Big Data for smart infrastructure: London Bridge Station Redevelopment

Sinan Ackigoz & Krishna Kumar
10.09.15 – Cambridge, UK

CSIC Cambridge Centre for Smart Infrastructure and Construction
EPSRC Engineering and Physical Sciences Research Council
Technology Strategy Board Driving Innovation
UNIVERSITY OF CAMBRIDGE

Redeveloping the redeveloped station

1972 Vision: Two old railway stations will be merged into one with a higher capacity, giving easy interchange between buses, tube and trains – and direct access to all service from the spacious concourse with new bars and shops.

2012 vision: The number of platforms will increase and track layout will accommodate higher capacity trains. Existing bus, train and underground services will be linked with the largest concourse in the UK

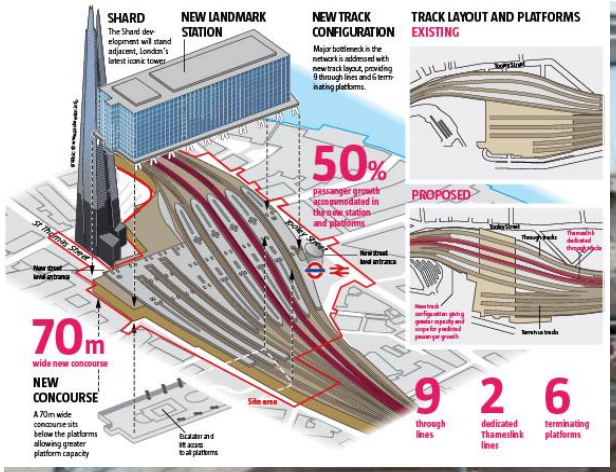
Challenges and Opportunities

Challenges:

- To forecast the future demands on current infrastructure – we need data!
- To build lasting infrastructure – we need data!

Opportunities with improved sensing (and a lot more data):

- Quantify the performance of current infrastructure
- Evaluate long term performance of our structures



Active train station (Network Rail)

Critical lifelines (Weston St. sewer) and important structures (Shard, Guy's Hospital, SER Offices) nearby

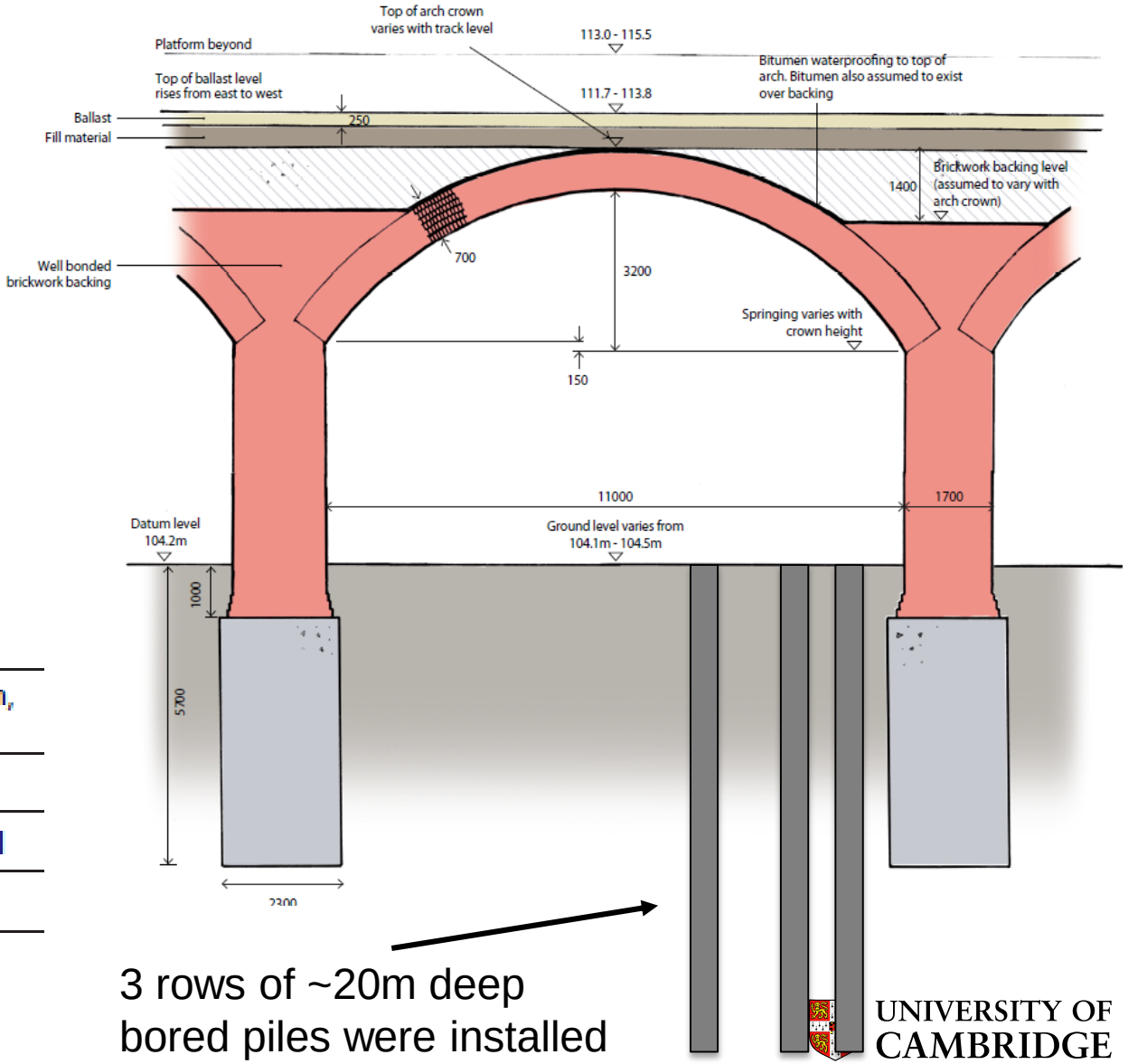
High risk construction activity (Costain)

Active Jubilee Line and tunnels 30m below (London Underground)

[1] retrieved from www.paulweston.info on 04.09.14

The structure – Arch E951

Built in:
1893
Span, Depth and Arch Thickness:
10m, 3.2m, 0.7m
State:
Uncracked and no water drainage problems



Key:

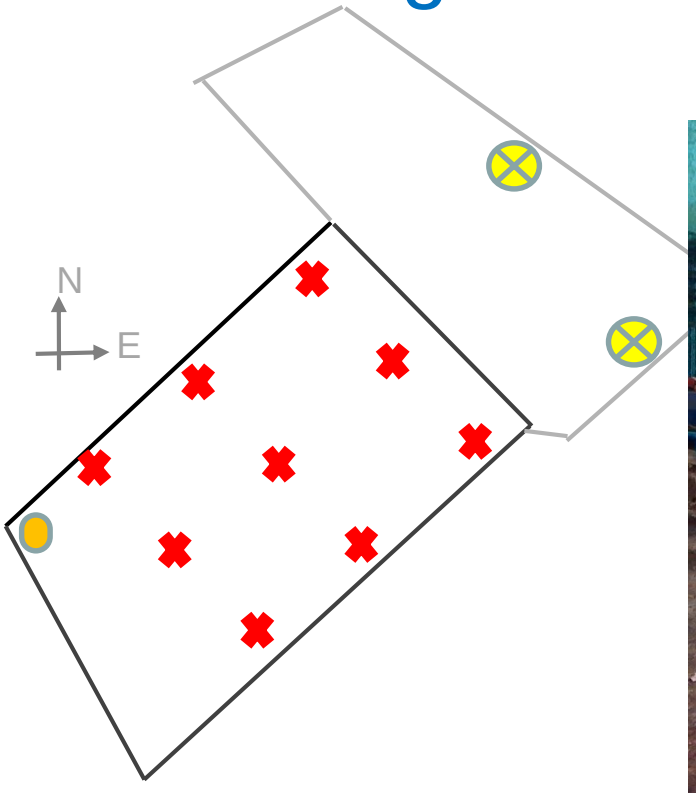
	Bricks (dark=in section, light=in elevation)
	Concrete (dark=in section, light=in elevation)
	Ballast
	Made ground/fill material
	Ground/earth general

3 rows of ~20m deep bored piles were installed

CSIC Cambridge Centre for Smart Infrastructure and Construction

UNIVERSITY OF CAMBRIDGE

Point sensing: which point exactly?



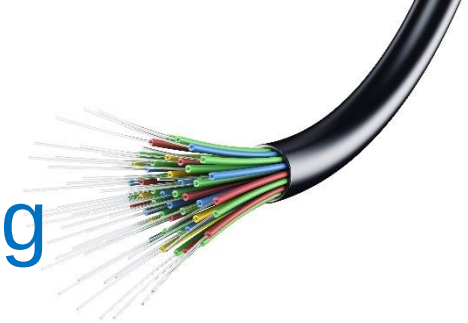
Plan view of Arch E951

Automatic total station

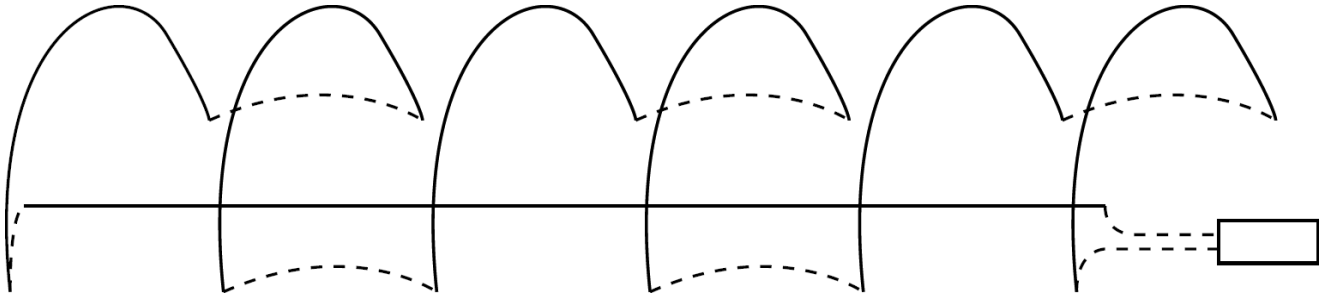
Prism target

Control target (no movement)

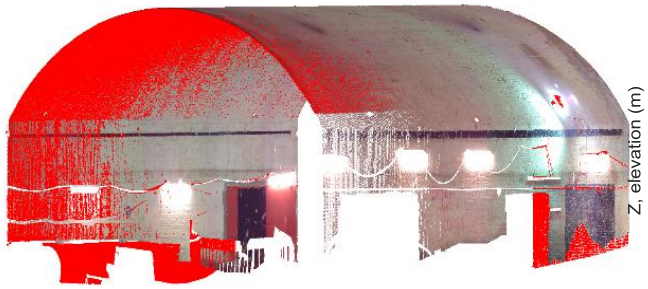
A paradigm shift:
Point Sensing to Distributed Sensing

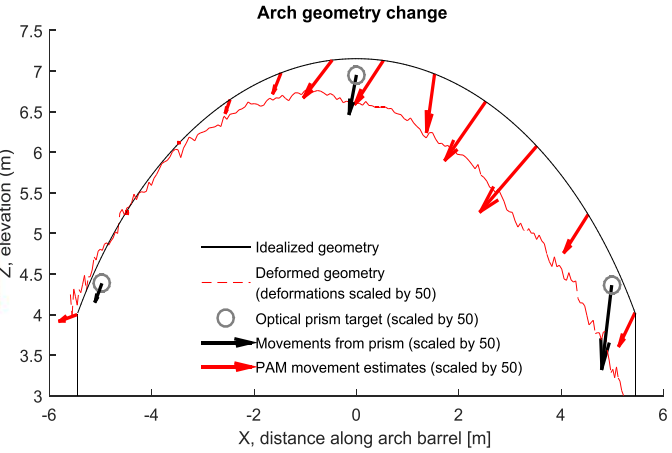


One sensor for all
Fibre optic sensors give us the opportunity to detect cracks and damage as they form



Learning from data
Displacement data from laser scanners give us a four dimensional understanding of complex masonry response

