PROJECT ICEBERG AUGUST 2017

A UNIFIED DATA - FRAMEWORK FOR MAPPING UNDERGROUND

Project Iceberg: Work package 2 - Defining the problem space for an integrated data operating system above and below ground.

CATAPULT Future Cities

Ordnance Survey

British Geological Survey
NATURAL ENVIRONMENT RESEARCH COUNCIL

Future Cities Catapult (2017) A Unified data framework for mapping underground. Project Iceberg: Work package 2 - Defining the problem space for an integrated data operating system above and below ground. Ordnance Survey: LIF 6996.1/E. British Geological Survey OR/17/053.

Nicole Frith Ordnance Survey
Sasha Catchpole Ordnance Survey
Rollo Home Ordnance Survey
Carl Watson British Geological Survey
Holger Kessler British Geological Survey

Additional contributions from the Project Iceberg team, specifically, Andy Clark and Izabela Hurst from the OS, Stephanie Bricker and Helen Bonsor from the BGS, and Rimal Likhari and Stefan Webb from the Future Cities Catapult are gratefully acknowledged. The project team express their gratitude to the industry and research experts who participated in the stakeholder survey, and offered insights on current issues for subsurface data management and expert advice on existing barriers and future opportunities. This report forms an output from NERC knowledge exchange fellowship (NE/M020673/1) held by S. Bricker undertaken in partnership with Innovate UK's Future Cities Catapult.

EXECUTIVE SUMMARY

Project Iceberg is an exploratory project undertaken by Future Cities Catapult, British Geological Survey (BGS) and Ordnance Survey (OS). The project aims to address the serious issue of the lack of information about the ground beneath our cities and the un-coordinated way in which the subsurface space is managed.

Difficulties relating to data capture and sharing of information about subsurface features are well understood by some sectors and have been explored in previous research and industry reports. This study does not replicate past work, but rather reviews outcomes and explores the barriers to wider uptake of subsurface management systems.

The long-term goal is to help increase the viability of land for development and de-risk future investment through better management of subsurface data and ground-related risk. To help achieve this, our study aims to enable a means to discover and access relevant data about the ground's physical condition and assets housed within it, in a way that is suitable for modern data driven decision making processes.

The project considers both physical infrastructure i.e. underground utilities and natural ground conditions i.e. geological data and is divided into three different work packages:

- **Work Package 1:** Market research and analysis
- **Work Package 2:** Data operation systems and interoperability for a subsurface data platform
- **Work Package 3:** Identification of use cases for a subsurface data platform

This report summarises the findings of work package 2.

CONTENTS

CONTENTS

1. INTRODUCTION

"THE NEXT FRONTIER, IN BOTH A LITERAL AND FIGURATIVE SENSE, IS UNDERGROUND… INTERFACING MULTIPLE MAP LAYERS FROM DIFFERENT SOURCES TO COME UP WITH VALUABLE INTELLIGENCE"
Alan Leidner, Director of the Fund for the City of New York's Center of Geospatial Innovation

Project Iceberg is an exploratory project being undertaken by Future Cities Catapult, British Geological Survey (BGS) and Ordnance Survey (OS), with objectives to paint a picture of the subsurface – what is there, who holds data about it, who accesses it and how a single digital subsurface platform (that is BIM-ready) could drive radical efficiencies in workflow.

The project aims to build a holistic picture and market analysis of the current way in which the subsurface and its data is currently accessed and to outline the technical, legal and financial features of a single digital platform that links surface and subsurface data. The project aims to make a robust case for transformational change, providing stakeholders with an early indication of the 'preferred way forward' (not the preferred option).

The subsurface is an incredibly complex environment upon which the society places an increasing set of needs, such as holding significant utility assets, infrastructure assets and buildings. We are also increasingly reliant on the ground for its environmental functions, for example, flood control, waste storage and extraction of natural resources. The difficulties relating to capturing and sharing data about subsurface features are well understood, having been explored in projects such as the National Underground Assets Group (NUAG), EPSRC's Mapping the Underworld and the ASK (Accessing Subsurface Knowledge) network. Links to these, and other relevant projects are provided in Appendix D.

Mounting pressures of affordable housing, infrastructure management and environment protection place significant pressure on the finite land resource in the UK. Late stage awareness of ground properties and physical constraints to planned development is costly – ground risks are one of main causes of project delay, and of insurance claims on completed projects (Chapman, 2008). Meanwhile, according to TfL, road works account for 38% of the most serious and severe traffic disruptions across London at a total cost of £752 million (TfL, 2010).

Our long-term goal is to help increase the viability of land for development and de-risk investment through better management of subsurface data and ground-related risks. To help achieve this, our study aims to enable a means to discover and access relevant data about the ground's physical condition and assets housed within the ground, in a way that is suitable for modern, data driven decision making processes.

PROJECT SCOPE & ACTIVITIES

This study does not try to replicate past work, but to review outcomes and explore the barriers to benefits being disseminated more widely. Given the multi-disciplinary nature of land and asset management, this review covers a spectrum of sectors as there are shared aims across these sectors that inform the scope of the study. These aims are:

- Optimisation of asset performance, maintenance and resilience
- Effective planning and utilisation of subsurface space to support multiple functions, and
- Regulatory oversight via a shared single version of the data (giving improved transparency, accountability and governance)

Furthermore, the scope of this project is not limited to utility subsurface infrastructure but also subsurface geological data obtained from ground investigations.

The project has been carried out in three different work packages:

Work Package 1 (WP1): Market research and analysis through extensive desktop research, online survey of sector experts, followed by interviews with selected experts. The work package was led by the Future Cities Catapult and aimed to:

- Understanding the current state of play in the UK
- Reviewing previous projects relevant to Iceberg
- Assessing international project case studies with similar objectives as that of Iceberg

Work Package 2 (WP2): Building on the deliverables of WP1, this work package, led by OS, aimed to:

- Evaluate the level of interoperability of the data standards and operating systems. Identify barriers (for example technological, data-IPR and economic) to implementation.
- Outline the technical, financial, security and legal parameters of a single subsurface data model which allows subsurface data (2D and 3D) to be archived, released, and visualised in a manner consistent with surface related data and that meets data standard requirements, software compatibility and organisational requirements.
- Focus on the framework and protocols for sharing data, the service layers and the security that needs to underpin all of this.

Work Package 3 (WP3): Identified potential use case applications of a subsurface data platform.

It should be noted that developing and building a surface/subsurface data sharing framework is beyond the scope of Project Iceberg. WP2 aims to learn from the lessons of relevant previous projects, and in this way, provide insights into how progress against some of the persistent barriers may be overcome going forward in future projects. It is hoped that the themes discussed in this report will help shape the outcomes of Work Package 3 (WP3), which aims to map current and future user journeys.

While the development of a functional surface/subsurface data sharing framework is out of scope it is possible to identify the high-level stages of data flow, that would occur within such an exchange. Once these components are identified, it is possible to foresee the likely barriers that would need addressing going forwards to the proof-of-concept stage of any future projects. These barriers cover a range of themes, e.g. technology, standards, liability, authority, security, and IP. In this report, each of these themes is discussed with reference to previous projects relating to subsurface assets. Each of these barriers or 'problem spaces' has first been defined and justified within the context of Project Iceberg. Following this is a short discussion into existing research and work relevant to each problem space, and possible recommendations and considerations going forward

PROJECT TEAM

Future Cities Catapult: Future Cities Catapult is the government's urban innovation agency, with a mission to advance innovation, to grow UK companies, to make cities better. For this project, we leveraged our Strategy, Markets & Standards (SMS) and Creative Design Services (CDS) teams to paint a picture of the sub surface and assess the current state of play, in the UK and globally.

Ordnance Survey (OS): Great Britain's national mapping agency. It carries out the official surveying of GB, providing the most accurate and up-to-date digital geographic data, relied on by government, business and individuals.

British Geological Survey (BGS): UK's provider of geoscientific data, information and knowledge. And custodian of the UK's national geological data archives. The BGS develop technology for the digital transfer of subsurface geological data (e.g. to BIM) and 3D geological modelling systems.

HOW TO USE THIS REPORT

Section 2 (Conceptual model) of this report describes the data life cycle for a data framework that integrates above and below ground datasets. The data life cycle diagram highlights the key components and key factors to consider. The diagram is interactive with hyperlinks and readers can link to key sections of the report by clicking on each item of the diagram. A description of the data life cycle follows the interactive diagram, these written sections again highlight key factors that need to be considered. The icons for the key factors are also interactive and allow readers to hyperlink to relevant sections of the report.

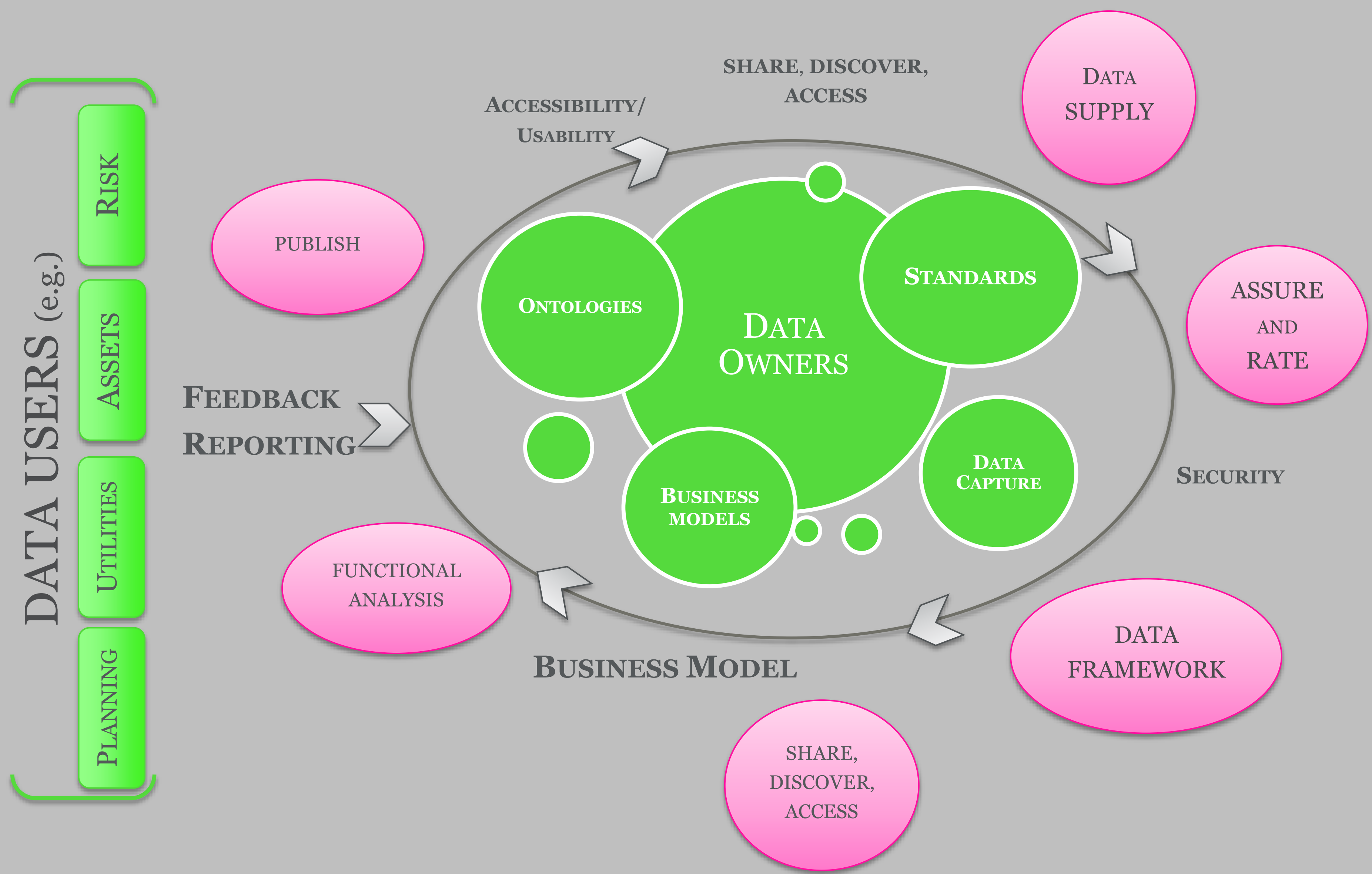
Section 3 (Key Factors) and Section 4 (General Problem Spaces) contain the detailed descriptions that underpin Section 2 (Conceptual model) and the interactive diagram.

