CDBB
Capability Framework and Research Landscape Scoping Workshop

Workshop: Capability to analyse and interpret data to create, and exploit the INFORMATION to understand, define, manage and realise the benefits of digital built Britain

Centre for Digital Built Britain

April 2018

UNIVERSITY OF CAMBRIDGE

cdbb
Centre for Digital Built Britain

This document captures the working notes from the workshop "Workshop: Capability to analyse and interpret data to create, and exploit the INFORMATION to understand, define, manage and realise the benefits of digital built Britain", held at Churchill College Cambridge on 10-11 April 2018

The summary sheets are assembled from the separate working groups from each of two streams; Research and Applications.

The details of the outputs from the individual working groups are captured in turn.

This material was used as a starting point for the creation and development of the Capability Framework and the Research Landscape. It is provided as source material for the interested reader.

Create, analyse and exploit information - Research Summary		
Rank order	Topic title	
1*	Visualizing, using information embedding context information into data	- Trends in messy data - Categorising
2*	Feedback loops (creating information)	- ML automated - Maximizing value
3**	Data integration and interoperability	- Different models, locations, very large scale, different organisations - Consistency
4	Quality of information	- Mis information, disinformation, fake news (security) - Provenance - Trust in AI (black box)
5*	How to search for / find information expertise	

Note:

* These are linked e.g. Info governance (competing and conflicting views)

** Linked

Research Topic: …	
Analyse and interpret data to create, and exploit the Information	

Scope:

Scope - In	Scope out	What sub-topics might overlap with other topics?
• Ways to make direct use of streams of raw data through ML • Maximize value extraction from data. Develop guidelines and tools to understand the process • Searching for and exploring DBB info • What are the social aspects of the information management in DBB? • Design intent tacit knowledge • Ontology • Linking data → context info → ontologies • Developing a lean information (knowledge) lifecycle • In-use data/information to inform design decisions • Information feedback to reduce non-conformance cost & lead time		• Attributes of BIM components. Cost/performance • Data lifecycle ↔ knowledge lifecycle

Step 2. Scope change by thinking about stakeholders

- Data: discover, search, browse
- Data structure. Structured - unstructured. Spectrum. Multiple ontology (scale)
- Info representation & visualization
- See Image 1
- User experience design (human computer interaction)
- Information governance - authority/validity. (Scales)
- Information ownership
- Misinformation. Disinformation.
- 1D - role. Classification of information based on their jobs

Region / Building, Road / etc in greater detail

Image 1

Step 3. Scope change by thinking about spatial differences

e.g. National/Regional	e.g. City/local	e.g. Asset specific
	• Issues; ownership, disclosure • Hard ↔ easy	

Step 4. Scope change by thinking about the lifecycle of assets and services

Articulate user needs and requirements	Conceive, plan and design (including optimisation and integration)	Build and commission (including optimisation and integration)	Manage and Operate (refine and enhance, optimise and integrate)	Provide valued services to users (and minimise downsides for non-users)	Retrofit / Renew / Decommission (with attention to the whole cycle)	…Assess, feedback and optimisation
		• Different info capture/created and used at each phase	• Information flow for enabling continuous learning and improvement			

Research Topic: …	
Analyse and interpret data to create, and exploit the Information	

Scope:

Scope - In			Scope out	What sub-topics might overlap with other topics?
• Incomplete data • Baseline performance - subsequent deviation • Centralised vs de-centralised data management • Data sharing; - what boundaries? - How open? - Which stakeholder? • Data context • Data "issues" sharing/interoperability • Processes across changing assets • Master data management • Legal/ethical issues • Data modelling plans vs reality • Public/private data/info boundaries • Public information requirements capture	• Analysis techniques: - Predictive capabilities - Machine learning - Statistical techniques • Dealing with differing/conflicting data quality requirements • Alternative integration architectures • Modelling change over time • Modelling time series • Mapping data between data models/ontologies • Alternative top level ontologies and evaluation of pros and cons	• Data presentation process; - How/what medium? - When? • Mapping information to usage. What framework elements do you need to support data sharing (Enterprise Architectures) • Data storage; - Who? / How? / Where? / Cost? • AI Black box • Data vs information • Information uncertainty • Qualitative information • Data de-trending - measured data - results of multiple trends/factors		