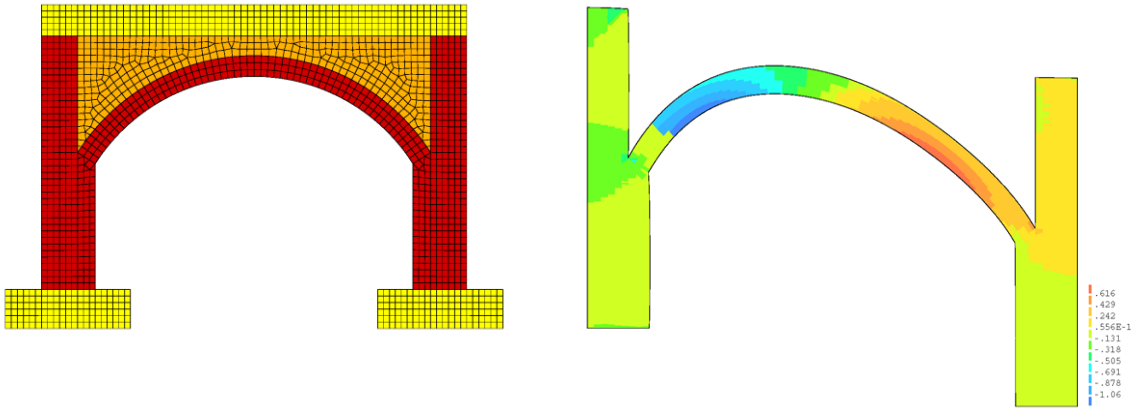


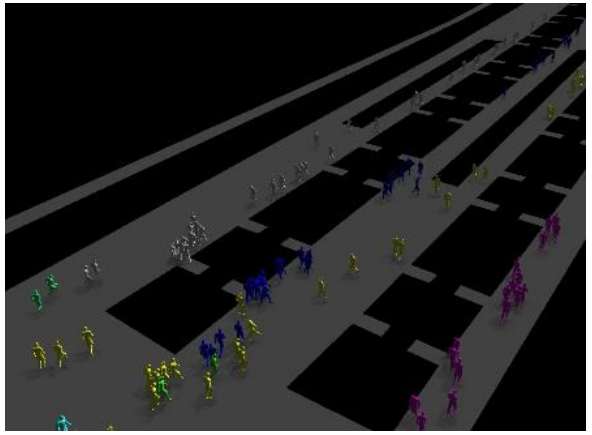


CSIC Cambridge Centre for Smart Infrastructure and Construction

UNIVERSITY OF CAMBRIDGE

Making unpredictable predictable

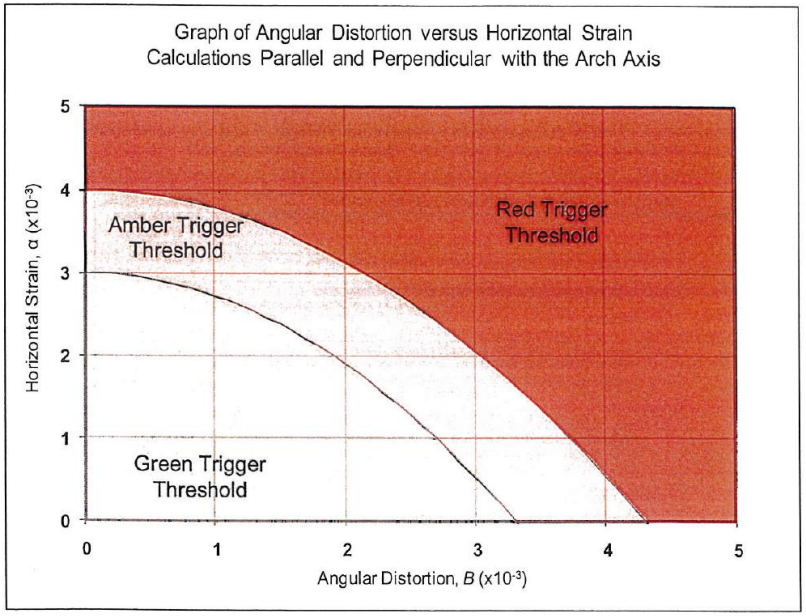




Improve our analytical/predictive capabilities with real data
Laser scanner data allowed us to formulate improved structural models which assess the safety limits for settlements.

Use data in real-time to predict events faster than time
By integrating depth sensing and RGB camera data, we can simulate passenger congestions in London Bridge.

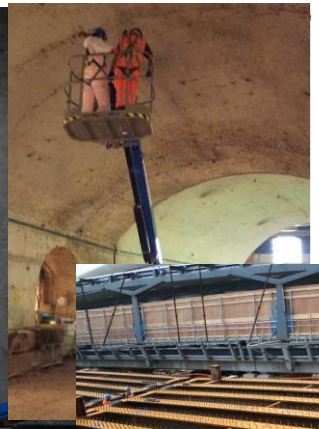
Manage assets for the future

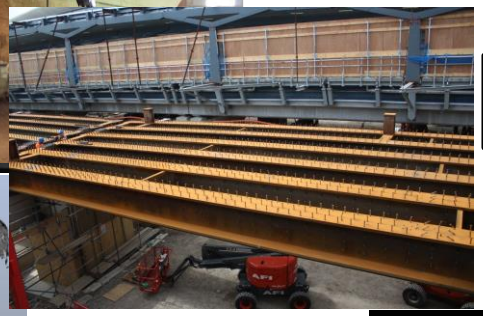


Improve assessment or design guidelines: The current masonry arch bridge assessment methods are insufficient – can we improve them?

Enhance public experience when using infrastructure: Predict crowd flows in city-scale and simulate options to manage traffic to prevent passenger congestion.

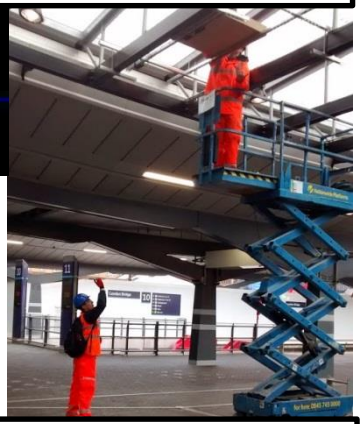
FO Monitoring

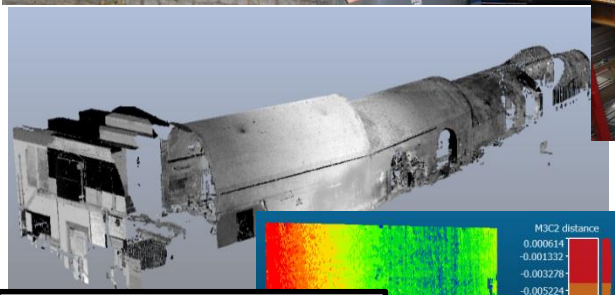




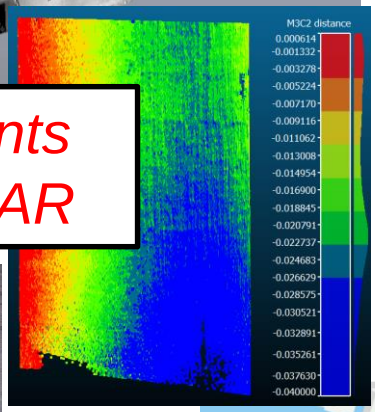
Real time 3D model construction

Real time people monitoring





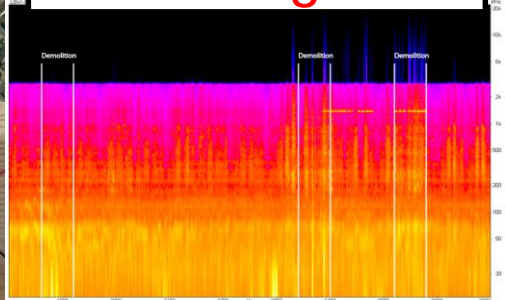
Movements from LIDAR



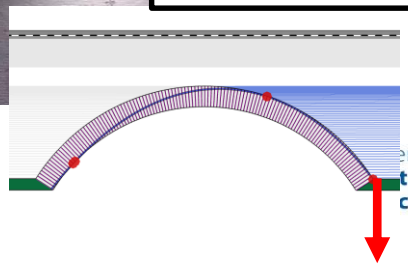
Real time people movement prediction

Social Media tracking

Wireless Noise monitoring



Analysis



Monitoring a 'complex infrastructure'