2. CONCEPTUAL MODEL

The diagram below represents the conceptual user journey through a data discovery, utilisation and updating/report work-flow (the data life-cycle). The high-level stages of data flow, which would occur within such an exchange can be classed as:

- **Data Supply**: Existing data must be fed into the framework. This will be complemented by future additional data capture.
- **Assure/Rate**: Uploaded data will need to be in some way assured and rated (or graded), to provide a measure of its reliability or suitability for different applications.
- **Data Store/Framework**: A “framework” must be developed to facilitate data sharing.
- **Share/Discover/Access:** Multiple users will need to access this “framework” to view and access subsurface data.
- **Functional Analysis/Analyse**: Users will need to derive value from shared subsurface data.
- **Publish/Action**: Based on shared data, users will need to take some form of action, e.g. decision-making, publishing, additional sharing.

Figure 1 Data life-cycle. [Hyper-link to report section].



2.1 DATA SUPPLY

KEY FACTORS TO CONSIDER:

(CTRL+CLICK ON KEY FACTOR TO LEARN MORE)

Data Standards

Data supply is the input of existing data to the 'framework' from those stakeholders with relevant information about the surface and subsurface. These stakeholders will come from various industries, government and quasi-government bodies to create a fully integrated view of both the surface and subsurface. Stakeholders may range from organisations such as BGS to Utilities to Local Authorities to Developers to Ordnance Survey. As the "framework" begins to demonstrate how value can be derived, the number of stakeholders will continue to increase.

Data models that can deal with uncertainty will be required. In addition, there will need to be an ability to indicate situations where no information is available, differentiated from situations where features from that stakeholder do not exist (i.e. avoiding false negatives). The ability to distinguish between observed data, modelled data and interpreted data is needed.

For a fully integrated solution all relevant data will need to be included, this can include additional elements outside the traditional scope of utilities and streetworks. This could include:

- Water infrastructure
- Sewer infrastructure
- Electricity infrastructure
- Gas infrastructure
- Telecommunications infrastructure
- Transport infrastructure
- Soils, surface and other underground features (for example physical properties and structure as well as observations relating to dynamic processes such as fluid flows through the subsurface)
- Surface nature (for example street, sidewalk, buildings and open space surface characteristics)
- Hydrography and bathymetry
- Surface elevation
- Water table
- Buried structures e.g. Foundations, basements, cellars, vaults, passage ways
- Sensor derived near real time observations

Back to top

2.2 DATA CAPTURE

KEY FACTORS TO CONSIDER:

(CTRL+CLICK ON KEY FACTOR TO LEARN MORE)

Data Standards

Capture Technology

3D (Depth or elevation)

The processes of collecting new data about relevant surface and subsurface features. This may be initiated by different events:

- Installation of new infrastructure, captured to agreed standards
- Initiatives to capture new data about existing infrastructure, perhaps in advance of planned works or new development or as a result of feedback from the framework users
- Capture of real changes to related information such as changes to ground cover or hydrography
- Improvements to data capture techniques employed resulting in higher quality or more refined data being available, for example an improved terrain model or improved knowledge of geology.

As with DATA SUPPLY the input of data to the 'framework' is from those stakeholders with relevant information about the surface and subsurface. These stakeholders will come from various industries, government and quasi-government bodies to create a fully integrated view of both surface and subsurface. Stakeholders may range from organisations such as BGS to Utilities to Local Authorities to Developers to Ordnance Survey. As the "framework" begins to demonstrate how value can be derived, the number of stakeholders will continue to increase.

For a fully integrated solution all relevant data will need to be included, this can include additional elements outside the traditional scope of utilities and streetworks. A list of example datasets can be found within section 2.1 Data Supply.

Back to top

2.3 ASSURE AND RATE

KEY FACTORS TO CONSIDER:

(CTRL+CLICK ON KEY FACTOR TO LEARN MORE)

Data Standards

Data Formats

Data Liability

Assurance

The process by which supplied data is assessed and validated to provide an indicative measure of its reliability and therefore how it can be used. ISO 19157 provides a set of measures (not values) that ensure that users can make valid decisions about how to use the data.

It is likely that much of the data available initially will be of unknown or 'low' quality. Pragmatically, there may be benefit even in sharing low quality data, in particular if the same data is already shared by less efficient methods.

This process is essential to provide confidence for users of the framework about how to use the data.

Data should be provided with its' provenance attributed explicitly. Data without known provenance needs to be identified as such.

Data Ownership

Within this activity, it is unlikely that data can be tested cost effectively but the concept of 'trusted suppliers' could be developed.

Auto-assurance of sensor data may also be considered, where data formatting, control measures and metadata are pre-prescribed.

Back to top

2.4 DATA FRAMEWORK

This refers to the framework or structure that provides a means of data sharing. There will be a requirement to manage large volumes of data from disparate sources to meet user needs.

Ingestion processes, timetabling, logical validation and feedback loops will need to be established. The framework will need to be able to supply multiple *ad hoc* requests in real time.

KEY FACTORS TO CONSIDER:

(CTRL+CLICK ON KEY FACTOR TO LEARN MORE)

- Data Ownership
- Hosting vs. Federation
- Intellectual property
- Security

Back to top

2.5 SHARE, DISCOVER, ACCESS

The process of multiple stakeholders accessing shared relevant data.

There will be a requirement to view and interrogate in real time for multiple users. In addition, feedback to the data originator, the data managers and potentially other users is required where issues are identified.

To recognise the maximum business benefit, these processes are likely to ultimately be required in 'real time' and be capable of being served to mobile devices.

KEY FACTORS TO CONSIDER:

(CTRL+CLICK ON KEY FACTOR TO LEARN MORE)

- Personal Data
- Accessibility/ Usability
- Data Standards
- Security

Back to top

2.6 FUNCTIONAL ANALYSIS

By accessing the data, the stakeholders will be expecting to derive business value, for example by identifying risk or more effective planning or work. At one end of the scale standardised queries could be run and responses generated – for example "there is no utility infrastructure in this area' or a geological report. At the other end of the scale the framework may supply data in a format for analysis 'offline' where the data owner agrees to this.

If charging is involved then this is likely to fall under PUBLISH/ACTION

KEY FACTORS TO CONSIDER:

(CTRL+CLICK ON KEY FACTOR TO LEARN MORE)

- What's Needed?
- Alert
- Feedback Reporting

Back to top

2.7 PUBLISH

This would be the publication of the data to external parties who may be expected to pay for a service. For example, a property being sold or a parcel of land being explored for development for reporting purposes.

KEY FACTORS TO CONSIDER:

(CTRL+CLICK ON KEY FACTOR TO LEARN MORE)

- Security
- Intellectual property

Back to top

> At LSBUD, we work with 60+ asset owners and respond to over 1.7 million requests for utility information per annum. Each asset owner has its own policy for sharing information as well as having the data in different formats.

LSBUD (LineSearchBeforeYouDig)

3. KEY FACTORS

3.1 CAPTURE TECHNOLOGY

***Definition*:** The process by which subsurface and surface data is created to known standards. This includes data that is created to record existing real world features.

Why it's important to consider: Data included within the data sharing environment will likely be derived via four main processes:

(1) Previously existing data will be uploaded to the sharing environment (see DATA SUPPLY).

(2) Authoritative data suppliers will continue to create key reference data sets (e.g. OS, BGS, asset owners e.g. Utilities, Transport Authorities, Environmental Regulators and contractors).

(3) Additional metadata, specifically required for the successful functioning of the framework, may need capturing or recording for example provenance, valid dates or similar. Capture of spatial and temporal metadata may also be used to expose un-processed data in such a way that enables smarter prioritisation of un-processed data for ingestion into the data framework.

(4) Data derived from the framework will become an additional data source via a feedback mechanism. This will include corrective updates as well as derived information.

As the framework develops and technology improves, there will be an increasing need to collect accurate data that can be more easily used and shared amongst stakeholders. In the shorter-term however, it is recognised that existing data is insufficient to provide a comprehensive view of the subsurface (EPSRC's *Mapping the Underworld and Assessing the Underworld*). Moreover we probably never will achieve a 'comprehensive view of the subsurface' – there will always be some level of uncertainty and interpretation involved. The availability and development of sufficient capture technology will be a critical enabler for this process.

Just 52% of the Iceberg stakeholder group (those that participated in the WP1 Survey) record the absolute location of subsurface assets referenced to a recognised geodetic datum (for example OSGB36 or ETRS89). The stakeholder group currently use a wide range of methods to collect subsurface asset data (ground-penetrating radar surveys [35%], radio-frequency location surveys [17%], ground probing radar surveys [17%], acoustic mapping surveys [9%] other methods e.g. tunnel surveys, intrusive ground investigation, surface level indicators [78%]). Full insights from the stakeholder survey are provided in Appendix C.

Relevant existing work: EPSRC's *Mapping the Underworld* is a project which aimed to address the lack of existence of any single sufficient data capture technology. They have undertaken critical enabling research and developed integrated geophysical tools into a single measurement vehicle, which can locate (in 3D) and record the position of all buried utility assets without excavation. The research findings provide a strong evidence base for a commercially-developed multi-sensor device and offer a proof-of-concept for a means to overcome this technological barrier.

Future recommendations: It is promising that the tools exist to overcome this technological barrier, however currently the approach taken by the surveyed stakeholders is disjointed. Research findings from the EPSRC's *Mapping the Underworld* project offer a strong evidence base for commercial development of a single, integrated tool. Continued development of such technologies, and as wide an uptake as possible of accurate locating and recording technologies, should be encouraged to continually develop the subsurface data that is captured and shared within any future framework. The benefits of developing such tools and technologies will likely be realised over longer timescales, and will be increasingly realised as any future framework gains traction within the stakeholder group.

In using such tools and technologies to create subsurface data where none currently exist, issues of data ownership may arise, and must be considered going forward.

Back to top

3.2 DATA STANDARDS

Definition:

Recognised and followed standards are crucial as a way of ensuring multiple stakeholders create, collect and maintain data in a consistent way so that data is fit for a defined purpose.

Similarly, the supply of data out of the system needs to be standardised where possible to minimise the costs of creation and validation and avoid the inadvertent loss of valuable information.

Standards will be relevant to format and structure as well as the quality metadata, all of which are required to facilitate effective sharing.

Why it is important to consider: By providing an accessible framework and standard for stakeholders to engage with, the data should meet the required minimum standard for data sharing within the context of this "solution/framework". If data is always being created to this standard, then all data should be able to be assured and ranked highly, providing more trust within data sharing practices. The Utility Strike Avoidance Group's (2014) report delivered the key message that asset owners need to improve the quality, accuracy and access to plans. However without a recognised standard, this process will vary between different stakeholders.

Whilst all data to a 'high' standard may be a laudable aspiration pragmatically any system will need to accommodate data of varying quality. Without the supply of data quality information then sharing will not recognise the full business benefits as users will have to assume all data is of the lowest quality.

The key related issues here are **Data Liability, Legal/IP, Classification/Format**

Relevant existing work: Recently, two sets of Publicly Available Standards (PAS) have been produced aimed at specifying standards for underground data capture. These are;

- *PAS128: 2014 Specification for Underground Utility Detection, Verification and Location*
- *PAS256: 2016 Buried Services – Collection, Recording and Sharing of Location Information Data – Code of Practice*
- *OGC Underground Infrastructure CDS - The OGC Underground Infrastructure Mapping and Modeling project will lead to improved public safety, project delivery and urban resilience from a secure 3D repository of urban underground infrastructure.*

- *KLIP & IMKL: Using INSPIRE to develop a regional utilities database in Flanders - https://portal.opengeospatial.org/files/?artifact_id=73811*

Those involved within the creation of the PAS standards include the British Standards Institute and the Institution of Civil Engineers, as well as input on the steering committee from OS, the University of Birmingham, NUAG, NJUG, TfL and variety of other organisations. Outcomes of the standards can be summarised as “Encouraging three-dimensional absolute locating of assets referenced to a national grid and datum”. PAS 128 does not cover any non-utility buried infrastructure or make recommendations for the data format, but is more focused on describing four levels of accuracy during a survey of underground utilities to a depth of 3m, it is likely that for some use cases data at greater depths than 3m will be required. PAS256 aims to outline a consistent, accessible data protocol to enable effective recording and sharing of the location, state, and nature of buried assets, and recommends how existing asset records should be updated, recorded and collated. The standard does not cover recommendations on how or where the data is stored, or how data integrity is assured. Other future stakeholders may include the Open Geospatial Consortium (OGC), whose current standards address the data architecture, (format/structure etc.) whilst PAS128/256 address processes whereby data is captured. OGC’s work to provide standards that facilitate interoperability may be crucial in designing a standard that addresses all the components of capture processes, data storage, sharing and integrity. Metje et al.’s study (2015) states that 84% of utility strikes were due to the plotting of an inaccurate location. *PAS256: 2016 Buried Services – Collection, Recording and Sharing of Location Information Data – Code of Practice* recommends that data should be available for sharing as soon as is reasonably possible, preferably within 30 days. A faster turnaround will be required if a truly dynamic system is desired.

