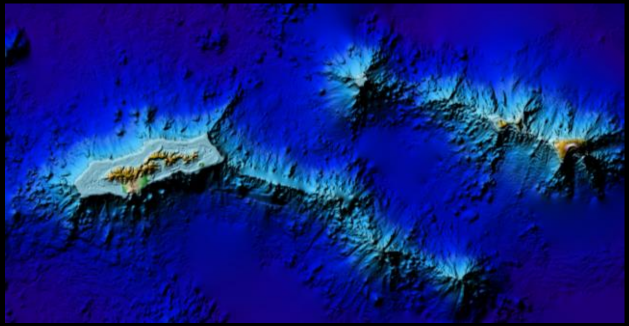
Data:
Abstract
Rich
Dynamic
Multi-dimensional
2D / 3D Representation



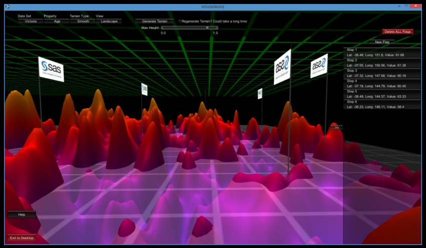
Theory
- Habitat theory
- Savanna theory
- Environmental aesthetics



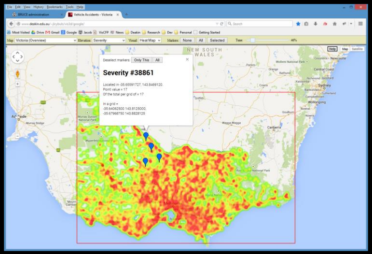
Technology Solution
- 2D and 3D
- Terrain Metaphor
- Immersion



Data Manipulation
- Interactivity
- Exploration
- Reporting



3D	Aspect	2D
Map and Landscape	**Data**	Map and Overlays
Data Point Proximity	**Similarity**	Data Point Proximity
Elevation	**Quantity**	Size
Colour Spectrum	**Quantity/Attribute**	Colour Spectrum
Viewing Angle Light and Shadow	**Perspective**	
Camera Angle Close and Far High and Low	**Aggregation/Detail**	Zoom
Layers	**Multi-Dimensions**	Overlays
Tuning	**Filter**	Tuning
Grids	**Selection**	
Flags Beacons Landing Spots	**Inspection/Annotation**	Markers
Trace	**Decisions**	
Tour	**Report**	
Movement	**Exploration**	Movement
Granular Terrain Smooth Terrain Columns	**Representation**	Bubble Chart Heat Map Choropleth
Landscape Peaks Lights	**Presentation**	Annotated Map
Keyboard / Mouse Body Motion Hand Motion	**Control**	Keyboard / Mouse
Multi-User	**Collaboration**	