Surveying data from Costain for Arches and Tunnels

Rail track displacement data from Network Rail

Fibre optics data from CSIC

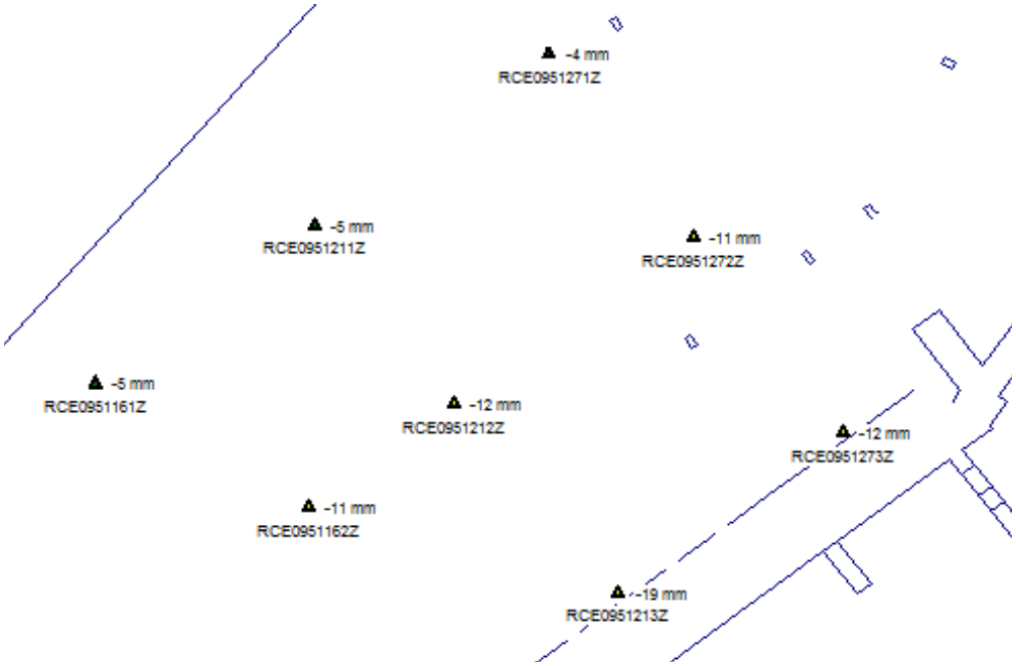
Design predictions from WSP

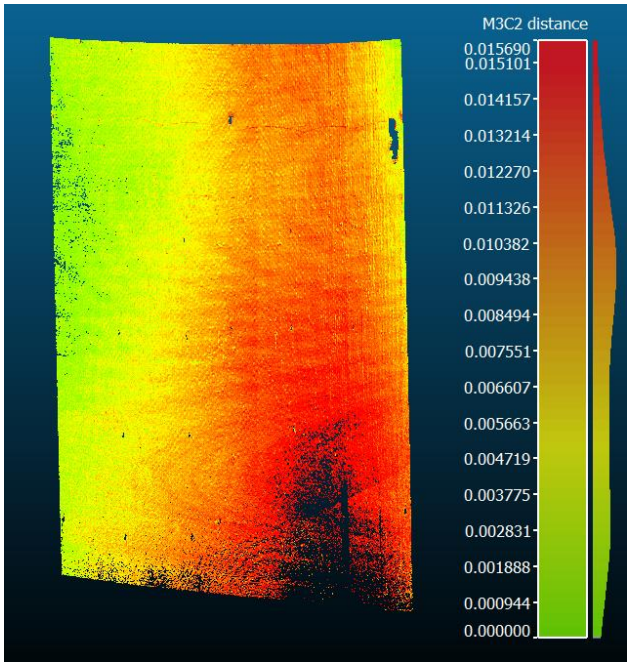
Cloud comparison data from CSIC

Can we link all this information to:

1) Retrieve a better understanding of response
2) Improve communication between agents

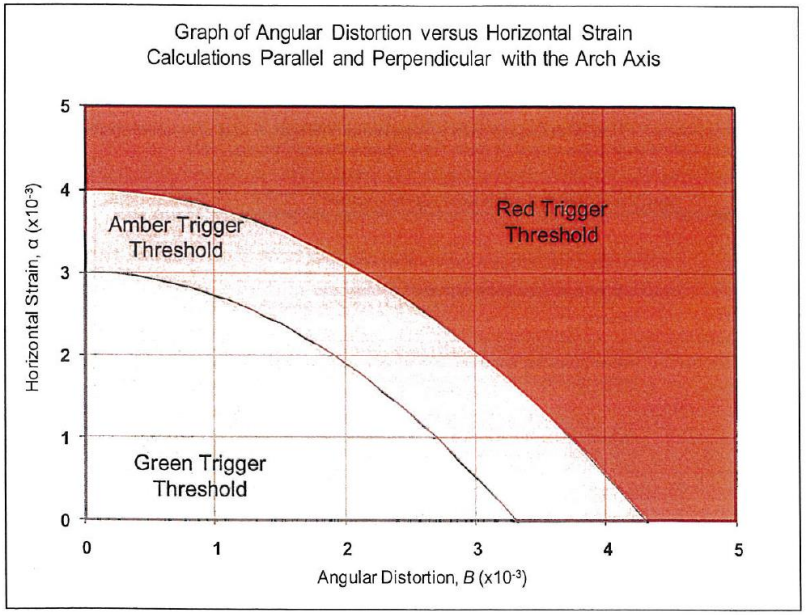
Data exists in different formats for different metrics, what do we do to link them?