Future recommendations: An industry agreed standard that builds on PAS128/256 is required, that ensures the standard is accessible to various stakeholders and addresses all aspects of data capture, storage and sharing practices. Wider involvement in OGC initiatives will allow a broader range of experiences to develop the standards. In particular, the Project Iceberg team should contribute to a real use case that informs the development of a pilot project in collaboration with the OGC Concept Development Study (http://www.opengeospatial.org/standards/requests) into underground mapping. Regulatory change may have to be enforced to ensure wide spread adoption of such standards.

Back to top

3.3 3D (DEPTH OR ELEVATION)

Definition: The positional element that describes the location of the asset with respect to a vertical datum and potentially the ground surface – either explicitly of by inference.

The third dimension at its’ simplest is simply an element of the data geometry specification. However, given that much existing data is only capture in two dimensions or possibly 2 ½ dimensions (X, Y and depth) then it has been drawn out specifically in this section.

Why it's important to consider: 2D is a not an accurate representation of reality (EPSRC's *Mapping the Underworld*). There are no cross-industry recognised enforced regulations that state that the third dimension (depth or elevation) must be captured. Within the context of a future "framework", for some stakeholders to derive value from sharing data, depth must be known. For example, as a stakeholder concerned with planning best use of the underground, being able to view the entire current situation is critical to make informed decisions related to urban development. Or, as another example, as a utility company planning a series of asset maintenance works near surface, knowing at what depth other companies' assets are located may affect the estimated length of works or the tools and procedures used. The relatively recent utilisation of GPS to create accurately located data also increased the range of inconsistency between different data owners, resulting in greater discrepancy (Beck et al., 2007).

> ***"Any 3D system requires the asset to be represented as a solid object. Minimum capture standard is a 2D line (xy) with a z attribute per segment or node plus an overall attribute of diameter."*** ***(Holger Kessler, BGS).***

The Government's *Digital Built Britain* agenda, combined with new mandated use of BIM level 2 on all public-sector projects by 2016, places greater importance on the need for 3D data.

It must also be considered how to address a lack of third dimension data in relating to previous records and how any future feedback loop can help to improve data.

Relevant existing work: *BIM for the subsurface*, an ongoing project, funded by Innovate UK and undertaken by Keynetix, Atkins, Autodesk and BGS, is currently looking at the integration of geological conditions into BIM models. In order to achieve a truly integrated view of the surface and subsurface, data must show its variations in depth, not just x and y. This need for an integrated 3D above and below ground model is also being explored through a NERC Fellowship, *'Integrating subsurface environmental data & knowledge into city planning'*. The fellowship aims to support a new volumetric approach to planning (planning, land use, delivery of infrastructure, housing, and asset management), in conjunction with Glasgow City Council.

Future recommendations: Capturing the third dimension in data will become increasingly crucial in developing an integrated and realistic view of the surface and subsurface. A future "solution/framework" must be able to demonstrate to data suppliers the potential value that can be derived and considered within the Standards problem space. Within the feedback loop, it would be useful to consider how improving any 3-dimensional data could occur.

Back to top

3.4 DATA FORMATS

Definition: The types and structure of data that can be included in the framework via DATA SUPPLY and DATA CAPTURE. Formats for Data Capture should be derived from Standards, however for legacy data sub-optimal formats will need to be considered for example pdf, scanned images or digital formats without internal metadata.

Data format is also relevant to PUBLISH/ACTION and SHARE/DISCOVER/ACCESS to supply the output of the framework is formats that users can utilise.

Why it's important to consider: Deciding what data formats to accept within the framework depends on many factors. To be a fully integrated view of both above and below ground, it would be impractical to exclude any data provided as it will create gaps in the overall picture. If the "framework" cannot be trusted to be a representation of all the available data, it may not be used. However, the technological capabilities may not allow multiple data formats (intelligent e.g. vector vs. unintelligent e.g. paper drawings) to be easily combined and interoperable. There is certainly value in creating processes whereby data can be extracted from unintelligent disparate sources (automated/manual) to feed into the framework, however this will also impact on and need to be addressed in any cost analysis/future business model.

Relevant existing work: The challenge of combining heterogenous data with no common format or standard was one that posed a risk to the *VISTA* (Visualising integrated information on buried assets to reduce street works*)* project. Its aims were to:

- Pull together current records of pipes, cables and wires and link them to new surveys
- Create the first 3D maps of underground utilities in the UK

It was trialled within both the East Midlands and Scotland, in collaboration with the Scottish Road Works Register (SRWR), however a progress update cannot be located.

Other projects that encountered this problem include the *'National Asset Records Exchange'* (National Underground Assets Group), the Climate Change Commission's 'Assessment of the impact of climate change on infrastructure' and the *GLA Network Utilities*.

Future recommendations: Within the creation of a set of universal standards ranging across all industries, there needs to be a consideration of whether to suggest/enforce a certain format and whether it is accessible to all data supply stakeholders. This will also impact on technology and the future capabilities within the "framework".

Back to top

3.5 ASSURANCE

Definition: A way of identifying the quality of data supplied to the "framework" and to reject data that is not fit-for-purpose, from which no value can be derived.

Why it's important to consider: Without assurance that the data within the "framework" is accurate (or the degree of inaccuracy or uncertainty provided), can be relied upon, or trusted (to a certain degree), participation and utilisation of the resource may be limited. Historical mindsets that companies do not need to share data or see no value in it are already a threat to the development of an integrated view of both above and below the surface. Lack of data, or uncertainty that the data is correct, has been shown to be a barrier to companies sharing data previously, as outlined within the results of the survey. Subsurface data however is almost always subject to some level of uncertainty (even man made material and buried assets). Measurements are subject to interpretations based upon conceptual models, which can change over time. Subsurface data is almost always subject to some level of uncertainty (even man made material and buried assets). Measurements are subject to interpretations based upon conceptual models, which can change over time.

- **“ad hoc processes based on information reported from project archives. Too often key data is not recorded”**
- What information relating to your organisation’s assets would you be **reluctant to share** and why? **“Data accuracy level”**
- What other subsurface data would you **like to have access to? “Data accuracy level”**

By providing an indication to a potential user whether this data is accurate, it places the responsibility with the user whether they choose to access and how they use the data or not. It has been discussed previously how all data has the potential to be useful to any stakeholder, depending on how value is derived. Therefore, no data should be excluded from the sharing environment but rather an indication made based on an agreed set of values, relating to a set of parameters that are linked to metadata.

These metadata could include; accuracy, update frequency, methodology of data capture etc.

It has been suggested that it may be more appropriate to only indicate whether the data has been quality controlled/quality assured by the data supplier. However, if this is to be the solution to this problem space, an agreed set of standards or requirements of these processes would need to be agreed with the various industries to ensure that everyone is working to the same parameters. Also, knowledge that every company already undertakes these processes as part of their workflow – otherwise this is another task to add to this, and the value of this extra cost will need to be demonstrated to ensure participation.

Data that is identified as not meeting the required quality levels should be fed back at the earliest opportunity to the supplier.

Relevant existing work: As part of the publication of PAS128, a means by which to assess and indicate confidences that can be placed in the data was included. However, lack of industry wide implementation means it is hard to assess whether it was a meaningful addition to a data sharing environment.

Few previous projects or initiatives appear to have incorporated any mechanism of “cleansing” during planning for an integrated view of the subsurface. However incomplete historical data and different records of the truth, i.e. the implication that the data is not truly representative or inaccurate, has been perceived as a barrier to creating a data sharing environment or framework (*GLA Network Utilities; VISTA* - Visualising integrated information on buried assets to reduce street works).

Future recommendations: Once again industry-wide participation and consensus is key to solving this problem space and providing a means by which users can assess and make decisions on providers’ data.

Back to top

3.6 DATA OWNERSHIP

Definition**:** Maintaining data owners’ intellectual property rights and identifying legal implications and responsibilities associated with data sharing.

Why it's important to consider: The sharing of subsurface data will lead to issues of IP and legal rights, responsibilities, and liabilities. Of 23 survey respondents, 22% identified 'Intellectual Property Rights' as a specific factor preventing their organisation from participating in a subsurface data sharing framework. Addressing the issue of legal liability was also identified as a possible reason for not participating in a data sharing platform. In particular, issues may arise with respect to where legal liability lies when decisions are made based on incomplete or incorrect information.

Relevant existing work: The National Underground Assets Group (NUAG) have previously proposed a national web-based solution service, called the *National Asset Records Exchange*, which entailed that asset owners remain responsible for managing and securing their own asset information. A London-based project proposed by *GLA Network Utilities* also proposed that liability should not be transferred but rather remain with data owners. Similar issues have also been identified as a barrier to the adoption to BIM as a process in the construction industry (for example see Understanding and facilitating BIM adoption in the AEC industry, Ning Gu and Kerry London 2010). The mandating of Level 2 BIM (2016) has had some impact, but adoption rates still remain low indicating potentially the need for further action.

Future recommendations: Metadata may offer a partial solution to this barrier. In providing metadata, data users can establish to what extent they trust the data and can opt to undertake additional survey if required or necessary. Legal liability must either remain with the data owners/suppliers, or be transferred to the user at the point of use. Opting for the first of these may limit involvement in a subsurface data sharing platform, as data suppliers are unlikely to expose themselves to additional risk in this way. Opting for the latter is likely to limit use of the platform, as data users will be wary of making decisions based on possibly 'untrustworthy' data. Most the Survey respondents stated that they either share (65%) or sell (17%) their subsurface data with other organisations. Processes for establishing how risk is, or is not, transferred during these processes may already be established and may offer learnings for how such a barrier can be overcome in a more holistic and comprehensive subsurface data framework.

Back to top

3.7 INTELLECTUAL PROPERTY

***Definition*:** Defining the legal requirements e.g. licensing, responsibility etc., once data has been extracted and analysed by stakeholders other than the data users.

Why it's important to consider: There are costs associated with creating and maintaining data. The costs of using, and making decisions based upon, other organisation's data must be determined. Selling data is a revenue stream for many organisations (e.g. South West Water, Anglian Water, Thames Water, Southern Water). 17% of the Survey respondents, for example, sell their data to other organisations, and so use this as a revenue stream. This must be protected and replaced within any future subsurface framework, if stakeholders are to engage and supply data more widely.

Relevant existing work: In 2011/2012, the National Underground Assets Group (NUAG) started the 12-month London Trial Project, which involved a web-based national asset record information sharing service (NRS) to improve access to information on buried assets. This was expected to be followed by nationwide implementation, where costs are shared amongst users based on an agreed apportionment model.

Other examples of how organisations address licensing and IP can be found at:

- BGS: http://www.bgs.ac.uk/data/licensing/home.html
- OS: https://www.ordnancesurvey.co.uk/business-and-government/licensing/index.html

Future recommendations: Terms regulating the costs of data use will need to be included within the terms and conditions of participating in a subsurface data framework. These will need to be sufficient to replace revenue streams generated under the current business-as-usual selling procedures, but not too high, so as to price a new means of data sharing out of the market.

Back to top

3.8 DATA LIABILITY

Definition: The legal responsibility where decisions made based on the input data give rise to a liability.

Why it's important to consider: The sharing of subsurface data will lead to issues of IP and legal rights, responsibilities, and liabilities. Addressing the issue of legal liability was identified as a possible reason for not participating in a data sharing platform. In particular, issues may arise with respect to where legal liability lies when decisions are made based on incomplete or incorrect information.

Relevant existing work: The *GLA Network Utilities* project outlined that within this area, any solution designed should ensure that legal liability remains with the data owners. "System owners will be required to ensure transparency, highlighting that asset owners will keep liability" – however the wider implications of this decision may limit buy-in from stakeholders if they believe they are opening themselves up to limitless legal, financial and organisational-image risk.