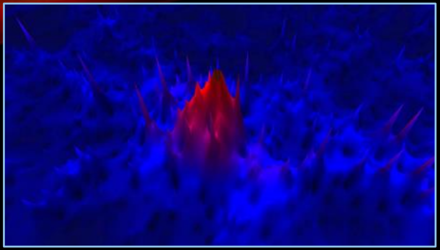


METAPHOR
2D VS 3D

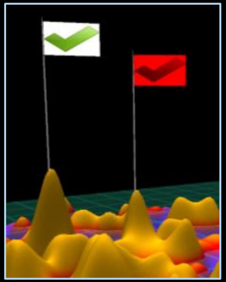
IN SEARCH OF
THE 3D METAPHOR

Experiments
with data terrains

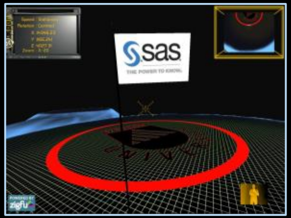
Peaks and
elevation



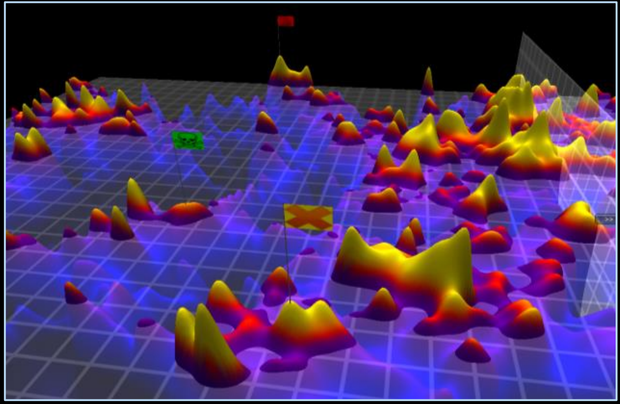
Flags
Markers

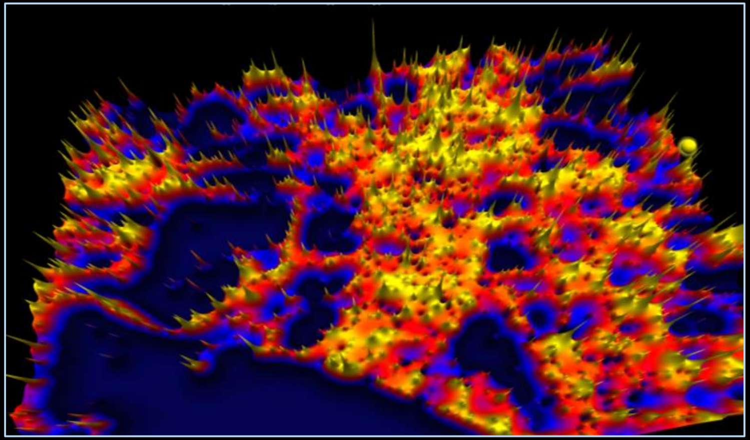


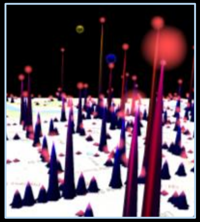
Landing spots



Grids
Surfaces

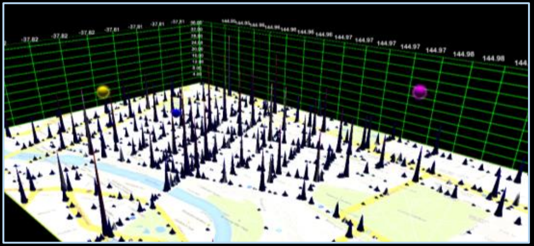




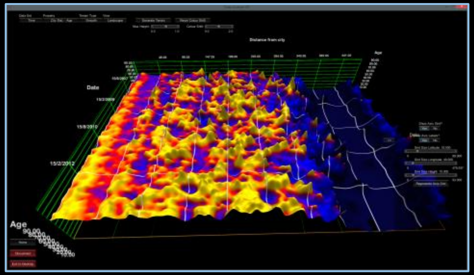


Lights

Maps &
Coordinates



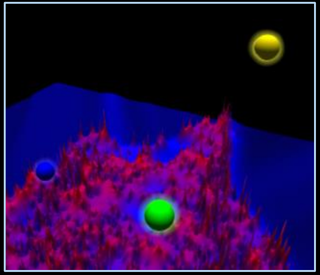
Heat, colour and
3D heat maps

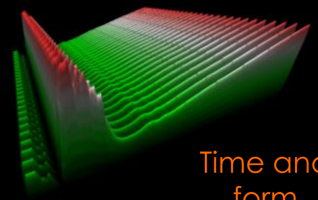


Time-Space

And more:
Traces
Beacons
Navigation
Text
Etc.

Multiuser
presence



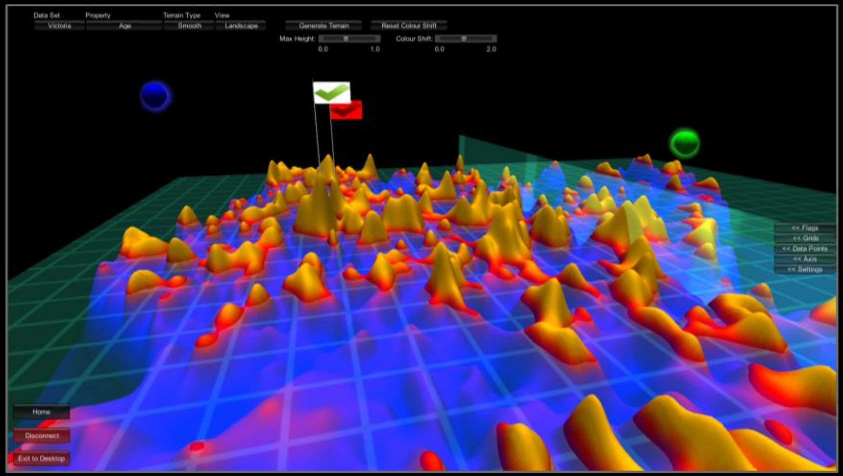


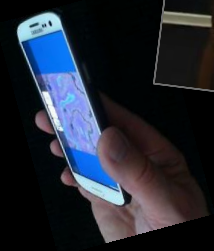
Time and
form

IN SEARCH OF THE NARRATIVE

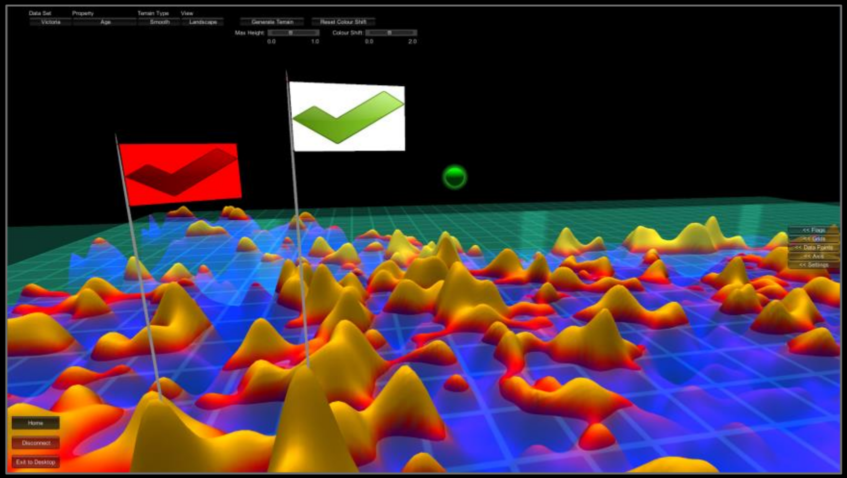
Narrative is a social phenomenon
that is constructed collaboratively