Step 2. Scope change by thinking about stakeholders

- Decision takers
- Those affected by decisions
- Those about whom info is used
- Multi-stakeholder consideration is collaborative decision making

Step 3. Scope change by thinking about spatial differences

e.g. National/Regional	e.g. City/local	e.g. Asset specific
• Degradation of data and provenance • What is a country??? defect? • Dealing with data overload. Keeping data small	• Data → ← summarisation • Disparate data sources - combining • What is a city defect? • What is a city optimisation? • Emergent & predictive behaviours • Information prioritisation at scale	• Movement of assets to different locations and contexts • Analyse to optimise performance and find cause of defects • Visualisation and interpretation to inform decision-making of asset owners

Step 4. Scope change by thinking about the lifecycle of assets and services

Articulate user needs and requirements	Conceive, plan and design (including optimisation and integration)	Build and commission (including optimisation and integration)	Manage and Operate (refine and enhance, optimise and integrate)	Provide valued services to users (and minimise downsides for non-users)	Retrofit / Renew / Decommission (with attention to the whole cycle)	…Assess, feedback and optimisation
• Methodology to facilitate discussion between key stakeholders to extract max value from monitoring system			• Obsolete sensors - how to keep data • Information for operation • Data migration		• Information for decommissioning (??? and design)	

Research Topic	
Analyse and interpret data to create, and exploit the Information	

Step 1. What are the major research clusters/themes?	What are capabilities and research that will be needed as DBB matures from 'deliver' to 'operate' to 'integrate'?					
	Deliver *(create the built asset)*		***Operate*** *(manage asset through life and deliver the services that derive from and depend on the asset)*		***Integrate*** *(deliver services and benefits based on integrated systems and organisations)*	
	What capabilities and enabling research?	***Which people / institutions are working on this?***	***What extra capabilities and enabling research?***	***Which people / institutions are working on this?***	***What extra capabilities and enabling research?***	***Which people / institutions are working on this?***
• Create info from data [or reuse]	• Generate valid information based on machine learning	• CSIS, Cambridge DIAL				
• Search and fund info	• Information retrieval • Text mining • Ontologies • Digital literacy	• Loughborough • CSEI Imperial College				
• Making sense and use info - Information governance	• Visualisation • Benchmarking • Collaboration and conflict resolution	• CSEI • CIBSE				

INFORMATION GOVERNANCE Links lifecycle

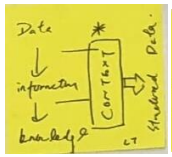
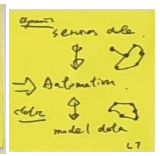
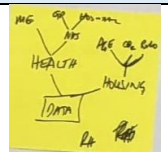
Research Topic	
Analyse and interpret data to create, and exploit the Information	

Step 1. What are the major research clusters/themes?	**What are capabilities and research that will be needed as DBB matures from 'deliver' to 'operate' to 'integrate'?**					
	Deliver *(create the built asset)*		***Operate*** *(manage asset through life and deliver the services that derive from and depend on the asset)*		***Integrate*** *(deliver services and benefits based on integrated systems and organisations)*	
	What capabilities and enabling research?	***Which people / institutions are working on this?***	***What extra capabilities and enabling research?***	***Which people / institutions are working on this?***	***What extra capabilities and enabling research?***	***Which people / institutions are working on this?***
• Data integration • Data contextualization	• Data modelling/ontology - change over time - time series - plan vs actual • Integration architectures		• Data integration at scale (e.g. towers with same tiles as Grenfell)			
• Communication of information inferred from data [exploring mediums] • Information visualization	• Information 'savvy' construction workers who can operate tools	• EHUD Reiter Aberdeen Uni natural language expertise	• Upskilling relevant workforce to use visualization software • Data reduction - extract key data		• Information aggregation • Translating 2D maps (e.g. OS maps) to 3D	
• Role of AI			• AI trust			
• Information quality			Dealing with multiple data quality requirements (different decisions) • Communicating information quality	• Philip Woodall, University of Cambridge		
• Decision making	• Collaborative decision making					

Create, analyse and exploit information - Application summary		
Rank order	Topic title	
1	Identifying value in information	- What data do you need@ ground level to support value - You might not know the different stakeholders value differently (value in use)
2	Understanding why different contexts are different	- Applying data science to understand behaviours - Local, regional and national feedback into design cycle
3	A House	- Cost challenge: Subsidies, mass market - Social drivers: age, technical comfort e.g. smart meters - ML , Big data, AI: current barrier of know-how
4	Feedback Loops - Data - information - knowledge - wisdom	- Making better assets in the future

Application Topic: …	
Analyse and interpret data to create, and exploit the Information	

Step 1. Scope: What topics should we include in this part of the framework – and what demonstrators would illustrate / stretch the boundaries?