Future recommendations: *These future recommendations are also outlined within the Legal/IP section of the Data Framework Component Summaries.* Legal liability must either remain with the data owners/suppliers, or be transferred to the user at the point of use. Opting for the first of these may limit involvement in a subsurface data sharing platform, as data suppliers are unlikely to expose themselves to additional risk in this way. Opting for the latter is likely to limit use of the platform, as data users will be wary of making decisions based on possibly 'untrustworthy' data. The majority of the Survey respondents stated that they either share (65%) or sell (17%) their subsurface data with other organisations. Processes for establishing how risk is, or is not, transferred during these processes may already be established and may offer learnings for how such a barrier can be overcome in a more holistic and comprehensive subsurface data framework.

A consistent and clearly defined legal framework must be created for both those stakeholders utilising and deriving value from data shared within the environment, and those supplying it.

Back to top

3.9 CONTROL OF DATA (AUTHORITY)

Definition: Stakeholders retaining autonomy over their own data.

Why it's important to consider: Once data enters a sharing environment, which can be accessed by a range of stakeholders, those supplying the data are at risk of losing autonomy over their own data. This may be mitigated through use of a federated data framework (see above), where data continues to be stored with the owner. At later stages of the data sharing process, issues of authority and autonomy may arise where stakeholders can derive additional data or value from data that was contributed to the framework by another stakeholder. Similarly, there may be questions and issues regarding data authority arising in the instance that those stakeholders using data contributed to the framework offer improvements and corrections.

Relevant existing work: The National Underground Assets Group (NUAG) have previously proposed a national web-based solution service, called the *National Asset Records Exchange*, which entailed that asset owners remain responsible for managing and securing their own asset information.

Future recommendations: One way of overcoming this problem may be to build the "framework" as a federated system, whereby data remains with the data owners.

Back to top

3.10 OWNERSHIP

Definition: Responsibility of maintaining and controlling the framework.

Why it's important to consider: There are several issues surrounding the theme of ownership:

(1) Given that no single data framework currently exists to facilitate a single view of the subsurface, there is, in the first instance a requirement that 'someone' (e.g. an organisation, government department, cross-industry committee) takes responsibility for creating and building a data sharing platform. Two projects, *Mapping the Underworl*d and *Assessing the Underworld* identify no single champion within England and Wales to take responsibility for creating such a system, which is therefore a barrier to implementation going forward. Once resolved, issues of IP and ownership may arise amongst those involved with creating and building the platform, and those contributing and supplying data to the platform.

(2) Once a data sharing platform exists, there will likely be a requirement for a custodian (e.g. organisation, government department, committee) responsible for maintaining the overall framework. An issue arising in this instance is to what extent the chosen custodian is responsible for data maintenance, and to what extent this responsibility lies with the data, or asset, owner. Again, issues of IP and ownership may arise amongst those maintaining the platform and those supplying and maintaining data.

Relevant existing work: To overcome these issues, previous projects have typically required some council/government involvement. The *ASK Network* project, Glasgow, demonstrates successful collaboration between over 20 industry partners and 12 public sector bodies in Scotland, however was primarily developed by just Glasgow City Council in collaboration with BGS. Glasgow City

Council are now responsible for ensuring reported data is compliant and reported according to a standard specification (Glasgow Specification for Data Capture, GSPEC). Similarly, in Malaysia, the Department of Survey and Mapping (JUPEM) developed a national underground utility database as part of its mandate.

Future recommendations: Examples of previous projects would suggest that to achieve the most successful outcome, there may be a requirement for a single, or limited number of companies, take ownership and responsibility for creating and maintaining a subsurface data framework. Responsibility for creating such a framework may need to be included as part of a government mandate, and may need to fall to a body or organisation that is, to some extent, impartial, to limit issues of stakeholder inequality. Given that this is envisioned as a national framework, there may also be a requirement that responsibility for creating and maintaining the framework itself falls to a stakeholder with a national coverage (e.g. OS, BGS).

Back to top

3.11 TECHNOLOGY

***Definition*:** The technological requirements of a system or framework that can provide a single view of the subsurface.

Why it's important to consider: A critical aspect of any future framework offering a single view of the subsurface is the technological development of the framework itself. 9% of the Survey respondents identified technology as a possible barrier to their involvement in future similar projects aiming to develop technological systems to facilitate a single view of the subsurface. Currently, many stakeholders from a wide range of industries use GIS (e.g. ArcMap, QGIS) and CAD software to understand spatial data. However there is no standardisation with respect to geospatial software programmes. Consequently BGS, for example, has approximately 20 geospatial software programmes that are interoperable with one another.

Middleware, such as FME (Feature Manipulation Engine), is a useful tool that can be used to convert between a wide range of data formats for use in multiple software packages.

The technological requirements of such a system or platform are difficult to define however, without provision of more detailed data inputs and use cases of data access and use, which will be developed in WP3 and future projects.

Relevant existing work: Technological barriers similar to those relevant to Project Iceberg have been overcome to achieve previous projects, albeit over much smaller spatial scales, and may offer insights into how similar barriers can be overcome in future.

At a local scale, the *Heathrow Map Live* project defined a Common Data Environment (a simple, easy-to-use web-based tool on an Oracle database) to reduce data duplication and facilitate data sharing at the organisation level.

At the city scale, the Glasgow-based *ASK Network* (Accessing Subsurface Knowledge) project developed an improved data exchange mechanism between the public and private sectors. This comprises both a web portal to check data compliance and a secure and centralised database, which can be accessed by government departures, local authorities, academic researchers, and the general public.

At the national scale, the University of Leeds developed the *VAULT* system as part of the *VISTA* project (Visualising integrated information on buried assets to reduce street works), which is now live across Scotland. This facilitates access to, and securely delivers, integrated information on utility and other underground apparatus to over 300 unique users across 47 different organisations from a centralised location.

Future recommendations: Projects such as *Mapping the Underworld* and *Assessing the Underworld* have shown that new technologies are developing and improving our current capability all the time. It is not within the scope of this report to define specific future technological capabilities, particularly without a greater understanding of current industry requirements.

Back to top

3.12 HOSTING VS. FEDERATION

Definition**:** The architecture of how the data sharing system operates.

Why it's important to consider: To provide a single view of the subsurface, there must be a means to bring disparate and disjointed datasets together from multiple data sources and suppliers. Hosting the data on a single server may cause issues to arise regarding data autonomy, intellectual property (IP), and inequality amongst stakeholders. A federated database is a more probable alternative to a single hosted server, where autonomous databases are connected via a computer network.

Relevant existing work: Previous projects have typically opted to use a centralised database as a means to combine multiple datasets. The *Heathrow Map Live* project defined a Common Data Environment, which constituted a simple, easy-to-use, web-based tool on an Oracle database. At the city scale, the Glasgow-based *ASK Network* (Accessing Subsurface Knowledge) project comprises both a web portal to check data compliance and a secure and centralised database. The *VAULT* system, developed as part of the *VISTA* project (Visualising integrated information on buried assets to reduce street works), securely delivers integrated information on utility and other underground apparatus to over 300 unique users across 47 different organisations from a centralised location. The coverage of these projects is not analogous with that of Project Iceberg however, which seeks to determine the barriers to implementation of a similar service at a national (e.g. Great Britain), rather than a local, scale. One example that does operate on a larger, regional scale is the Flanders underground utility location system – KLIP, which uses Windows Azure cloud storage to power an information sharing portal used by map requestors (construction, consultancies, companies and citizens), network authorities and public domain authorities.

Future recommendations: The desired coverage of any future solution is much wider than that of most of the previous projects listed above, however the KLIP solution is approaching a scale anticipated. Given the extensive range of stakeholders (e.g. OS, BGS, Utilities, Transport Authorities, DEFRA, Archaeologists), a hosted data framework is unlikely to be realised amongst issues of IP and data autonomy. In this instance, an impartial, regulatory, or governmental organisation/department may be required to host the server. A federated data framework constitutes a more likely solution, which allows asset and data owners to maintain authority and does not promote issues of stakeholder inequality. A federated system will also facilitate an improved data currency with respect to a hosted server.

Back to top

3.13 SECURITY

Definition**:** Addressing how to secure data accessible via the framework.

Why it's important to consider: Within the framework, there will be a lot of data, which must remain secure. This may range from the location of critical infrastructure to underground tunnels and access routes, to commercially sensitive data regarding asset condition and maintenance. Given that companies are valued based on their asset value, making such information more widely available may threaten their business model and market position. As with any data sharing framework, which can be accessed remotely, there will be a requirement for sufficient security to prevent uninvited users. There will also be sensitive data held within the framework, which may not be appropriate for widespread dissemination amongst stakeholders and/or the wider public. Over 40% of Survey respondents cited "commercially sensitive data" as a reason for them not sharing or selling subsurface data with other organisations in the current business-as-usual scenario. "Security concerns" was also a reason provided by an additional respondent. This is therefore viewed as an important barrier, which must be addressed and sufficiently overcome, to promote widespread stakeholder engagement.

Relevant existing work: As part of their work, BGS gain access to sensitive geospatial information about public water supply boreholes, which would be inappropriate for public dissemination in its raw form. To overcome this issue, BGS have access to sensitive information in-house for use on the project, and then deliver a publicly-accessible model which doesn't include sensitive sites.

The *VAULT* portal, developed as part of the *VISTA* project, has methods in place to try and increase security:

> **"By default data is placed upon the map on the Vault layer and is visible immediately. It is, however, possible to have data that is potentially sensitive placed on a hidden layer instead. This layer is not shown unless the user makes a request to view a section, at which time information is revealed (only for the area requested) and details are logged on the SRWR of the user requesting the data. If the user pans their view of the map to a new location, no new data is shown from the hidden layer unless the user makes a further request for data from the new section of map. This provides an audit trail of which users viewed the data contained within the hidden layers. This audit trail can be shared with appropriate data providers on request."**

(The Office of the Scottish Roadworks Commissioner (2015) available at: http://www.roadworksscotland.gov.uk/LegislationGuidance/Guidance/Vault.aspx).

Future recommendations: Security concerns amongst stakeholders will severely inhibit engagement with any future data sharing framework. This may require reduced data granularity at the point of access and dissemination. In this instance, it will therefore be important to ensure that the granularity of the data remains sufficient to meet stakeholder needs and improve decision-making processes (relative to the current business-as-usual scenario).

Back to top

3.14 PERSONAL DATA

***Definition*:** Sharing of the personal data that is held within the framework.

Why it's important to consider: Datasets uploaded to the platform will likely include, to some extent, personal information e.g. addressing. There are likely to be legal barriers to sharing such information, and consequently there will be a requirement to anonymise such data. In removing granularity and details such as this, there will be a loss in the value (to some extent) of the data that is shared through the framework. Whilst this may be inevitable, enough value must remain to encourage stakeholder engagement in a data sharing framework. How best to anonymise personal data, and how to retain value once anonymised are two key questions that will need to be addressed going forward.

Relevant existing work: BGS offer web-links to the Freedom of Information Act on their website (http://www.legislation.gov.uk/ukpga/2000/36/section/40), which provides legal regulations with respect to sharing personal data.

Future recommendations: The problem in this instance, is not the sharing, or lack of sharing, of personal data. There are already legal policies and procedures related to this in place, which must be followed. Ensuring that data is valuable once anonymised is the issue which must here be resolved. Understanding stakeholder needs and requirements will be critical to addressing this, so that data granularity can remain as close to that which is needed by stakeholders.

Back to top

3.15 ACCESS RULES

***Definition*:** Determining the level of access appropriate for/required by a stakeholder, whilst ensuring data owners can comfortably address legal/IP/commercial sensitivities.

Why it's important to consider: The ways in which stakeholders may participate in and interact with a subsurface data sharing framework can be classified into three main groups.

(1) Stakeholders that supply data to the framework (e.g. OS, BGS, utilities, DEFRA, transport, etc.)

(2) Stakeholders that access and use data uploaded to the framework (e.g. contractors, developers, academics, construction, etc.)

(3) Stakeholders that supply data to the framework and use additional data uploaded to the framework (e.g. utilities, BGS, transport, etc.)

To satisfy the range of engagement outlined above, there will be a requirement for access rules and regulations, which determine to what extent stakeholders can access data, the data they are able to access, and the costs associated with this access.

Relevant existing work: There are a range of previous projects, which have a wide range of users, from the general public through to industry through to public bodies. The *ASK Network*, Glasgow, for example, brings together over 20 industry partners and 12 public sector bodies in Scotland, and offers a data service which can be accessed by government departments, local authorities, academic researchers, and the general public alike.