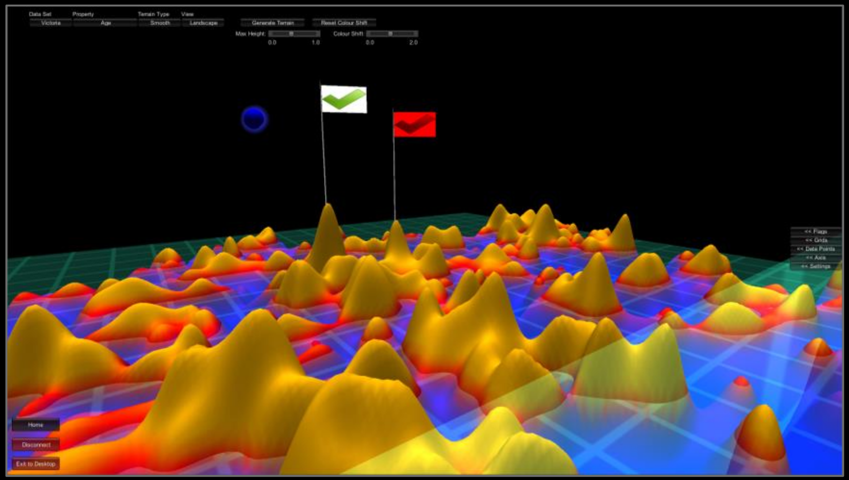
Average Age of Accident Victims in Victoria