Dilemma about distributed sensing -> How do we describe it and link it to other data

Raw data is not of interest, evaluated data with trigger levels interest stakeholders more



Measurement		Status	Network Rail Trigger Threshold[1+3]		Actions Required
Track 8, 9, 10, 11, 12, 13, 14, 15, 16 [+2]	Cant	Green	<+/-20		No Action.
		Amber	>+/-20 and <+/-50		DOE to Inform London Bridge Track Maintenance Engineer (TME) & DPE, no other action required.
		Red	>+/-50		DOE to call Infrastructure Control Centre (ICC - 02076206336) and inform of a 'Cant Intervention Level' has been Exceeded.
	Twist	Green	Level 1	<+/-15	No Action.
		Amber	Level 2	>+/-15 and <+/-24	As per level 2 fault.
		Red	Level 3	>+/-24 and <+/-33	As per level 3 fault.
		Black	Level 4	>+/-33	As per level 4 fault.
	Horizontal Distance Between Platform Edge and Track (kinematic envelope)	Green	<+/-20		No Action.
		Amber	>+/-20 and <+/-50		As per level 2 fault.
		Red	>+/-50		As per level 3 fault.
	Displacement Fault	Green	<+/-25		No Action
		Amber	>+/-25 and <+/-35		As per level 2 fault.
		Red	>+/-35		As per level 3 fault.

Dilemma about assessment of the data – should we link alarms and triggers? That requires data processing but is an important aspect that needs to be discussed.

Querying

- At London Bridge Arch E951 area, has there been a black alarm triggered during the piles constructed 12.13-01.14?
- Is there a crack opening greater than 5mm in the arches, when the settlement values reached 50mm?
- What is the maximum displacement observed during tunneling (or other construction activity) within x meters from the pile.
- Number of data readings required for a strain profile – within a tolerance of X?
- Effect of analyzer X on scatter / noise in soil layer Y

Spatio-Temporal Query Processing

- Semantic analytics tools have primarily focused on thematic relationships, but spatial and temporal relationships are often critical components in analytical domains.
- *Current GIS and spatial database technology does not support complex thematic analytics operations*. Traditional data models used for GIS excel at modelling and analysing spatial and temporal relationships among geospatial entities but tend to model the thematic aspects of a given domain as directly attached attributes of geospatial entities. Thematic entities and their relationships are not explicitly and independently represented, making analysis of these relationships difficult.

Why conventional relational dbs won't cut it?

Sensor	Location	Date/Time	Displacement X	Displacement Y	Displacement Z
42	51°N 0°30W	20150910120000	2	1.5	3
42	51°N 0°30W	20150910130000	2.1	1.5	3

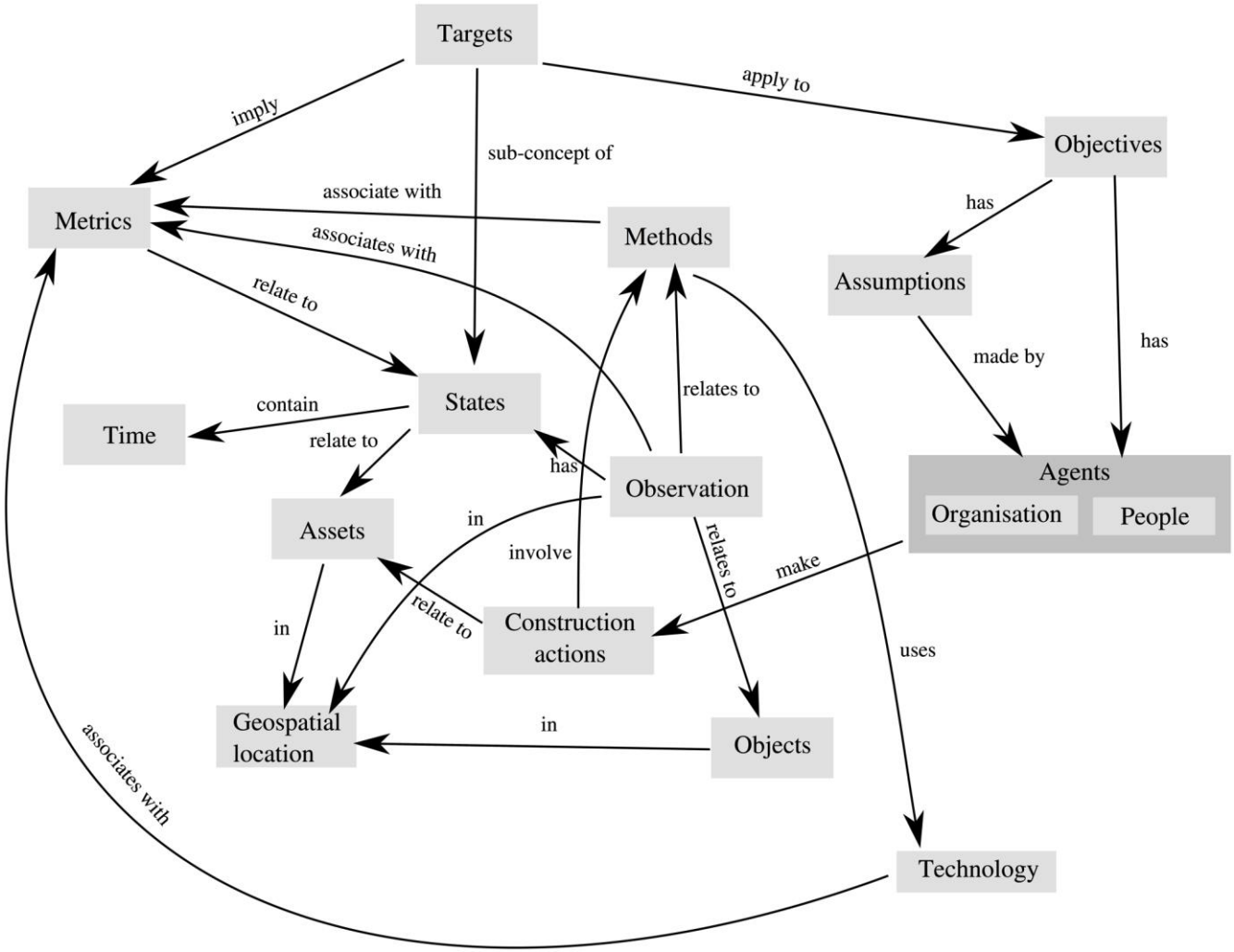
Sensor	Location	Date/Time	Temperature
57T	51°N 0°30W	201509101230 0	24

Humidity? Vibration? 3D point cloud data? …

What are smart cities standards?

- Smart cities standard (PAS 182) offers a handful of generic concepts (such as *place*, *observation, metric* etc.) to formulate a common language for linking data across organizations in a city.
- The main goal is to seamlessly integrate information from multiple organisations using *Linked Data (Semantic Web technology)* that can be shared and edited. Different from BIM, Smart Cities allow for the sharing of critical but less tangible aspects that relate to the data: such as related *Assumptions* and *Objectives*.