Future recommendations: Given the wide range of possible data users that may have access to a future data shearing framework, different levels of access will be necessary to ensure data security. To successfully determine these levels of access and promote optimum engagement, a full analysis of user requirements will be needed. This will include those requirements of data suppliers, who will need to protect data supplied to the platform, and those of the data users, who will need data to be sufficient to meet their needs. An alternative approach would be to accept only those data which are suitable for all levels of use. This would reduce the value of such a framework for those stakeholders, who require more detailed information, and limit widespread engagement.

Back to top

3.16 ACCESSIBILITY/USABILITY

Definition**:** Ensuring that the data sharing framework addresses the needs of multiple stakeholders, with respect to accessing others' data and sharing their own.

Why it's important to consider: With so many data inputs, users, and sensitivities involved with a subsurface data sharing platform, there is a risk that it will become complex to use. Complexities such as this may inhibit interaction and engagement with the platform, which will reduce the traction any such framework may gain throughout the stakeholder group. There are a wide range of stakeholders involved in any project such as this. These range from BGS and utility companies, who deal with subsurface data regularly, to local councils and planners, who may only interact with subsurface data on an irregular basis. Members of the stakeholder group will also have had a range of previous exposure to traditional methods of displaying spatial data. Some may still predominantly use very traditional methods e.g. paper maps, whilst others will use GIS as part of their day-to-day life. For any subsurface data sharing framework to gain traction, this wide range of technical proficiencies must be addressed, if future use is to be as wide-reaching as possible.

Relevant existing work: Many stakeholders currently using and visualising subsurface information use CAD and GIS software, such as ArcMap and QGIS. *VAULT* was developed as part of the *VISTA* project, and is now live across Scotland. Demonstrations of the *VAULT* portal suggests it functions as a typical GIS, with data layers that can be turned on and off as required. This is similar to that of the PADU project driven by JUPEM, the Department of Survey and Mapping Malaysia. Users of these systems will have likely used interfaces similar to this before, and so should be able to easily adapt to this new system. Less technical users, with a more limited exposure to geospatial software, may not be as comfortable using an interface such as this.

Future recommendations: It will be important to ensure that any framework that is developed is user-friendly and accessible. This will require input from those qualified in User Design and User Experience and additional input from the users themselves. Whilst specific architecture solutions are not identified it is likely that web service and APIs to discover, access and potentially process data will be critical and a range of audience tailored end clients would need to be built upon these framework services to suit users.

Back to top

3.17 WHAT'S NEEDED?

Definition: The practicalities of the framework/system requirements.

Why it is important to consider: This problem space can essentially only be answered once all other problem space solutions have been defined. Decisions relating to how stakeholders will provide data, share and access information and the legal technicalities can only be made once a definitive framework/system has been decided upon. There are learnings however to be taken from previous initiatives that attempt to be fully prescriptive: the Digital National Framework (DNF) was an interoperability exercise enabled through the use of common references (TOIDS) associated with geospatially represented features. Where the system failed, in the opinion of the authors, was in its inability accommodate new feature types. Extensibility, as well as the ability to handle multiple versions or representations of the same object, is therefore a critical requirement.

Relevant existing work: Previous projects may have tried to define the parameters of a solution/framework prior to assessing all other problem spaces and have not been able to achieve project aims thus far. This may serve as a warning as to why it is important to listen to all stakeholder input, to have an over-arching holistic view of requirements, before creating any system architecture. In Flanders the KLIP project involved over five years of developing and implementing manual procedures that changed the data sharing culture and provided the necessary time to refine the business model(s). It wasn't until phase 2 when the system requirements for KLIP Digital were fully understood and a technical solution was developed iteratively and released. See: https://youtu.be/cyMJROmt7Eo

Future recommendations: The only recommendation is that user-led design and experience are considered during the practical arrangements of a framework. For example, if a certain software capability is required, to understand what stakeholders will interact with the data and to choose an appropriate package that is easily accessible, rather than expensive solution-specific programmes that may require additional expense. However, this does not preclude the adoption of key architectural principles to guide decision processes, such as a preference for distributed data storage, a high level dataflow and a belief that standards driven web services will be involved.

User led design is a key component of making such a scheme workable. We should strongly recommend that a real use case is required to inform the development of a pilot project. This may be something that can be developed in collaboration with the OGC Concept Development Study (http://www.opengeospatial.org/standards/requests) into underground mapping. Several Scandinavia authorities are working with the KLIP team to develop similar systems; we could also learn lessons from their experiences and potentially short cut the learning curve.

Back to top

3.18 FEEDBACK REPORTING

Definition: Providing a means by which users of the system may report problems or observations of the data to suppliers.

Why it is important to consider: As data input into the framework is shared, accessed and used by stakeholders, it is inevitable that problems or mistakes within the data will become apparent. Therefore, it is important to have a mechanism in place so that this may be reported back to framework owners/data suppliers and improved upon. As use and utilisation of the framework increases, it is likely that there will be increasingly greater numbers of “submissions” made through this process. By providing this route to the data suppliers, an open dialogue surrounding data capture and standards can be facilitated. An important aspect to consider is the level of expectation around whether this mechanism will be actioned by the party responsible for supplying the data. It may need to be a mandated part of the “framework” that works in tandem with legal liability, e.g. if you are submitting to this framework, you are liable to correct data based on user’s observations, or the rating of accuracy of the data will decrease. If not, users may feel that their concerns or observations are being ignored and therefore lose trust in the quality of the data/framework.

One other factor that will be likely to disrupt a simple flow of information between user and supplier is when the owner of the asset is not aware that it belongs to them or cannot be identified correctly. A mechanism for feedback and reporting within the data sharing environment must also have a way of making relevant stakeholders (e.g. companies that operate underground infrastructure within the area) aware of any underground unidentified objects or assets.

It is also important to consider validation processes from the feed-back loop. This can take the form of ‘crowd review’ (where information is exposed to a wider [potentially public] group that can then comment on that information) or a ‘real world’ ground-truthing of the information provided to the supplier highlighting inaccuracies within the data. If data is to be provided to the supplier as corrected and updated data, then there must be systems in place whereby data still moves through a quality control process before becoming accessible to other users. In this instance, it may be worth providing a set of standards for recaptured data that must be met. If the user only highlights that there may be some inaccuracies within the data but does not provide any additional corrected/updated data, then a different set of procedures and timescales must be put into place to ensure that companies do not engage with the Reporting/Feedback loop.

The GLA Network Utilities project demonstrated that currently limited feedback of site discrepancies, through a lack of a standard way of reporting incorrect records, was a barrier to sharing of information.

Relevant existing work: Within the newly released *PAS256 Buried assets – Capturing, recording, maintaining and sharing of location information and data – Code of practice*, a section provides a recommended workflow for the recording of both wrongly recorded objects (WRO) and unidentified buried objects (UBO). It clearly defines a process which should be adhered too, depending on whether it is a WRO or UBO. This procedure outlines how to record discrepancies, using a pre-determined template. Explicitly it must record general information such as “Data asset owner/potential asset owner informed", location information, description of the UBO and any additional comments. It also provides common descriptive terms that may be used as a guide for the person discovering the WRO or UBO, whose knowledge may not be appropriate for the recording/data capture. The creation of this template may provide a basis for the requirements of a feedback loop/reporting mechanism within a future data sharing environment, ensuring that all incorrect or missing data is provided to the same standard.

Future recommendations: In order to increase participation and trust in and use of a data sharing environment or framework, it is important that any discrepancies or inaccuracies within shared data can be reported. A feedback loop that allows users of the data to make this known to suppliers will work towards improving quality of data, allowing for new opportunities in asset management and reduction of delays to projects caused by previously unknown infrastructure to be realised. A standardised form of reporting is needed to ensure that data suppliers can act quickly and appropriately to improve their data, whilst providing users reporting a clear structure to adhere too. PAS256 moves some way towards providing this framework, however it may need to be tailored to be of most use and relevance to a specific data sharing environment and the stakeholders involved with it.

Back to top

3.19 ALERT

Definition: A way to ensure that relevant information/data is provided to stakeholders by pushing data when required.

Why it is important to consider: Within the data sharing environment, a way to make user experience more valuable could be to create an additional feature of optional alerts. This could ensure that maximum benefit could be extracted from data sharing, through intuitively providing stakeholders with relevant information. By identifying connections between stakeholders over a shared business need e.g. shared location/operating area, benefits, such as minimised disruption, could be realised through collaborative working. It could also provide further value to the user through identifying areas or needs to the business and making those stakeholders aware when there is either change or a lack of change. For example, housing developer A has purchased an area of land. Developer B has just submitted borehole log data to the "framework" extracted during recent construction of buildings that provides new insight into geological conditions in the area. Developer A would receive an alert that makes them aware that new data that could be of potential value to them has just been shared. This new information may affect projected timescales or construction requirements that would have been costly to discover further into the project lifecycle.

This passive alerting system means that stakeholders do not feel pressured to have to check manually, but rather are always engaging within the system.

Conversely, this must be explored through stakeholder user experience design to ensure that it is a mechanism that will assist and support engagement, rather than alienate users. Appropriate options should be available so that stakeholders feel free to engage as much or as little as is needed for their business need, whilst still encouraging collaboration between different industries. This may be

Relevant existing work: Previous projects such as The Scottish Road Works Register (SRWR), a system used by all roads authorities and undertakers in Scotland to coordinate their works, have facilitated collaboration and stakeholder engagement through a shared framework. The National Underground Assets Group (NUAG) presentation to the Highway Authorities & Utilities Committee (HAUC) in 2011 set out a system whereby stakeholders with underground asset data shared data through a web-based system, with one of the projected benefits to be greater coordination and collaborative working of street works. These previous projects have largely focused on coordination of street works, however these alerts can become more intelligent through enriched personalisation to provide stakeholders with greater insight of the areas or aspects that are relevant to them.

Future recommendations: By providing a feature that allows users to interact more intelligently with a data sharing environment, it is likely that user experience will be more positive and realise the benefits of working collaboratively more quickly. It may not be a feature within the initial beginnings of such a framework, but should ultimately be incorporated to facilitate cross-industry communication and collaboration.

Back to top

3.20 EXTERNAL ACCESS RULES

***Definition*:** Defining the regulations, with respect to what sharing/publishing/decision making is allowed based on data owned by other stakeholders.

Why it's important to consider: The above section (section 5.c) discusses the various levels of access that may need to be regulated in a subsurface data sharing framework. Data granularity will likely vary at these different levels of access. In this instance however, this term relates to the rules regarding how data is shared and disseminated more widely. Such regulations may result from a fear of misuse, such as that which (partially) drove the *BIM for the Subsurface* project, led by BGS, Keynetix, Atkins and Autodesk. For example, through a subsurface data sharing platform, Utility Company A may have access to detailed information about subsurface assets owned by Utility Company B and C, which can enable them to make more effective decisions. Using this information internally is not analogous with sharing this information with a contractor, for example, who may need to carry out work on behalf of Utility Company A. There may therefore, be need for additional permissions, regulations and licenses associated with how data is shared more widely by those stakeholders directly accessing data from the framework.

Relevant existing work: NUAG have previously proposed a national web-based solution service, the *National Assets Records Exchange*. As part of this proposal, they suggest that asset owners remain responsible for managing their own information and interacting with other NUAG stakeholders. Their solution entailed a map-based platform, where a point is selected and then information returned to the user, *post contact with the asset owners' customers*. This ensures the asset owner remains fully responsible for permitting the sharing of their data.

OS, for example, offer a publishing licence (https://www.ordnancesurvey.co.uk/business-and-government/licensing/licences/publishing.html), which permits data to be published in print and via digital mediums e.g. websites.

Future recommendations: Following on from the issues associated with licensing data use (section 7.a) there may need to be additional licensing terms relating to publishing information. If stakeholders using data provided via a sharing framework wish to publish or share this more widely, there may be a requirement to purchase a publishing licence. This will need to cover the granularity and detail that can be shared, depending on the ways in which users intend to publish data. This will need to be sufficient to protect data suppliers, but not too stringent to limit additional sharing where necessary.

Previous projects, which have involved the sharing of data from multiple sources, have made data available through use of web feeds and GIS software (e.g. ArcMap, QGIS). *VAULT*, a system developed as part of the *VISTA* project that is now live in Scotland, has an interface very similar to that of ArcMap, which would be appropriate for many users in more technical, geospatially-focused roles, as they will most likely have previous exposure to such interfaces. Those with reduced exposure to such systems may not be comfortable using such an interface however.

As part of its mandate, the Department of Survey and Mapping Malaysia (JUPEM) developed a national underground utility database (PADU). Similar to the *VAULT* system, this collated subsurface utility information, and supplied a more co-ordinated view in a GIS format. Japan's ROADIS project, which collates diverse information related to roads and facilities that occupy spaces both above and below ground, also used GIS technology to communicate information.