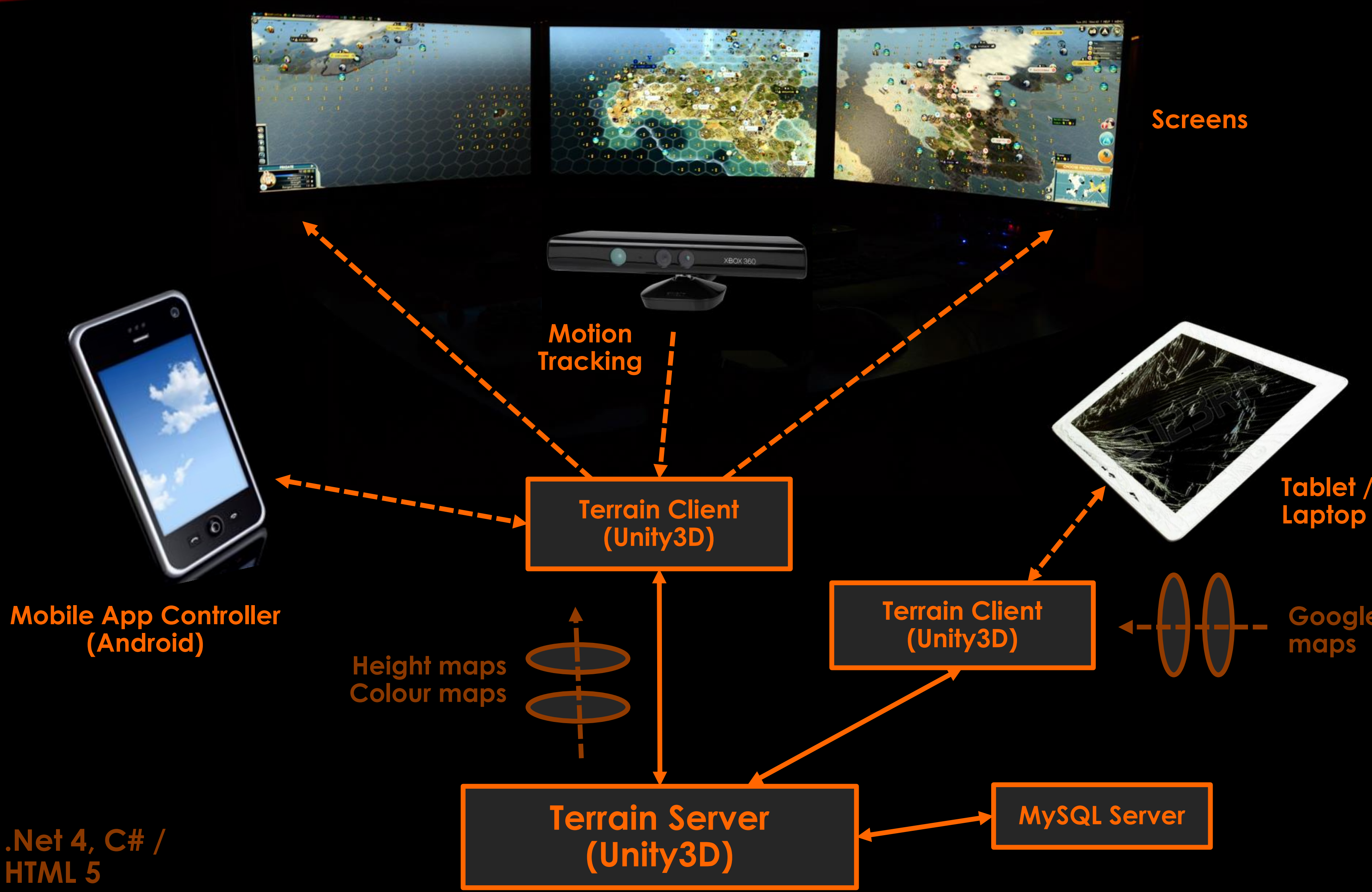


Visual Analyst 3D





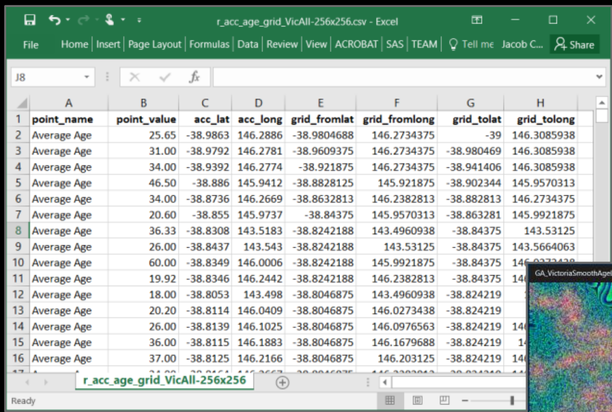
IMPLEMENTATION
VISUAL ANALYST 3D



.Net 4, C# /
HTML 5

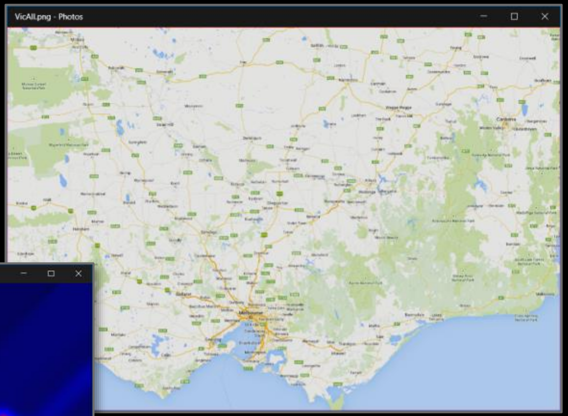
VTK?
ParaView?
VisIt?
VisNow?

DATA RENDERING
UNITY3D

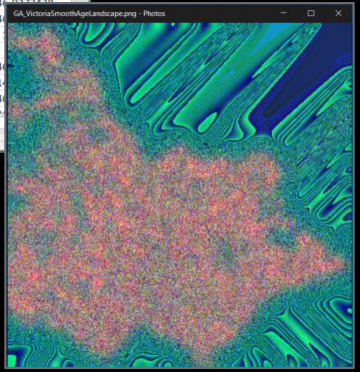


	A	B	C	D	E	F	G	H
1	point_name	point_value	acc_lat	acc_long	grid_fromlat	grid_fromlong	grid_tolat	grid_tolong
2	Average Age	25.65	-38.9863	146.2886	-38.9804688	146.2734375	-39	146.3085938
3	Average Age	31.00	-38.9792	146.2781	-38.9609375	146.2734375	-38.980469	146.3085938
4	Average Age	34.00	-38.9392	146.2774	-38.921875	146.2734375	-38.941406	146.3085938
5	Average Age	46.50	-38.886	145.9412	-38.8828125	145.921875	-38.902344	145.9570313
6	Average Age	34.00	-38.8736	146.2669	-38.8632813	146.2382813	-38.882813	146.2734375
7	Average Age	20.60	-38.855	145.9737	-38.84375	145.9570313	-38.863281	145.9921875
8	Average Age	36.33	-38.8308	143.5183	-38.8242188	143.4960938	-38.84375	143.53125
9	Average Age	26.00	-38.8437	143.543	-38.8242188	143.53125	-38.84375	143.5664063
10	Average Age	60.00	-38.8349	146.0006	-38.8242188	145.9921875	-38.84375	[illegible]
11	Average Age	19.92	-38.8346	146.2442	-38.8242188	146.2382813	-38.84375	[illegible]
12	Average Age	18.00	-38.8053	143.498	-38.8046875	143.4960938	-38.824219	[illegible]
13	Average Age	20.20	-38.8114	146.0409	-38.8046875	146.0273438	-38.824219	[illegible]
14	Average Age	26.00	-38.8139	146.1025	-38.8046875	146.0976563	-38.824219	[illegible]
15	Average Age	36.00	-38.8115	146.1883	-38.8046875	146.1679688	-38.824219	[illegible]
16	Average Age	37.00	-38.8125	146.2166	-38.8046875	146.203125	-38.824219	[illegible]

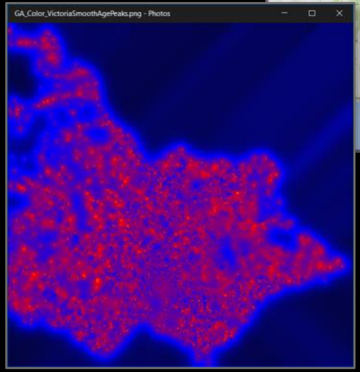
Selected Variables



Map

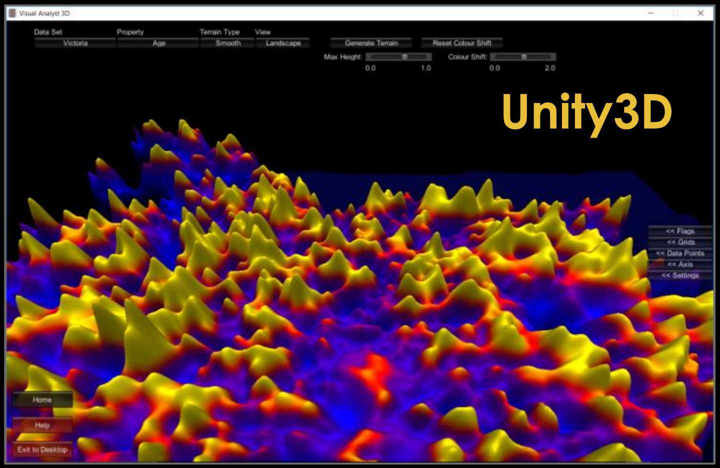


Height Map



Colour Map

As much of data to be rendered is missing, a genetic algorithm is used to create a smooth surface spanning the data points. This results in a 3D heat map.



Unity3D

Due to very large number of data points, rendering with height maps rather than meshes in Unity3D allows using efficient physics engine and surface interactions

EXPERIMENTS

Research aims

To determine the preferred methods of visual interaction with data to support sensemaking, problem-solving and decision-making. In particular:

1. Identify choices of data visual representation and the methods of data interaction.
2. Determine decision processes which rely on the visual cues provided by the data.
3. Identify the preferred combination of visual representations and interactions with the data.
4. Ascertain the impact of group decision making and individual experience on analytic performance.