Concept model

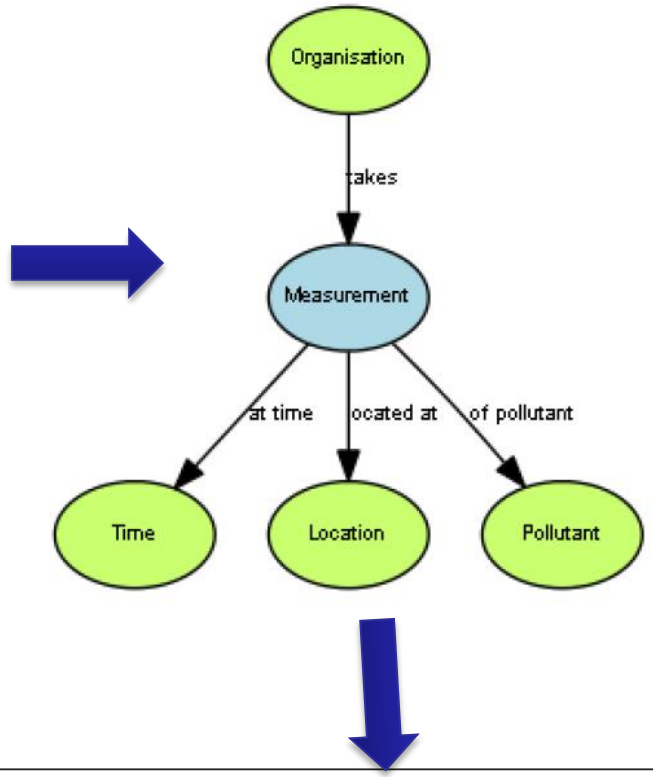


Potential benefits?

1. Real-time infrastructure monitoring and prediction system – save lives and infrastructure.
2. A smart framework that can correlate sensor data with design and analysis to understand the interactions between infrastructures and people – Better engineering understanding
3. Multiple organizations/public benefit from gathered data that is in a **universal format** and organically linked to other critical information
4. Improve communication between agents by uniformly representing **data/related aspects** (eliminate redundancies and misunderstandings)

Bristol Air Pollution Measurement

	A	B	C	D	E	F	G	H	I	J	K
1	SiteID	Date	Time	NO2raw_ppb	NOxraw_ppb	NOraw_ppb	NO2rat_ppb	NOxrat_ppb	NOrat_ppb	CAQI	
2	203	01/01/2001	00:30:00	13.5	65	50.5	4.73	19.44	14.71	2	
3	203	01/01/2001	00:45:00	13	55.5	41.5	4.58	16.57	11.99	2	
4	203	01/01/2001	01:00:00	-10	-10	-10				2	
5	203	01/01/2001	01:15:00	-10	-10	-10				2	
6	203	01/01/2001	01:30:00	12.5	46	32.5	4.43	13.7	9.27	2	
7	203	01/01/2001	01:45:00	12.5	36.5	23.5	4.28	10.83	6.55	2	
8	203	01/01/2001	02:00:00	11	54.5	43	3.82	16.26	12.44	2	
9	203	01/01/2001	02:15:00	9	44	35.5	2.92	13.09	10.17	2	
10	203	01/01/2001	02:30:00	10.5	35.5	24	3.82	10.52	6.7	2	
11	203	01/01/2001	02:45:00	12	34	22	3.97	10.07	6.09	2	
12	203	01/01/2001	03:00:00	11.5	32	20	3.97	9.47	5.49	2	
13	203	01/01/2001	03:15:00	11.5	40.5	28	4.13	12.03	7.91	2	
14	203	01/01/2001	03:30:00	13	32	18	4.58	9.47	4.89	2	
15	203	01/01/2001	03:45:00	11	42.5	31.5	3.67	12.64	8.97	2	
16	203	01/01/2001	04:00:00	12.5	31.5	18	4.43	9.31	4.89	2	
17	203	01/01/2001	04:15:00	11	23	11.5	3.82	6.75	2.92	2	
18	203	01/01/2001	04:30:00	11	17.5	6	3.82	5.09	1.26	2	
19	203	01/01/2001	04:45:00	11	19	7	3.97	5.54	1.56	2	
20	203	01/01/2001	05:00:00	10	12.5	2.5	3.37	3.58	0.2	1	
21	203	01/01/2001	05:15:00	8.5	10	1.5	2.92	3.07	0.15	1	
22	203	01/01/2001	05:30:00	10	12.5	2.5	3.37	3.82	0.45	1	
23	203	01/01/2001	05:45:00	8.5	11	2.5	2.92	3.37	0.45	1	



<http://bristol-data-epimorphics.dyndns.org/data/environment/air-quality/observation/conc/location/00203/date/1999-01-01/time/00:00:00/duration/PT15M/substance/NOX>

<http://bristol-data-epimorphics.dyndns.org/def/environment/air-quality/core/pollutant>

http://bristol-data-epimorphics.dync

NOx - Nitrogen oxide (mono or di)

Show Search Form

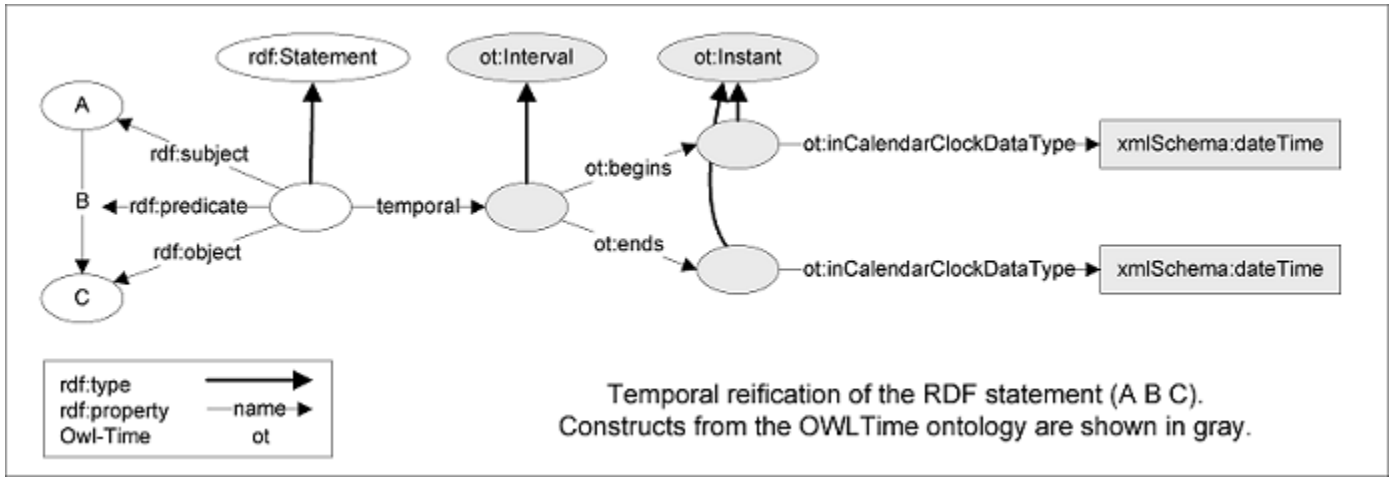
http://bristol-data-epimorphics.dyndns.org/def/environment/air-quality/air-pollutant/NOX