The Sydney Down Under Project, which integrates 3D BIM and geospatial and 2D data, was delivered as both a 3D model of the CBD. The *Emergency Services Spatial Information Library (ESSIL*) is front-ended by the *Spatial Information and Mapping System (SIMS)*, which offers an easy-to-use interface for decision makers.

BGS notes that user requirements vary by project, and they often offer a bespoke data discovery approach for most clients (e.g. GIS, CAD, etc.), which may not be possible within a co-ordinated data framework with such a variety of stakeholders.

Back to top

4. GENERAL PROBLEM SPACES

These problem spaces are running themes across the entire data flow.

4.1 BUSINESS MODEL

***Definition*:** Addressing the issue of ‘ownership’ of the data sharing framework (see section 4.e) to decide who absorbs the costs and who receives the (financial) benefits.

Why it’s important to consider: Within this theme, there are several issues to consider:

(1) There have been many financial benefits identified from a subsurface data sharing platform. For examples, there is approximately £150 million of third party damage to utility assets per year (Mayor of London, 2013). It has previously been recognised that improving subsurface asset location information could offer a return on investment (ROI) of up to US$21.00 for every US$1.00 spent (Zeiss, 2014). An issue that may arise here, however, is which, and to what extent, of the many stakeholders invested and involved in a subsurface data sharing platform recoup these costs, should they be reduced via a data sharing framework.

(2) Creating and maintaining subsurface data costs money. Currently, asset data owners sell subsurface data to generate revenue, with Survey respondents disclosing that they typically spend between £0-500,000 per year acquiring subsurface data from third parties. These revenue streams may need to continue to exist for those stakeholders supplying data (via the framework rather than via private contact), to promote optimum engagement with the framework.

(3) There will also be additional costs associated with the initial creation and development of a framework. Previous projects focusing on subsurface data sharing have received funding ranging from £540,000 (*BIM for the Subsurface*) to £3.5 million (*Mapping the Underworld*). The initial costs of creating and encouraging use of a data sharing framework can therefore not be ignored when looking at possible barriers to implementation. These high costs may only be limited to a period of framework creation and development however. The *Sydney Down Under* project, Australia, combines utility and building infrastructure information into a single database. This project initially proposed a pilot project in the north-western corner of the CBD, however realised that most of the cost would be in setting up the project, with expansion to the rest of the city possible with only a little extra capital.

(4) Finally, there are costs associated with stakeholder involvement in such a framework. Of the Survey respondents, >50% identified a “lack of resources (human or financial)” as a barrier to involvement in a subsurface data sharing framework.

It cannot be denied that the creation, development and maintenance of a subsurface data framework will incur significant costs. However, the possible savings associated with a functional data sharing environment (financial, and otherwise) are great enough to provide a business case for investment.

A possible issue arising in this instance is the timescale over which a return on this investment will be realised (see section 8.a.ii below).

Relevant existing work: Previous projects, such as *Mapping the Underworld* and *assessing the Underworld* have received research grants and funding so little information regarding costs of a business model area available. In 2011/2012, the National Underground Assets Group (NUAG) started the 12-month London Trial Project, which involved a web-based national asset record information sharing service (NRS) to improve access to information on buried assets. This was expected to be followed by nationwide implementation, where costs are shared amongst users based on an agreed apportionment model. The *GLA Network Utilities* project identified three possible scenarios with approximate costing attributed, however estimates range from £16 million - £51 million (2014-2019) and are specific to the future solutions they outlined. It was also shown that public sector intervention is unlikely to be justified on economic grounds in the period to 2018 without mitigation of some factors.

Future recommendations: Given that many companies sell their data and receive revenue from doing so, stakeholder engagement will be limited without some mechanism to replace this financial stream, or financial incentives to overcome this difference. The costs of developing and creating a subsurface data framework may be overcome in several ways. It has been previously discussed that there may be a need for a single custodian, who is responsible for its creation. If this custodian is a government department, then government funding may be sufficient to cover development costs. If this custodian is an organisation (e.g. OS, BGS), then the business model and pricing structure of data supply and use will need to incorporate a means to return capital to the original investor.

Back to top

4.2 IMPLEMENTATION TIMESCALES

Definition**:** The timescales over which a framework will be developed vs. the timescales over which returns on investment will be realised.

Why it's important to consider: The costs of developing a subsurface data sharing framework are great, as are the potential cost savings associated with a functioning solution to subsurface data sharing. All of those stakeholders currently creating and accessing subsurface data will have processes in place which are sufficient for their business-as-usual activities. Realising the benefits of a subsurface data sharing framework may be realised over much longer timescales than typical asset lifecycles or investment cycles. Water utilities only have asset management plans for 5-7 years, and gas and electric companies work over 8 year cycles (Revenue = Incentives + Innovation + Outputs). Even governance periods, should a government mandate occur in the future, are only secure for 5 years. Short term thinking may therefore limit stakeholder involvement, both with respect to their engagement and any financial commitments.

Relevant existing work: The *GLA Network Utilities* project outlined some future solution possibilities, but the timescales vary according to the scenario. In the long term vision of the project supporting a Smart London agenda, the vision of a fully mapped underground with a system fully complete and operational being realised was envisioned over the period of 2014-2035.

Future recommendations: Realising some benefits (financial, time-saving and health and safety) within short timescales analogous to stakeholder investment cycles will be critical to gain traction for a subsurface data sharing framework. If benefits can be realised, even just locally, over short timescales, this will secure stakeholder interest and investment. As short an implementation timescale as feasibly possible (but still long enough to ensure successful development) will reduce crossover between current business as usual practises and the more efficient data sharing processes associated with a functional data sharing framework.

Back to top

4.3 VALUE RECOGNITION

Definition**:** Recognition of value by stakeholders that is required to promote and encourage engagement with a subsurface data sharing platform.

Why it's important to consider: Stakeholder engagement is critical to the success of the proposed data sharing framework, as stakeholders will be required to supply and maintain data to known standards, and engage as users. One of the possible reasons for limited engagement is a lack of value recognition by involved stakeholders, who will have developed a tolerance to the frustrations of the business as usual scenario, and who will have to continue with this business-as-usual scenario until a framework is developed and functioning to meet their needs. This may be due to a lack of recognition of the value of data (e.g. a global survey of utilities in 2013 found that up to 60% do not consider big data analytics a significant opportunity to improve the delivery of their services; Capgemini Consulting, 2015); or may be due to the specific value they associate with a project such as this.

Mapping the Underworld, *Assessing the Underworld* projects, and a London-based project by *GLA Network Utilities* identified a lack of value recognition as a challenge to engagement and stakeholder buy in. This has also been echoed by 17% of the Survey respondents, who identified a lack of value as a specific factor limiting their involvement.

A possible issue here is that whilst there are many business cases which promote a subsurface data sharing framework, there is no single business case that can be employed to portray the value in developing a framework such as this.

Relevant existing work: Previous work shows that it is possible to overcome this barrier, given that there has been a wealth of previous projects which incorporate a wide range of stakeholders. The *ASK Network*, Glasgow, for example, has accumulated over 20 industry partners and 12 public sector bodies in Scotland. A wide range of organisations expressed an interest in participating in a previous *NUAG* project proposing to create a *National Asset Records Exchange*. Interested organisations include: Thames Water, Virgin Media, Southern Gas, TfL, National Grid, Mayor's Office, City of London, Sutton and east Surrey water, Crossrail, Network Rail, BT (not an exhaustive list). Value recognition may also be aided by a shift from isolated CAPEX (Capital Expenditure) and OPEX (Operational Expenditure) budgets to TOTEX (Total Expenditure) budgets, which promote a more holistic organisational perspective.

Future recommendations: Increased education about the benefits and value of data analytics, and a clear cost-benefit analysis of how and when (see sections 8.a.i and 8.a.ii) stakeholders will recoup their costs of involvement and realise the financial/time-saving benefits of a functional subsurface data sharing framework is a possible resolution to this barrier. It is hoped that the outcomes of WP1 of Project Iceberg may go some way to addressing this. If a successful framework is to be developed, it will require input and engagement from as wide a range of stakeholders as possible. If value recognition is a factor which limits participation, there may be a requirement for government involvement or regulations which promote engagement and involvement.

Back to top

4.4 ONTOLOGY/SEMANTICS/TERMINOLOGY

Definition: Throughout Project Iceberg, previous projects concerned with sharing subsurface data and even throughout varying industries there is a disparity between what different stakeholders mean. In order to have a single, coordinated view of the subsurface, to allow the data to be interoperable and to derive value from sharing data, there must be a common understanding based on shared definitions or dictionary.

Why it is important to consider: In order for varying stakeholders that all have an interest in accessing subsurface information to derive value from shared data, every framework participant must share the same view and interpretation of the data. A perfect example of why cementing a shared ontology within the "framework" can be found within the utilities industry. Beck's 2009 study highlighted that every utility company, apart from those that have sewers within their asset portfolio, refer to the measurement of depth as from the ground surface level to the top of the asset. However when measurements are made of sewer depth, these data refer to the distance between the ground surface level to the bottom internal measurement of the asset. The differences between these two measurements could be of considerable difference and therefore affect any planned works. Generally cost of works increases with depth and the conditions that affect project requirements could change resulting in delays. This is just one example of varying terms and definitions across one industry; the scale of this interoperable terminological barrier increases as the numbers of industries increases.

Relevant existing work: Fu and Cohn's (2008) paper, "Semantic Integration for Mapping the Underworld" examines techniques for reconciling semantic heterogeneity within the utility domain, through the creation of a utility thesaurus. The thesaurus was designed to work as a reference vocabulary on which to base the identification of mapping between utility data and subsequent resolution for semantic heterogeneity.

Techniques outlined within their paper move some way towards working around the problems associated with integrating heterogeneous data and the issues of interoperability associated. The key benefit identified by Fu and Cohn is of removing domain inconsistencies, resulting in improved understanding by employing transparent naming and representation standards.

Future recommendations: In the future development of a data sharing environment, value cannot be derived from sharing data alone – it needs to be both useful and understandable to the user.

During construction of the “framework” and associated standards, an awareness of how to improve interoperability, common terminology and heterogeneity of data must be considered.

Back to top

4.5 INTEROPERABILITY

Definition: Common theme/problem that has stopped previous attempts at encouraging sharing of subsurface data from achieving their aims and often cited as one of the biggest barriers to data sharing.

Why it is important to consider: The interoperability problem space can broadly be split into three sections;

(1) Syntactic
(2) Semantic
(3) Schematic

Within these sections, there is incredible difference in the way in which industries, organisations and individuals, capture, store and interpret subsurface data. The data is encoded in an uncoordinated way i.e. without consideration of compatibility and interoperability with other utility systems (Beck et al., 2008).

(1) Syntactic: Syntactic heterogeneity refers to the difference in data format e.g. ESRI shapefile, Geography Markup Language (GML). These differences arise from the software that individual utility companies use to both use and store the data, ranging from ArcGIS to MapInfo and Oracle to SQL Server.

(2) Semantic: Semantic heterogeneity refers to differences in naming conventions and conceptual groupings in different companies. This can be further split into synonym and homonym mismatch. Synonym mismatch occurs whereby identical items are named differently. Homonym mismatch occurs whereby different data items are identified identically in different systems. Examples of this can be seen in the figures in the **Ontology/Semantics/Terminology** section of this document. The degree of schematic granularity can also exacerbate this issue, with two different companies identifying the same item but with different levels of detail e.g. with fine granularity – Medium Density PolyEthylene pipe – or as coarse granularity – Plastic pipe.

(3) Schematic: Schematic heterogeneity refers to the differences in data model between organisations. These occurs due to different business needs and interests. This impacts the type of information recorded, the ways the information is represented, how various information and data relate to each other and various semantics attached to records. This is demonstrated through a number of different issues; record structures, type mismatch (same class of data assigned with different data types e.g. what is described through a text field in one organisation could be described with a numerical field in another), range mismatch (different value ranges in fields) and granularity (different levels of details e.g. mains pipes in one organisation but the addition of services pipes in another).

Obvious barriers to data sharing are simply whether the data is digital or not and can originate from historical data. For example, Water utility companies inherited assets when the industry was privatised in the 1970s, and asset management records were passed to new companies on paper documents. Booth et al. (2016) estimate that if data sharing processes were digitised and automated, utility company profits could increase by as much as 30% and staff productivity by 15%.