Research participants

In total, 53 study participants were invited to conduct 29 experiments. 13 experiments were to study individual analysts. 8 experiments were conducted with co-located analytic teams and 8 with distributed teams. 10 experiments dealt with 2D and 19 with 3D.

All participants were asked to provide demographic information, plus to state their experience with data analytics and with complex interactive systems (such as games).

Tasks / Questions

About Victorian motor-vehicle accidents

Level 1 Learner

1. By looking at the data terrain, tell us in what areas of Victoria accidents involve the oldest people? Why do you think this is the case?
2. What are the towns or neighborhoods where the elderly are often involved in motor vehicle accidents?

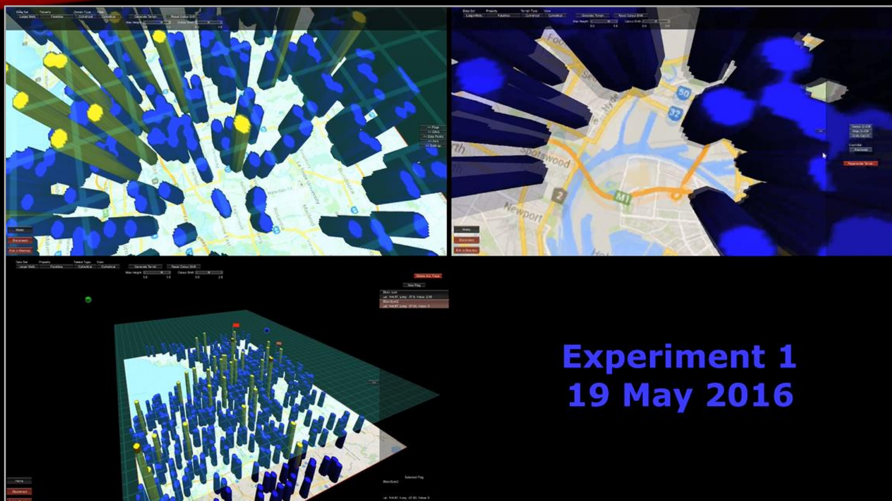
Level 2 Explorer

3. What 3 locations in Victoria have the oldest people involved in accidents? What is the number of people involved in each one?
4. What intersections in the City have the highest number of accidents?
   Why do you think this happens?