type	Pollutant Concept
is defined by	air-pollutant.ttl

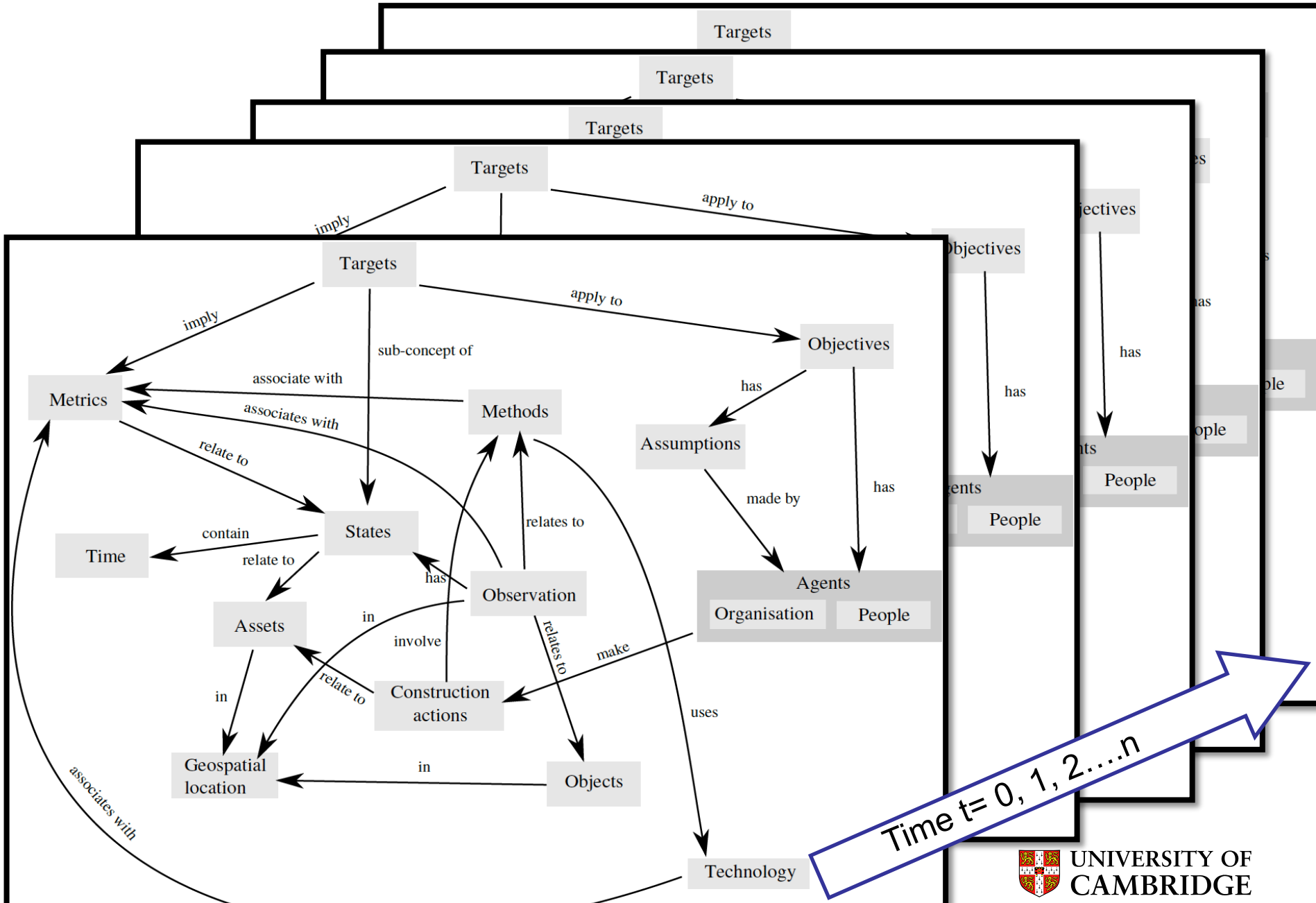
CSIC Cambridge Centre for Smart Infrastructure and Construction

Temporal RDF

- Temporal RDF extends the RDF statement from a triple to a quad where the fourth element is the valid time of the RDF statement



Time Slices

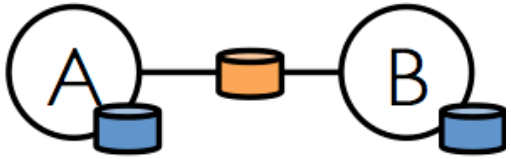


UNIVERSITY OF CAMBRIDGE

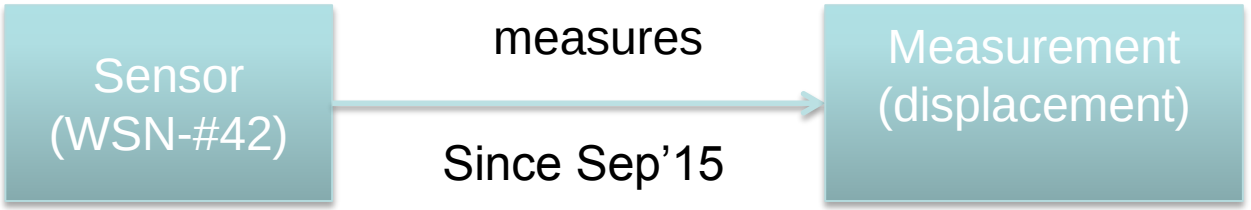
Graph processing

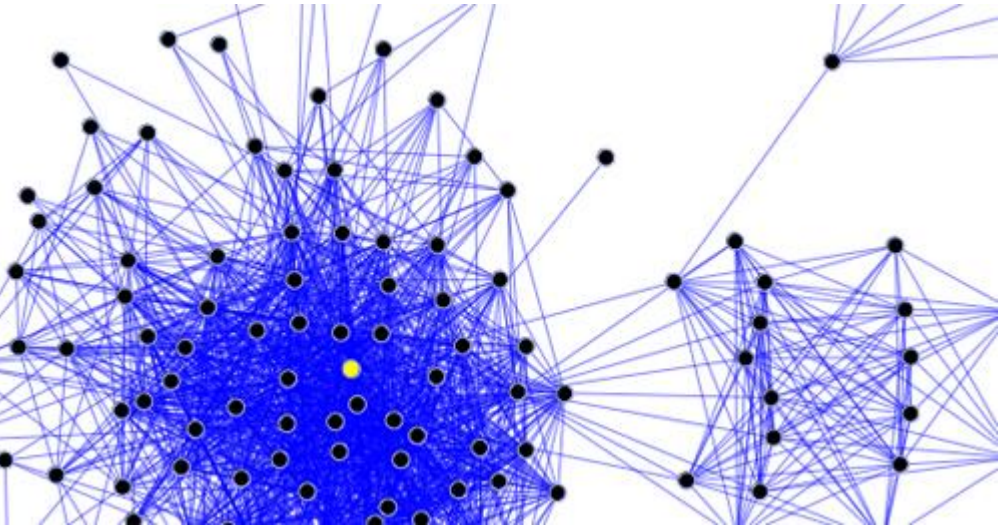
Vertices: 