Methods of integration are available currently, such as the use of Feature Manipulation Engine (FME), however this is not an automated process and carried out on a project-by-project basis. However in order to be an integrated view of both below and above the surface, this will need to be scaled up to encompass all data held by various stakeholders about the subsurface.

Relevant existing work: Beck et al.'s (2008) paper investigates the varying types of data heterogeneity that are found within the utility space and provides potential guidance and solutions to integrate different data systems. The *VISTA* project (Visualising integrated information on buried assets to reduce street works), undertaken by academic institutions (Universities of Nottingham and Leeds) and funded by the Department for Trade and Industry, looked at how to resolve semantic heterogeneity within the UK utility domain.

SYNTACTIC:

- Recommended use of Open Geospatial Consortium (OGC) approved formats.
- Use of OGC approved syntactically interoperable formats and services such as GML and Web Feature Services (WFS).
- Recommended use of proposed global schema approach that uses Oracle Database 10g architecture, that is highly scalable and grid enabled, importing data through FME.

SEMANTIC:

- Proposed synonym homogeneity through creation of a universal thesaurus.
- Homonym homogenisation can be achieved through ontology mapping, but a future recommended standard created in collaboration with stakeholders for more intuitive and timely integration.

SCHEMATIC:

- Development of a better integration tool may be difficult as current capability is relatively powerful already (COMA+). Currently most successful integration is achieved manually or semi-automatically.
- The *VISTA* project developed a conceptual framework to support utility data integration, which includes a number of specific processing and validation steps.

Ongoing development of this process will have to take into account that in order to be adopted by industry, an awareness that organisations will be unlikely to change its internal data structures and business process to facilitate such integration must be recognised.

Future recommendations: Any future solution to the above problems must always be developed with the context of industry participation in mind. Integration tools are likely to improve with ongoing research, including that of the *VISTA* project, however it may take regulation to enforce global schema participation within industry if it calls for altering current business workflow.

Back to top

4.6 COMMERCIAL SENSITIVITY

Definition**:** An unwillingness to share subsurface data that may be of commercial value to an organisation.

Why it's important to consider: Many organisations will be unwilling to share (or share widely) any data they hold which gives them a business advantage, such as data about the location, conditions, and risk of their assets. Survey respondents repeatedly raised such issues. Over 40% of the respondents cited "commercially sensitive data" as a reason for not sharing or selling data in the current business-as-usual scenario and "data confidentiality" was given as a possible factor limiting involvement in a subsurface data sharing framework (35%).

Within *PAS256: Buried assets – Capturing, recording, maintaining and sharing of location information and data – Code of practice*, there is acknowledgement of security-related concerns and an awareness that these must be taken into account when discussing and promoting subsurface data sharing.

Relevant existing work: *PAS1192-5* is a specification dealing with security mindedness relating to BIM, digital built environments and smart asset management. This specifies processes that:

> **"*will assist organisations in identifying and implementing appropriate and proportionate measures to reduce the risk of loss or disclosure of information which could impact on the safety and security of:***
>
> - ***personnel and other occupants or users of the built asset and its services;***
> - ***the built asset itself;***
> - ***asset information; and/or***
> - ***the benefits the built asset exists to deliver.*"**

(CNPI, 2015, Available at: https://www.cpni.gov.uk/system/files/documents/18/6f/BIM-Introduction-To-PAS1192-5.pdf)

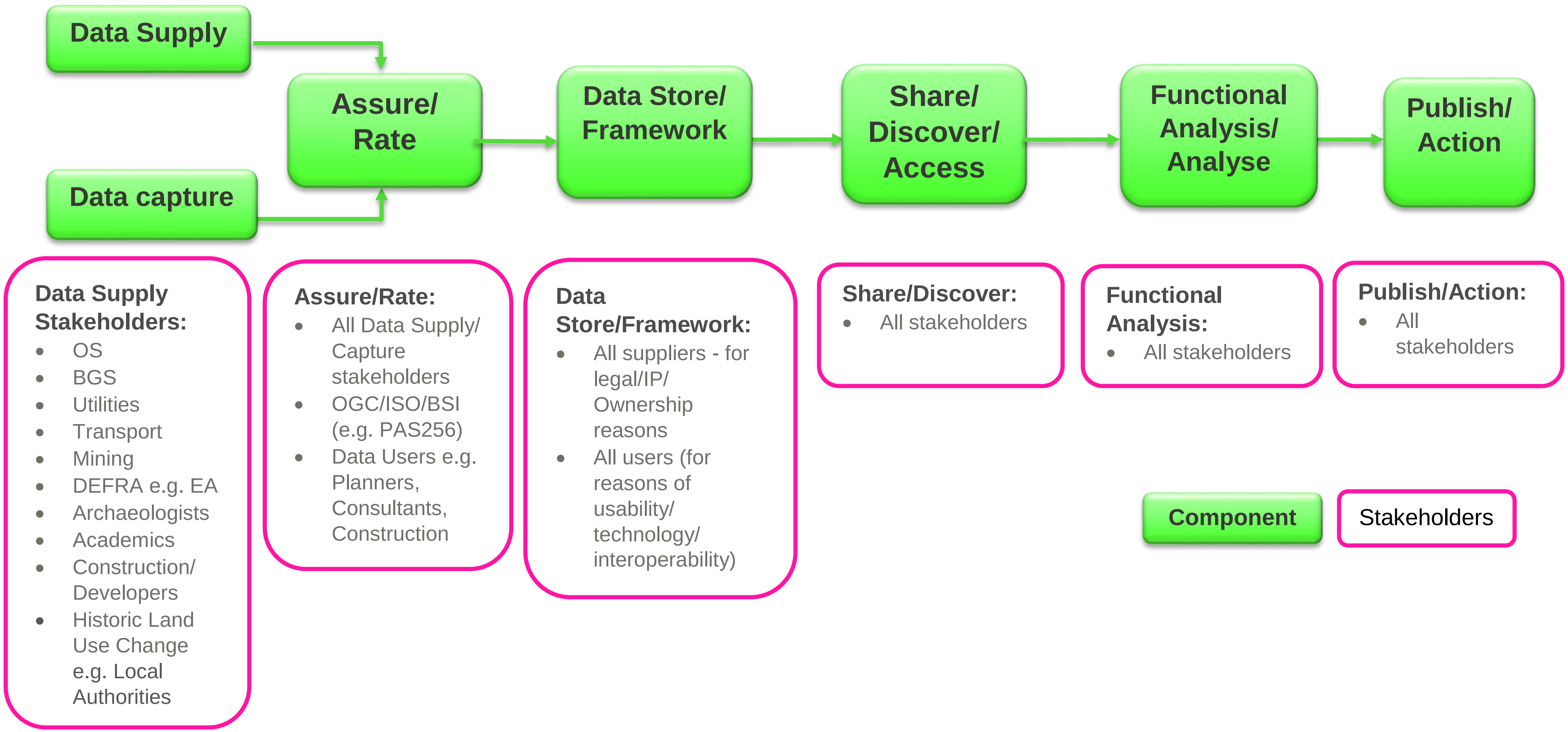
It notes, however, that these processes can also be applied to protect IP commercially valuable information.

For more information about external security issues, see section 3.13 Security.

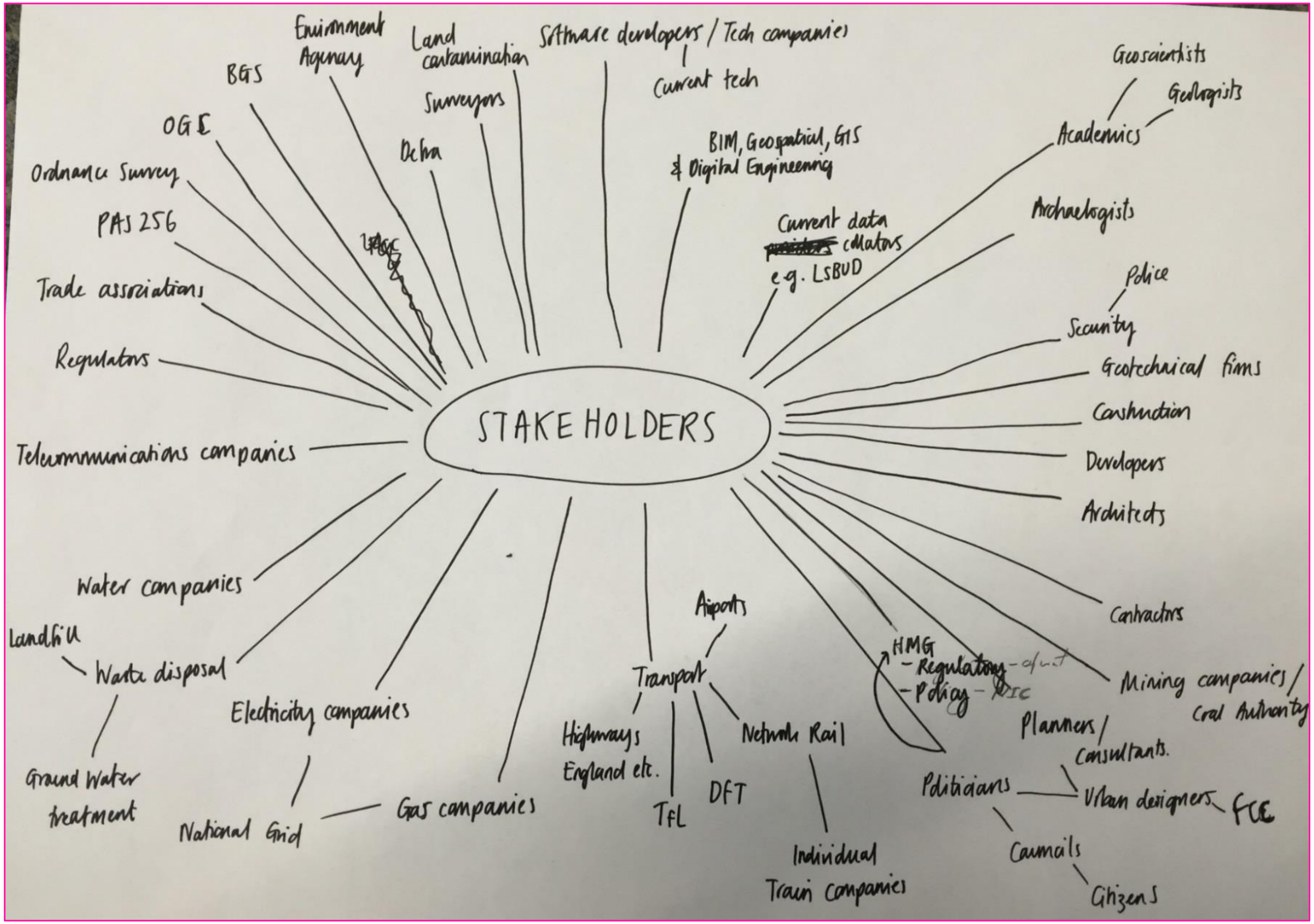
Future recommendations: Commercial sensitivity must be a barrier that is addressed to encourage widespread stakeholder engagement. Commercially sensitive data could be handled at the point where data suppliers submit data into the shared environment. This would enable suppliers to maintain control over data that they feel is of commercial value, however may limit the value of a data sharing framework, depending on their decisions over this. Commercially sensitive data could also be handled at the point of use. For example, at the point of use, and in line with a user's access level (see sections 5.c and 7.c).data could be delivered at a reduced granularity, or could be omitted completely. The risk here is around connecting a wide range (and a large volume) of commercially sensitive data via a federated network (or in a centralised database), which can be accessed via a single point.

Back to top

Appendix A. A conceptual model of how a flow of data could be constructed and relevant stakeholders.



Appendix B. Stakeholder diagram.



Appendix C. Stakeholder Expert Insights

A wide range of stakeholders are engaged in activities that aim to better understand the subsurface ground conditions and the buried infrastructure contained within in it. As part of Project Iceberg's market research, several stakeholders, working within the planning, utilities, mapping and research sectors, were interviewed and invited to take part in a survey.

The aim of the market research was to capture information about existing investment and capability and enable experts to share learning and offer insights on this topic. All those that took part in the survey are either owners of subsurface data or users of third-party subsurface data.