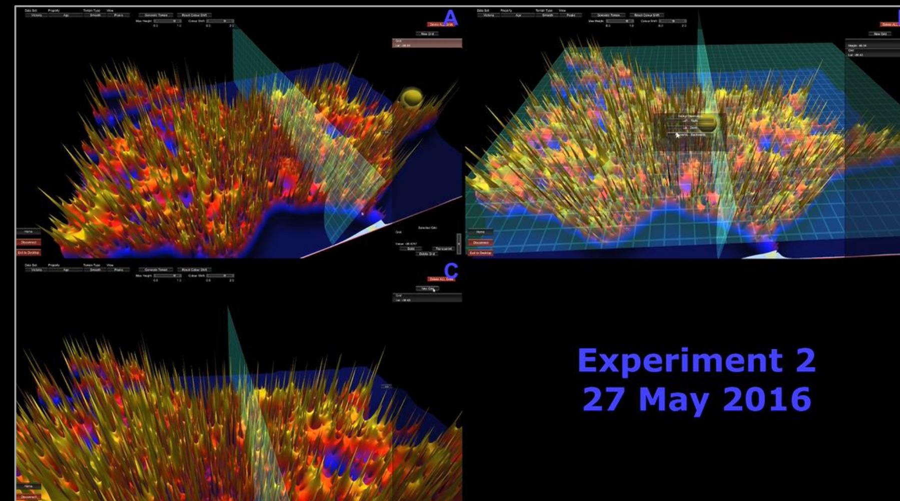
Level 3 Challenger

5. Where are the accidents black spots of Larger Melbourne? What criteria did you use?

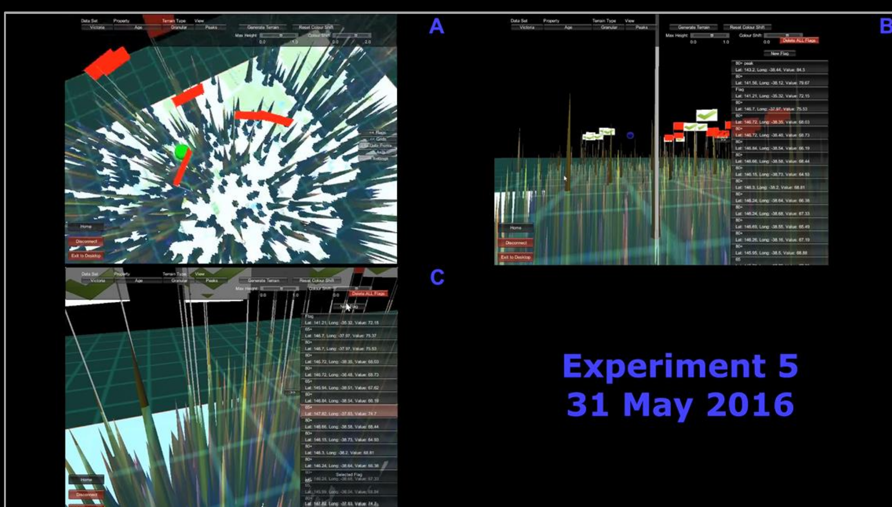
TYPICAL 3D INTERACTIONS



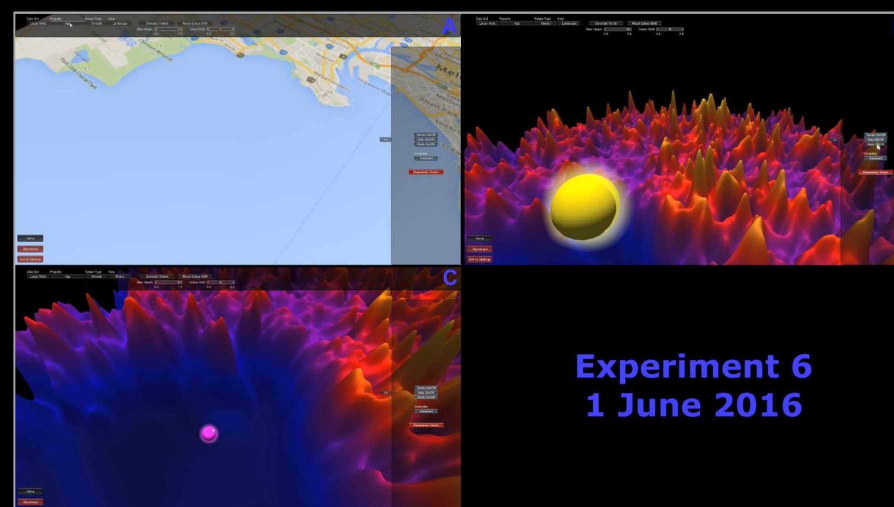
Experiment 1
19 May 2016



Experiment 2
27 May 2016



Experiment 5
31 May 2016

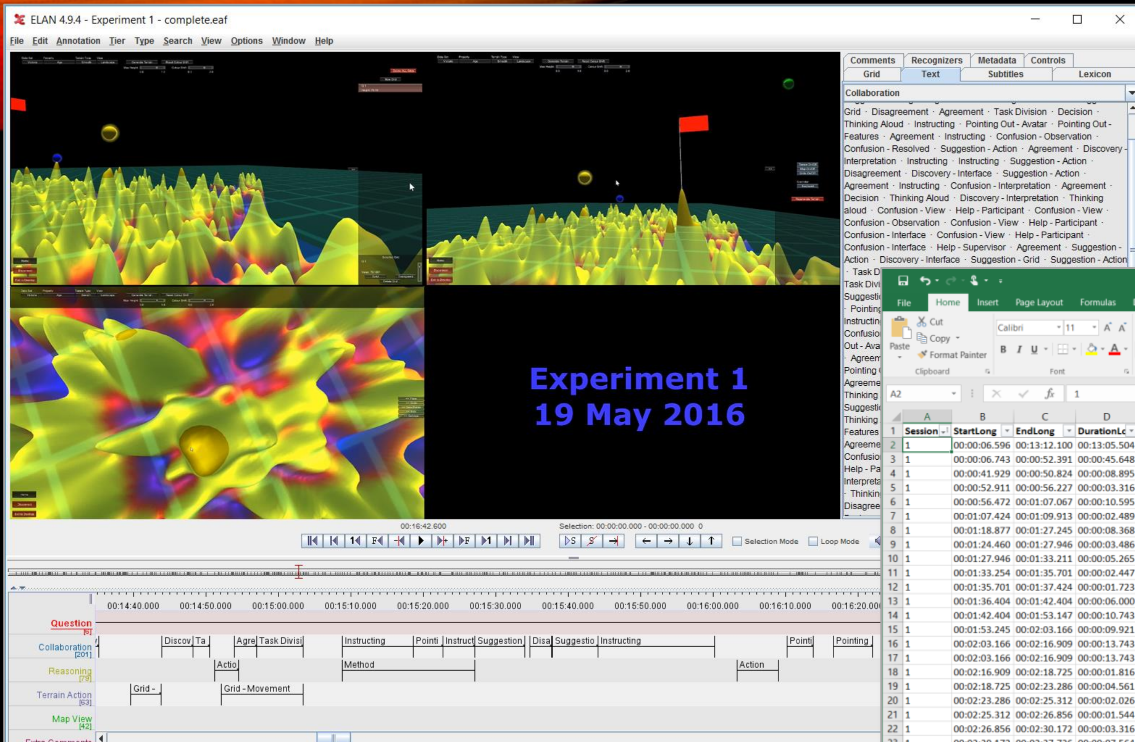


Experiment 6
1 June 2016

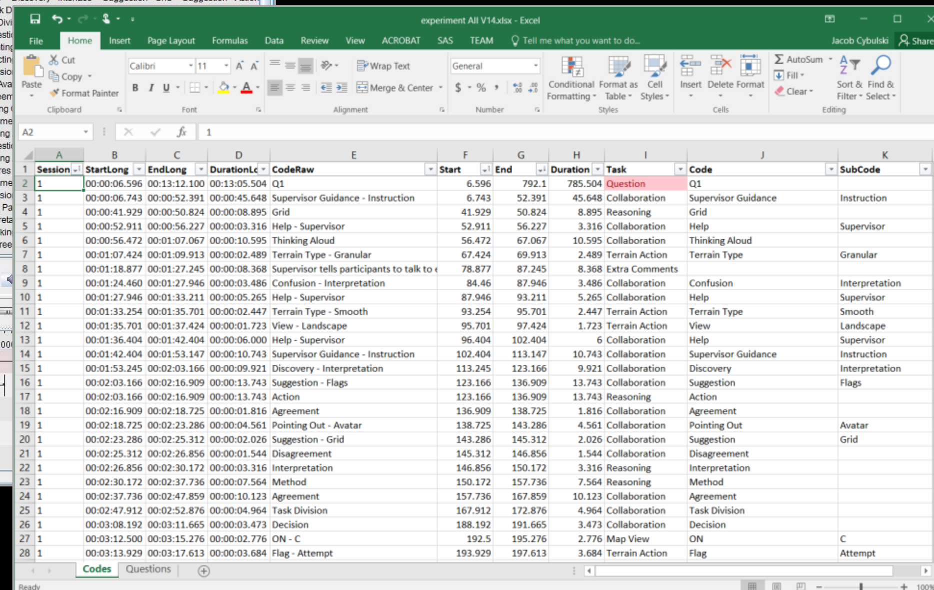
Teams opted for vastly different approaches to problem solving