Scope - In		Scope out	What sub-topics might overlap with other topics?
• Use machine learning or AI to automatically structure information in a coherent manner • Structured data (ontology & schemas) • Metadata • Consistent • Analysis of space usage against intended purpose • Increasing policy capacity • Data combinations; - weather - building temps - (& forecasts)	• Impact on hospital services    	• Actual collection of information	• Security & access rights & permissions

Step 2. Scope change by thinking about stakeholders (Are there new / different aspects of the topic and its demonstrators?)

- Provider
- Manager
- User
- Creating a focus on what information is required from the data
- Increase in resources
- Increase in environmental

Step 3. Scope change by thinking about spatial differences (e.g. to consider how can scale make a difference to the demonstrators we would propose)

e.g. National/Regional	e.g. City/local	e.g. Asset specific
• No change!	• Impact on local infra • Timeframe. Data requirements for future scenarios/modifications	• Increase in footfall

Step 4. Scope change by thinking about the lifecycle of assets and services: Are there new / different aspects of the topic and its demonstrators if we think through the lifecycle of the assets and the services?

Articulate user needs and requirements	Conceive, plan and design (including optimisation and integration)	Build and commission (including optimisation and integration)	Manage and Operate (refine and enhance, optimise and integrate)	Provide valued services to users (and minimise downsides for non-users)	Retrofit / Renew / Decommission (with attention to the whole cycle)	…Assess, feedback and optimisation
	• Levels of serviceability • How to link interpretation of data to desired outcome					• Social benefits

Application Topic: …	
Analyse and interpret data to create, and exploit the Information	

Scope: What topics should we include in this part of the framework – and what demonstrators would illustrate / stretch the boundaries?

Scope - In		Scope out	What sub-topics might overlap with other topics?
• Common language and ontology for data/info reqs especially in operational timeframe • User centred design - a useful tool (1 to scope data requirements?) • Aggregation and operation of data to create information • Big data • CAFM • Define needs and outcomes required 'in life' which need data/info to inform • How will information be used? Design, construct, operate, maintain • Simplification of data choices to support "complex decision" - tooling and processes • Machine as client for data and info • Interpretation of information mobility • Skills gap: data understanding is not wide enough! • Comparing data sets "apples with apples"	• Statement: "Info, like data is not neutral" • Data science definitions and skills; "its an 'art' not a science ..." • Data accessibility framework between parties: - data protection / commercial / security/trust / aggregation / retention • Data + process = information • Information + cognition = knowledge • Boundary later between information and knowledge • Big data small information big insight • Data info predicting and influencing behaviours which impact built environment	• Services layers - tools to allow access and filtering of data • Automated information mobility	• Trends and drivers • Systems • Everything! • Social

Scope change by thinking about stakeholders (Are there new / different aspects of the topic and its demonstrators?)

• Active contributors vs passive contributors eg Open street map or Waze • Generators of data. People. Sensors • Consideration of OIR & individual requirements • Social media (automated) input and output of information • Service providers; - telecoms - utilities - transport • Regulatory use of data and information	• Consumers of data. People. Machines. Applications • Eg: - Estate manager - Energy manager - Sustainability manager - Security - IT - Fire

Scope change by thinking about spatial differences (e.g. to consider how can scale make a difference to the demonstrators we would propose)

e.g. National/Regional	e.g. City/Local/Rural/Coastal	e.g. Asset specific	
• Aggregation, Anonymising, Pseudo-anonymisation (data protection) (incentives) • Moore's Law: Computing/Storage/ Processing capability		• Language for structured, open data static and dynamic • Static: - Power use - Manufacturer - Location - Sub component - Install date - Warranty duration	• Dynamic - Temperature - Power (in use) - Vibration - Lux • Design reviews; - defined "snap shot" views of data - legally identifiable point in demand

Step 4. Scope change by thinking about the lifecycle of assets and services: Are there new / different aspects of the topic and its demonstrators if we think through the lifecycle of the assets and the services?