Edges: 

Triplets: 

Property graph

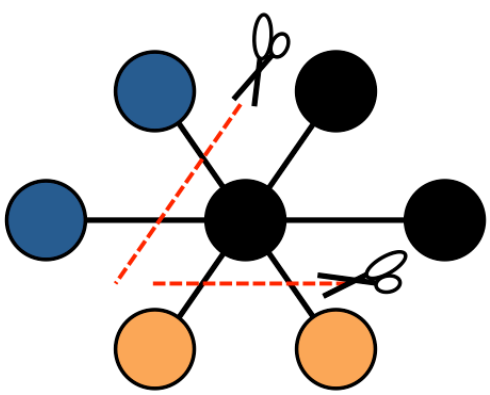




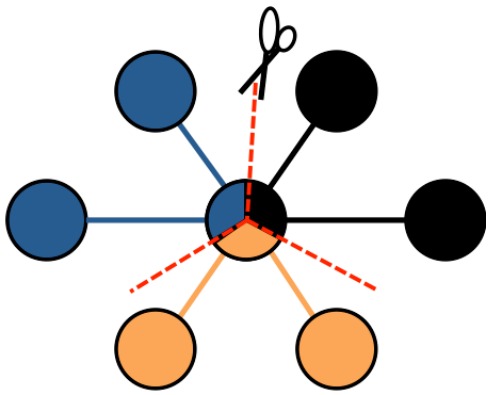
Key	value
12:00	53.5 mm
12:05	53.6 mm
12:10	53.6 mm

UNIVERSITY OF CAMBRIDGE

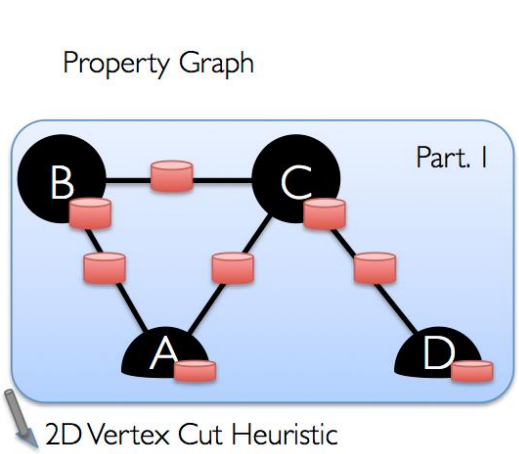
Optimisation

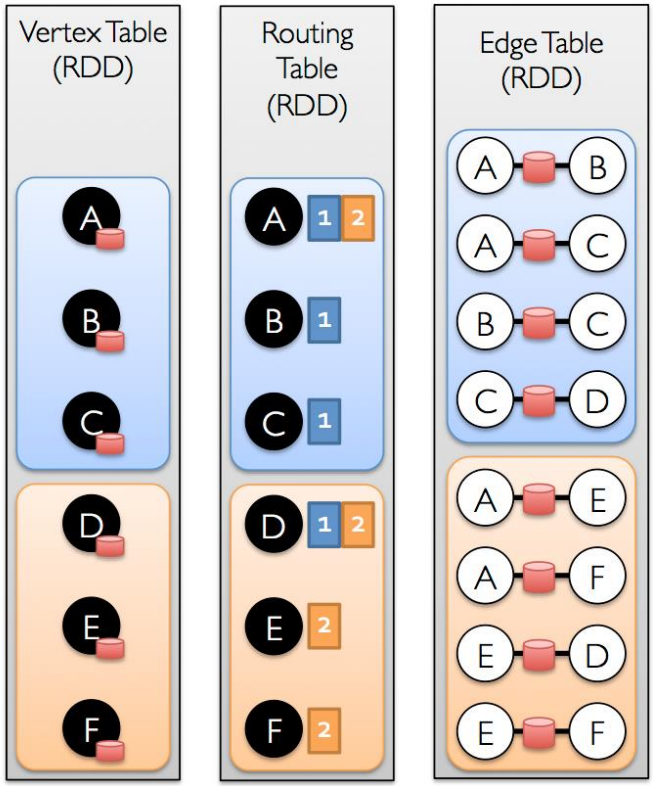


Edge Cut



Vertex Cut



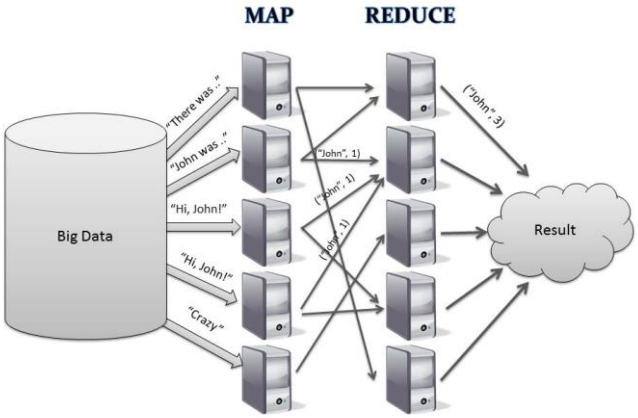


CSIC

UNIVERSITY OF CAMBRIDGE

Map Reduce (Batch Process) or In memory (real-time) computation?

- Map Reduce
  - Batch processing
  - Larger set of data to extract features and correlations



- In memory
  - Real-time predictions
  - Uses features and correlations from statistical analysis
  - Works on a smaller set of data and focused output.

Spark

Future direction

- Applying Machine Learning algorithms such as ‘deep-learning’ techniques to extract features and correlations.
- Real-time prediction of infrastructure behaviour based on statistical correlations and relationships.

“Transform the future of infrastructure through smarter information”

www.centreforsmartinfrastructure.com

Thank you!

CSIC Cambridge Centre for Smart Infrastructure and Construction

UNIVERSITY OF CAMBRIDGE