RESEARCH METHODS

Video encoding in Elan



Field notes and codes



Session	StartLong	EndLong	DurationLc	CodeRaw	Start	End	Duration	Task	Code	SubCode
1	00:00:06.596	00:13:12.100	00:13:05.504	Q1	6.596	792.1	785.504	Question	Q1	
1	00:00:06.743	00:00:52.391	00:00:45.648	Supervisor Guidance - Instruction	6.743	52.391	45.648	Collaboration	Supervisor Guidance	Instruction
1	00:00:41.929	00:00:50.824	00:00:08.895	Grid	41.929	50.824	8.895	Reasoning	Grid	
1	00:00:52.911	00:00:56.227	00:00:03.316	Help - Supervisor	52.911	56.227	3.316	Collaboration	Help	Supervisor
1	00:00:56.472	00:01:07.067	00:00:10.595	Thinking Aloud	56.472	67.067	10.595	Collaboration	Thinking Aloud	
1	00:01:07.424	00:01:09.913	00:00:02.489	Terrain Type - Granular	67.424	69.913	2.489	Terrain Action	Terrain Type	Granular
1	00:01:18.877	00:01:27.245	00:00:08.368	Supervisor tells participants to talk to (	78.877	87.245	8.368	Extra Comments		
1	00:01:24.460	00:01:27.946	00:00:03.486	Confusion - Interpretation	84.46	87.946	3.486	Collaboration	Confusion	Interpretation
1	00:01:27.946	00:01:33.211	00:00:05.265	Help - Supervisor	87.946	93.211	5.265	Collaboration	Help	Supervisor
1	00:01:33.254	00:01:35.701	00:00:02.447	Terrain Type - Smooth	93.254	95.701	2.447	Terrain Action	Terrain Type	Smooth
1	00:01:35.701	00:01:37.424	00:00:01.723	View - Landscape	95.701	97.424	1.723	Terrain Action	View	Landscape
1	00:01:36.404	00:01:42.404	00:00:06.000	Help - Supervisor	96.404	102.404	6	Collaboration	Help	Supervisor
1	00:01:42.404	00:01:53.147	00:00:10.743	Supervisor Guidance - Instruction	102.404	113.147	10.743	Collaboration	Supervisor Guidance	Instruction
1	00:01:53.245	00:02:03.166	00:00:09.921	Discovery - Interpretation	113.245	123.166	9.921	Collaboration	Discovery	Interpretation
1	00:02:03.166	00:02:16.909	00:00:13.743	Suggestion - Flags	123.166	136.909	13.743	Collaboration	Suggestion	Flags
1	00:02:03.166	00:02:16.909	00:00:13.743	Action	123.166	136.909	13.743	Reasoning	Action	
1	00:02:16.909	00:02:18.725	00:00:01.816	Agreement	136.909	138.725	1.816	Collaboration	Agreement	
1	00:02:18.725	00:02:23.286	00:00:04.561	Pointing Out - Avatar	138.725	143.286	4.561	Collaboration	Pointing Out	Avatar
1	00:02:23.286	00:02:25.312	00:00:02.026	Suggestion - Grid	143.286	145.312	2.026	Collaboration	Suggestion	Grid
1	00:02:25.312	00:02:26.856	00:00:01.544	Disagreement	145.312	146.856	1.544	Collaboration	Disagreement	
1	00:02:26.856	00:02:30.172	00:00:03.316	Interpretation	146.856	150.172	3.316	Reasoning	Interpretation	
1	00:02:30.172	00:02:37.736	00:00:07.564	Method	150.172	157.736	7.564	Reasoning	Method	
1	00:02:37.736	00:02:47.859	00:00:10.123	Agreement	157.736	167.859	10.123	Collaboration	Agreement	
1	00:02:47.912	00:02:52.876	00:00:04.964	Task Division	167.912	172.876	4.964	Collaboration	Task Division	
1	00:03:08.192	00:03:11.665	00:00:03.473	Decision	188.192	191.665	3.473	Collaboration	Decision	
1	00:03:12.500	00:03:15.276	00:00:02.776	ON - C	192.5	195.276	2.776	Map View	ON	C
1	00:03:13.929	00:03:17.613	00:00:03.684	Flag - Attempt	193.929	197.613	3.684	Terrain Action	Flag	Attempt

- The research adopted Hermeneutic Ethnographic approach (Harvey and Myers 1995, Myers 1999)
- Data was collected in the form of video recordings of analytic sessions
- Videos were coded using ELAN - software tool for video annotation
- Thematic analysis of coded annotations resulted in structural representation of analysts' behaviour

Resulting themes and codes

Question:	Q1, Q2, Q3, Q4, Q5, Feedback
Terrain Action:	Terrain Type, View, Data Set, Data points, Property, Flag, Grid, Axis, Colour Shift
Collaboration:	Task Division, Confusion, Agreement, Disagreement, Confirming, Decision, Discovery, Pointing Out, Suggestion, Thinking Aloud, Help, Instructing, Supervisor Guidance
Reasoning:	Action, Flags, Grid, Inaccuracy, Interface, Method, Interpretation, Result
Map View:	ON, OFF

EXPERIMENTAL RESULTS
(NOT INCL. DISTRIBUTED TEAMS)

Retrospection and Ongoing Improvement:

- Analytic process never stops, thus participants constantly modify answers to earlier questions, after developing new insights in the subsequent analyses.
- In general, a healthy degree of skepticism is maintained toward results, motivating participants to improve answers.