Key insights from the survey are highlighted below:

- While the **exact costs of acquiring subsurface data** have not yet been quantified or were unknown to survey respondents, they **are deemed to be quite high** by some of the respondents as they usually require in-house experts, external consultants and liaisons with data owners for a comprehensive view of the subsurface.
- Two-thirds of stakeholders say that their organisation **incurs indirect costs** as a result of incomplete information about the subsurface.
- The two **major impacts** of incomplete subsurface information are **delays to projects** and the need for **additional surveys**.
- Around half of the responses quoted positional accuracy for their buried asset **locations measured at metre scale** – highlighting the low level of accuracy currently in place across asset owners.
- Respondents also mentioned the continued use of traditional GIS - data transfer from (normally) '**quite poor databases'**.
- With **75% of respondents using their own and third-party subsurface data**, the need for a more efficient, data exchange framework is more apparent.
- Lack of subsurface information means that the **land value is not being realised**. For example, developers will avoid land where there is high uncertainty on risks or costs.
- **Wide customer base** for the datasets exists and some organisations are realising the commercial opportunities of subsurface data products and services already.
- The existing subsurface **datasets are highly variable** in terms of coverage, accuracy, format, scales which limits accessibility and usability.
- One of the main barriers to sharing subsurface data relate to **security for data of national importance**; Other perceived barriers include intellectual property rights for data of commercial interest, lack of awareness of the benefits that subsurface data brings, lack of demand within utility sector for subsurface data services and a lack of time and resources to invest in resolving the issues.
- Despite the barriers, **two-thirds** of respondents **would like to see a** subsurface **data exchange** platform and increasing open access to data.
- The data exchange system **needs a geospatial interface**; GIS/Web formats are preferred with open and closed functionality; 3D/4D elements need to be considered; Open to commercialisation of services/products.

APPENDIX D PROJECTS OF RELEVANCE TO UNIFIED DATA FRAMEWORK FOR INTEGRATED MAPPING ABOVE-BELOW GROUND

PROJECT	OVERVIEW	MORE INFORMATION
ASK Network	The Accessing Subsurface Knowledge (ASK) Network is a data and knowledge exchange network between public and private sectors developed by BGS and Glasgow City Council (GCC) with support from other partners in the public and private sectors. The project aimed to develop and exchange high quality systematic subsurface data sets and models. A web-portal provides a data transfer mechanism to a centralised repository for raw subsurface borehole data in standardised formats.	http://www.bgs.ac.uk/research/engineeringGeology/urbanGeoscience/Clyde/askNetwork/home.html
BIM for the subsurface	BIM for the subsurface project aims to address issues such as project delays due to unforeseen ground conditions by applying the BIM process directly to ground investigation & subsurface infrastructure design. The Geotechnical BIM suite, allows historical geological data to be served digitally via APIs, through to BIM systems such as AutoCAD.	http://www.keynetix.com/bimforthesubsurface/
CityVerve	CityVerve is the UK IoT demonstrator project, aiming to build and deliver a smarter, more connected Manchester. CityVerve aims to create a blueprint for smart cities worldwide using IoT sensors and collaborative platform technology.	http://www.cityverve.org.uk/
Digital Built Britain	Digital Built Britain is a government-led strategy, utilising Building Information Modelling (BIM) in combination with the Internet of Things (IoT), advanced data analytics and the digital economy, to enable better planning of new infrastructure, at lower costs, with improved efficiencies in operation and maintenance.	http://digital-built-britain.com/
GLA Network Utilities	Greater London Authority (GLA) project to evaluate how a co-ordinated system of utility mapping could be implemented across London. This included reviewing and summarising existing smart utility mapping projects across London.	https://www.london.gov.uk/what-we-do/business-and-economy/better-infrastructure/london-infrastructure-map?source=vanityurl
Greater Manchester Open Data Infrastructure Map (GMODIN)	The GMODIN is an open map of relevant public and private infrastructure data from open public sector and environmental assets to energy utility networks. One of the issues many projects have faced is fitting their datasets into a pre-agreed, top-down schema. The GMODIN took away that	http://mappinggm.org.uk/gmodin/

	hassle for local authorities, and instead asked for any data they had, in any format. A series of schemas were built up, which were then passed back to the local authorities to use in future. The OS Maps API is used to deliver the tool.	
Heathrow Map Live	The Heathrow Map Live project aimed to reduce infrastructure strike incidents involving utilities during excavation by improving data reliability and accessibility. Heathrow defined a Common Data Environment (CDE) where data is created once only (i.e. single owner) and shared across organisations.	https://geospatialworldforum.org/2012/gwf_PDF/Nigel Stroud.pdf
JUPEM – Malaysia National Underground Utility Database	The Malaysian Department of Survey and Mapping Malaysia (JUPEM), has developed a national underground utility database (PADU) to act as a repository of underground data provided by utilities in a GIS format.	
Mapping the Underworld	Mapping the Underworld (MTU) is a 10-year research programme led by University of Birmingham, funded by EPSRC, which seeks to develop the means to locate, map in 3D and record infrastructure assets, using a single shared multi-sensor platform, so that the position of all buried assets can be known without excavation.	http://www.mappingtheunderworld.ac.uk/
NUAG - National Underground Assets Group	NUAG, is an independent organisation set up in 2005, to represent stakeholders with an interest in, or affected by, capturing, recording, storing and sharing of information on buried and associated above-ground assets such as pipes and cables. NUAG established the standards and processes for information creation and exchange to ensure consistency in referencing and recording asset information. NUAG proposed a national web-based solution service, the National Asset Records Exchange.	http://www.energynetworks.org/assets/files/electricity/engineering/Street%20Works%202012/11%20Mike%20Farrrimond.pdf
Open Geospatial Consortium (OGC)	The OGC led an Underground Concept Development Study (CDS) to document the progress made by OGC and its members to build a complete picture of the present situation and develop a conceptual framework for action to improve underground infrastructure data interoperability. The report also identifies the most important steps to be taken next in order to develop the necessary data standards and foster their adoption. Following this review an Underground Pilot to verify, standards-based interoperability for 'smarter' underground projects in cities is underway.	http://www.opengeospatial.org/projects/initiatives/undergroundcds
PAS Standards (128 & 256)	British Standards Institution PAS 128, provides specification for underground utility detection, verification and location, enabling the utility survey industry to deliver its services to a recognised level of accuracy.	http://www.pas128.co.uk/ https://www.bsigroup.com/en-GB/

	PAS 128 (2014) focuses on levels of accuracy – referred to as Survey Category Types – that can be specified when requiring a PAS 128 compliant underground utility survey. PAS256 (2017) sets a consistent, accessible data protocol to enable effective recording and sharing of the location, state, and nature of buried assets, and recommends how existing asset records should be updated, recorded and collated.	
ROADIS – Japan Road Administration Information System	Japan's ROad ADministration Information System (ROADIS) enables a central oversight of the locations of on-ground and below-ground critical infrastructure sing GIS systems. Online connections between host computers installed at each ROADIC branch office and the terminals and mapping systems of road administrators and utilities enable mutual utilisation of data.	
Sydney Down Under	The NSW Emergency Information Coordination Unit (EICU) and the City of Sydney collaborated to develop an intelligent 3D model of buildings and infrastructure, above and below ground in the central business district (CBD). It supports full attribute and 3D spatial queries on all features: buildings (both above and below ground), utilities and tunnels. The data was held in a mix of databases with their own data models and was integrated using data integration software (Safe Software's FME solution). This is front-ended by the Spatial Information and Mapping System (SIMS), which bundles applications and data into an easy to use interface for decision makers.	http://geospatial.blogs.com/geosp atial/2011/08/gita-anz-2011-sydney-down-under-compiles-comprehensive-digital-model-of-urban-infrastructure.html
VISTA - Visualising integrated information on buried assets to reduce streetworks	Project VISTA (Visualising integrated information on buried assets to reduce streetworks) was a follow-on activity to the MTU project to develop visualisation techniques which integrate subsurface data, and enhance their legacy - disseminating the information to digging teams and network planners. Similar objectives underpinning the two projects VISTA and MTU led to the collaboration of 22 utilities and partners to create and trial one combined system, now commercially realised as VAULT.	http://www.roadworksscotland.gov.uk/LegislationGuidance/Guidance/Vault.aspx

5.REFERENCES & RESOURCES

ASK Network: https://www.bgs.ac.uk/askNetwork/ [Accessed November 2016]

Assessing the Underworld: http://assessingtheunderworld.org/ [Accessed January 2017]

Beck, A.R., Cohn, A.G., Sanderson, M., Ramage, S., Tagg, C., Fu, G., Bennett, B. and Stell, J.G., 2008, October. UK utility data integration: overcoming schematic heterogeneity. In *Sixth International Conference on Advanced Optical Materials and Devices* (pp. 71431Z-71431Z). International Society for Optics and Photonics.

Beck, R., Fu, G., Cohn, A., Bennett, B. and Stell, J. (2007) 'A framework for utility data integration in the UK', in Coors, M., Rumor, M., Fendel, E. and Zlatanova, S. (eds) Urban Data Management Society Symposium, (Stuttgart, Germany, October 2007).

BIM for the Subsurface: http://www.keynetix.com/bimforthesubsurface/ [Accessed January 2016]

Capgemini Consulting (2015) Big Data BlackOut: Are Utilities Powering Up Their Data Analytics? Capgemini Group: Paris.

Chapman, T. (2008) The relevance of developer costs in geotechnical risk management. In: Foundations: Proceedings of the Second BGA International Conference on Foundations, ICOF2008. Brown M. J., Bransby M. F.,Brennan A. J. and Knappett J. A. (Editors). IHS BRE Press, 2008. EP93, ISBN 978-1-84806-044-9.

Department of Survey and Mapping Malaysia, JUPEM: 'Underground Utility Mapping In Malaysia – Moving Beyond 2020,' Department of Survey and Mapping Malaysia, Oct 2015

Digital Built Britain: http://digital-built-britain.com/ [Accessed February 2017]

Digital National Framework:
http://lisg.bcs.org/lisg/sites/default/files/downloads/DNF0001_3_00_Overview_1.pdf

Fu, G. and Cohn, A.G. (2008) October. Semantic integration for mapping the underworld. In *Sixth International Conference on Advanced Optical Materials and Devices* (pp. 714327-714327). International Society for Optics and Photonics. Booth et al. (2016)

GLA Network Utilities: ARUP GLA Presentation, Available on Project Iceberg shared Google Drive

Heathrow Map Live: https://www.linkedin.com/pulse/20140522175912-840956-accelerating-world-wide-initiatives-to-map-underground-utilities [Accessed February 2017]

Mapping the Underworld: http://www.mappingtheunderworld.ac.uk/ [Accessed January 2017]

Mayor of London (2013) Smart London Plan: Using the creative power of new technologies to serve London and improve Londoners' lives, Smart London Board: London.

Metje, N, Ahmad, B and Crossland, S. (2015) 'Causes, impacts and costs of strikes on buried utility assets', Institution of Civil Engineers, Municipal Engineer, 168, 3: 165-174.

National Underground Assets Group: https://www.waterbriefing.org/home/technology-focus/item/4086-new-national-underground-asset-information-service-launches-in-london [Accessed February 2017]

(Also see NUAG Presentation, available on Project Iceberg shared Google Drive.)

Open Geospatial Consortium: http://www.opengeospatial.org/standards [Accessed February 2017]

PAS128: Personal Communication with Izabela Hurst (2016)

PAS 1192-5: CNPI (2015) Available at: https://www.cpni.gov.uk/system/files/documents/18/6f/BIM-Introduction-To-PAS1192-5.pdf)

PAS256:2017 Buried assets. Capturing, recording, maintaining and sharing of location information and data. Code of practice http://shop.bsigroup.com/ProductDetail?pid=000000000030336926

ROADIS Project: http://www.gita-japan.com/pdf/English.pdf

Scottish Road Works Register, VAULT: The Office of the Scottish Roadworks Commissioner (2015) http://www.roadworksscotland.gov.uk/LegislationGuidance/Guidance/Vault.aspx

Sydney Down Under Project: http://www.spatialsource.com.au/gpsnav/mapping-the-sydney-underground

TfL (2010). Minimising disruption from roadworks, Surface Transport Panel, Transport for London. May 2010. http://content.tfl.gov.uk/Item07-Minimising-Disruption-Roadworks-STP-19-05-2010.pdf. [Accessed October 2017].

Utility Strike Avoidance Group 2014 report: http://www.utilitystrikeavoidancegroup.org/uploads/1/3/6/6/13667105/usag_2014_report_july.pdf

VISTA (Visualising integrated information on buried assets to reduce street works): http://www.comp.leeds.ac.uk/mtu/vista.htm [Accessed November 2016]

Zeiss, G. (2014) Accelerating worldwide initiatives to map underground utilities. [Online]. Available on Pulse's LinkedIn page [26 July 2016]