Articulate user needs and requirements	Conceive, plan and design (including optimisation and integration)	Build and commission (including optimisation and integration)	Manage and Operate (refine and enhance, optimise and integrate)	Provide valued services to users (and minimise downsides for non-users)	Retrofit / Renew / Decommission (with attention to the whole cycle)	…Assess, feedback and optimisation
• Define data security • Business needs define data requirements but client is not well equipped to define	• Design sensors in & demonstrate value	• Commission sensors to give good outputs • Validate static data	• Perform validation in live data • Store and share data to give information and insight	• Combining datasets • What? Energy Consumption • What? Usage profile • Why? Behaviours and incentives	• Replace system when the need comes	• Turn knowledge into wisdom • Align with new and upcoming regulations • Case-based reasoning

Application Topic	
Analyse and interpret data to create, and exploit the Information	

Step 1. What are major demonstrators that are required?	**What capabilities / functionalities of the demonstrators illustrate the maturing of DBB from 'deliver' to 'operate' to 'integrate'?**					
	Deliver *(create the built asset)*		***Operate*** *(manage asset through life and deliver the services that derive from and depend on the asset)*		***Integrate*** *(deliver services and benefits based on integrated systems and organisations)*	
	What would be the big challenges?	***How?***	***What would be the big challenges?***	***How?***	***What would be the big challenges?***	***How?***
Small scale demonstration project, e.g. A house	• Cost/drivers for change • Energy consumption	• Government subsidies / Tax relief • Incentive take up • Apps/technology/smart stuff	• Age • Familiarity • Intergenerational skill levels	• Smart meters	• Appropriateness of machine learning/AI output	• Smart grid enabled
•Large scale: National roads network						

Application Topic	
Analyse and interpret data to create, and exploit the Information	

Step 1. What are major demonstrators that are required?	**What capabilities / functionalities of the demonstrators illustrate the maturing of DBB from 'deliver' to 'operate' to 'integrate'?**					
	Deliver *(create the built asset)*		***Operate*** *(manage asset through life and deliver the services that derive from and depend on the asset)*		***Integrate*** *(deliver services and benefits based on integrated systems and organisations)*	
	What would be the big challenges?	***How?***	***What would be the big challenges?***	***How?***	***What would be the big challenges?***	***How?***
1. Prove data has a value … 2. Define value! How? Tools? 3. Test value in real case, baselined against "numeric-data"	• Understanding the up-stream information needs? "What do I need to do to provide value to … the contractor/client"	• BEP should capture all supply chain info needs	• Reduction of operational & maintenance cost improving environment space quality/energy efficiency • Capture long-term operator info needs noting changes in use of asset over time (i.e. need flexible data demands - extensible data models)	• Quality assurance & accuracy tolerance identification • Comparison between similar building types; where 1 has good data, 2 does not.	• Organisations sharing use/ maintenance/ performance data to provide larger data set and enable comparisons • Defining data requirement to allow useful "network/system" view - what data is needed @ asset level to make good 'system' decisions?	• e.g. CIBSE/UCL online live energy benchmarks • User centered design tooling to define "needs" to then define inputs
• Data & data science. To understand and influence behaviours which impact built environment, e.g. energy usage for a complex campus	• What data? • Data format and quality • Data access • Systems/sources • Validation of information recorded. Captured in what format	• Agree on a standardised graphical and non-graphical format	• Data science/algorithms - modelling - behaviours analysis - contextual datasets		• Who/how can it be used on a bigger scale? Sharing across communities. Aggregate value of data	
• Feedback loop in operation. Generates wisdom • Using OSA-CBM as an example for demonstration • Using the online scan of utilities as a feedback to utilities owner (demo)	• Understanding value of specifying sensor infrastructure	• Demonstrator 1	• Sensors infrastructure & sharing systems		• Exposing live performance data	• Validation processes • Data 'authorities' or custodians