Cue extraction:

- In spite of different data manipulation experiences, there were significant commonalities in extracting visual cues.
- Extracting cues in 3D environment needed lots of interaction with the tool and data which consumed considerable amount of time.
- There was tendency to reuse visual settings from one problem to another, often leading to deployment of inappropriate methods, regardless of the problem space, dimensionality or question asked.

Social:

- Engagement with visualised data is highly dynamic with control passing between team members. This highlights need for good user interface design to support effective team communication and collaboration, otherwise it limits collaboration.
- By interacting with visualisations, participants discovered more data attributes and solved problems better and more effectively when working in teams as compared with working individually.

Plausibility:

- Participants are commonly aware of issues associated with ambiguous, vague and insufficient data.
- Resulting answers were often imprecise, yet they were realistic, which highlighted the need for quality data to generate quality results.
- Skepticism towards the tool prevents clarification and precision.

Enactment:

- Participants working with 3D representations were highly dependent on interaction with visualised data.
- The aim of 3D interaction was to create a visual environment that directly provided information that made sense of the data.
- Answers derived from the 3D data visualization were highly cohesive, in spite of vastly different methods of visualization and interaction, as well as, personalization of user interfaces.

Identity creation:

- Identity is deeply rooted in personal experience and personality.
- Those skilled in computer games are comfortable with 3D interaction, unlike those with experience with traditional 2D business charts.
- Skeptical users lack motivation and interest in working with complex or novel interfaces. Moderate skepticism, nevertheless, can drive some 3D users to explore the tools and its available options.

Analysis of results obtained from experiments with distributed teams has recently been completed but not reported as yet.

RESULTS & REFLECTION

The main purpose of collaborative data visualisation is to assist teams in making sense of the presented data and the the past or current events it represents.

Sensemaking provides opportunities for users to enact their identity on data environment via visualisation tools. Users can shape and adapt the environment or the methods of interaction to suit their skills, experience and personality.

Visual interaction with data assist sensemaking processes in extracting cues from visual forms, such as colour, elevation or proximity. These may include markers or flags, grids, filters, sections, navigation and zooming, as well as, annotations and traces.

Extracted visual cues and derived insights are merely plausible, and so experience, repetitions and refinements are needed to become proficient in an ongoing process of sensemaking.

Sensemaking is a social phenomenon and team sensemakers usually perform better and generate higher quality insights. Yet, collaboration needs to be carefully planned, as it may lead to positive or negative effects.

Trust in data and tools promote effectiveness of analytic tasks. A degree of skepticism in obtained results assist experimentation and seeking behavior. Total skepticism hinders clarification and improvement of insights.

Recent projects:

Teaching Data Analytics (OLT)
Sensemaking and Legitimisation (ICAA)
Oil and Gas Exploration Project (SAS)
Collaborative analytics (SAS VAC / OLT)

Data sets:

Sales of movie tickets
Music distribution
Motor vehicle accidents
Oil exploration data
Share market dynamics
Real estate auctions

Current and future work:

Study comparing 2D and 3D IVA
Shared and private analytic spaces
Passer-by analytics
Dynamic data and text
Using HoloLens for 3D vis and control
Utilisation of eye tracking for 3D coding

Visual Analyst 3D with Accidents Data - Gender balance

PRIMARY REFERENCES

- Boland, R.J. (2008). Decision making and sensemaking. In F. Burstein & C. W. Holsapple (Eds.), *Handbook on Decision Support Systems* (Vol. 1, pp. 55-63). New York: Springer.
- Cybulski, J.L., Keller, S., Nguyen, L. and Saundage, D. (2015). Creative Problem Solving in Digital Space Using Visual Analytics. *Computers in Human Behavior* 42: 20–35.
- Cybulski, J.L., Keller, S. and Saundage, D. (2015). Interactive Exploration of Data with Visual Metaphors. *International Journal of Software Engineering and Knowledge Engineering* 25, no. 02, 231–52.
- Grady, J. (2007). Metaphor, in *The Oxford handbook of cognitive linguistics*, D. Geeraerts and H. Cuyckens, Eds. Oxford; New York: Oxford University Press.
- Harvey, L.J. and Myers, M.D. (1995): Scholarship and Practice: The Contribution of Ethnographic Research Methods to Bridging the Gap. *Information Technology & People* 8, no. 3 13–27.
- Lakoff, G. and Johnson, M. (2003). *Metaphors we live by*. Chicago: University of Chicago Press.
- Myers, M.D. (1999). Investigating Information Systems with Ethnographic Research. *Communications of the AIS* 2, no. 4es, 1.
- Namvar, M., Cybulski, J.L. and Perera, L. (2016). Using business intelligence to support the process of organizational sensemaking. *Communications of the Association for Information Systems*. 38(1), Article 20.
- Weick, K. E. (1993). *The Collapse of Sensemaking in Organizations: The Mann Gulch Disaster*, Administrative Science Quarterly, 38(4), 628-652.
- Weick, K. E. (1995). *Sensemaking in organizations*. Thousand Oaks, CA: Sage.
- Weick, K. E. Sutcliffe, K. M., & Obstfeld, D. (2005). Organizing and the process of sensemaking. *Organizational Science*, *16*(4), 409-421.