Session 2 Notes – Data and Digitisation

Table 1

Nataliya Aleksieva – NR, Amar Sohanpal-MOD, Monika Kreitmar CSIC, Felipe Rojas Parra CSIC, Matthew Foote-WTW, Fergus Harradence-BEIS, Carmen Muriana Cobo-TfL, Joanna Bonnett-Cowi, Brian Sheil-CSIC, Robert Mair, CSIC

Flipchart Notes

Question 1 how is industry transforming data into knowledge?

1. Is data availability really increasing – lack of data + silos/barriers to dissemination
2. Can be useful for specific applications – but only focus on data needed for specific project needs.
3. Some engineering disciplines do this better, eg. Aviation, motor – infra less so -> downside risk of measuring + reporting awkward issues.
4. Data gathering is not trivial – lack of capacity, capability or willingness.
5. External pressure eg. Building safety standard? -> force reg – carbon ?
6. Regulatory drivers to require increased data capture and disclosure.

Questions 2 What practical constraints exist?

1. Industry is not nature enough
2. Asset level – can measure maturity – systems level more problematic
3. Data should support risk identification and risk management 'diagnosis -> cure'.
4. Tailored data – and use of appropriate data
5. Size of data- how to process + curate?
6. Legislation drivers to capture of data –eg H+S -> similar for (infrastructure?) in future?

Questions 3 what makes data + digitalization challenging?

1. Heterogencity of project/asset/teams makes this difficult – difficult systems
2. Resource constraints /lack of prioritization to capture, curate +provide data?
3. Risk assessment should be central to projects – auctioning critical.
4. Data could shift focus of risk + responsibility through supply chain.
5. Data ownership – for public good? -> or commercial? – data sensitivity + security?
6. Interoperable data – standards?
7. Weak regulation? Major challenge? No shared vision?

- Liability for poor quality data
- Ownership of risk + consequence -> can influence desire to capture data?
- Different ownership groups

Other notes- Felipe

Table 2

Daniele Fornelli- Geo Obs, Haris Alexakis- Aston Uni, Nikolas Makasis CSIC, Viviana Bastidas Melo CSIC, David Pocock-Jacobs, Chris Campbell-Skanska, John St Leger-HS, Chrysoula Litina-NH, Manar Alsaif, UoC, Dongfang Liang-CSIC

Flipchart

Question 1 Data to knowledge

- Is it happening in industry
- Instrumentation and data gathering still a challenge. Why ?
- Adoption of new methods – higher perceived risk
- Convincing individuals/confidence
- Need proven testing of technology
- Trust in old ways – redundancy – risk management
- Knock-on effects?
- There is movement into digital approaches – but many moving parts
- Digitally capable differently at different stages of projects
- Is age group a factor?
- Need overlap and knowledge share
- Instrumenting existing structures with confidence?
- Asset safety & identifying benefits
- How do you ensure safety within and beyond lifetime without monitoring?
- Need proper understanding of how data will be used/be helpful?
- Many examples of data gathering abandoned
- Multiple layers of benefit from data need to be accounted for.

Lesson not to do excessive instrumentation
Need critical/useful measuring – failure prevention
Monitoring to enable quick response & prevent bigger issues, eg. Tunnel closures

- Mentality. We collect data, in this way
- Instrumentation field is a learning curve
- Certain assets not designed to be replaced and very expensive to do so
- Confidence and monitoring can determine severity of issues and inform action
- Asset lifecycle : Benefit, cost, avoidance of issues
- Challenges stem from > data quality /amount/ownership > available methods
- What is total rationale? Problem? Benefits?
- Not yet business-as-usual
- Extra cost/efforts to incorporate data-driven approaches
- Knowing what to look for – early signs of problems
- A lot of data collection but unsure if max benefits
- Benefits depend on who you ask

Maturity levels

- Holistic cost/benefits over x years

Often benefits come many years later for someone

Other notes

1. **Given the increasing presence of data in infrastructure applications, in your opinion, to what extent is the industry transforming this data into knowledge – providing actionable insights and informing decision-making?**

There are key areas that have been involved in instrumentation and have been used to gather data. However, the adoption of such technologies and the gathering of data are still a challenge. For example, there is a lack of acceptance of data and its interpretation. Part of the data specification is missing, and there are barriers in terms of knowledge and how is it perceived. In general, the adoption of new methods is higher perceived as a risk.

People generally do not trust new technologies and miss the opportunity to use the data to make decisions. In part, this is because there is a lack of testing, for example, testing the application of new technologies by using use cases that help people to measure the risk of the unknown. Testing can help people to have enough confidence to adopt new technologies that collect data from infrastructure. Then, if you change something in the next years you can test it first. It will allow you to understand what are the skills and the connectivity needed. This can help stakeholders trust the design and demonstrate value as well as engage with other stakeholders in the future.

There is a need to find a way to balance the risk, or we (practitioners) can keep doing the same. So, there is a need to show what is happening and it is important to always the question, what if? For example, companies are moving into the use and analysis of data, but this is a part of cultural change. Because there is a lot of data, should we (practitioners) collect everything, with the needed quality, is it enough?

Design in the construction sector is more complex, then we need other capabilities and skills. New people coming in, should be part of the jobs to keep learning new things, what is the knowledge required? How do you measure? Are they the right measures? For example, how do we measure confidence? We need to have enough evidence to quantify the benefits.

For example, in the context of safety, there are regulations that guarantee we develop systems safely. But how we can install very high-quality sensors in the bridges for the safety of the people?

2. **At the present time, what are the practical constraints preventing the civil engineering industry from achieving better solutions that are driven by data, given the current maturity of sensing technologies and data science methods? Do the current main challenges relate to the amount or quality of data being collected, the data ownership/availability, or the methods to make use of these data?**

The main practical limitations that prevent the civil engineering industry from achieving better data-driven solutions are related to the fact that people do not know what to do with such data. People are required to first have the correct justification based on problem framing and cost-benefit analysis: to understand what is happening, what are the problems, what are the needs and what are the benefits.

We need to embed the design as a monitoring plan and invest extra money for that. This implies that monitoring data is not business as usual and for that, we need a strategic plan.

The monitoring of different physical infrastructures in advance can help us to identify possible problems. Although there are a lot of data, we do not know how to convert actual data into knowledge. There are problems with quality, we ask for data, but we do not know whether we are asking for the right data.

Furthermore, in terms of current challenges, there are some related to the ability to show tangible benefits of data-driven solutions in the civil engineering industry. It is sometimes difficult to provide different investors with the benefits they can get in the early stages. The benefits can come quite late in the process, so it's hard to engage with investors and convince them early on.

Is there a useful framework for measuring the level of maturity in terms of digitalisation, not for individual assets but from the systems perspective (e.g., a framework that resembles the L1-L5 levels for autonomous vehicles)?

There should be a way to measure the benefits quantitatively. We have a data ontology, and we have the sensors, but how to measure the level of maturity in terms of digitalisation is an open question. So there is still a problem with maturity. Below there is an example of the levels of maturity (DIKW pyramid), considered by the participants as a good example of the relationships among data, information, knowledge, and wisdom and the levels of maturity. Organisations should be able to convert data into information, knowledge, and finally, wisdom to help stakeholders to make better decisions.

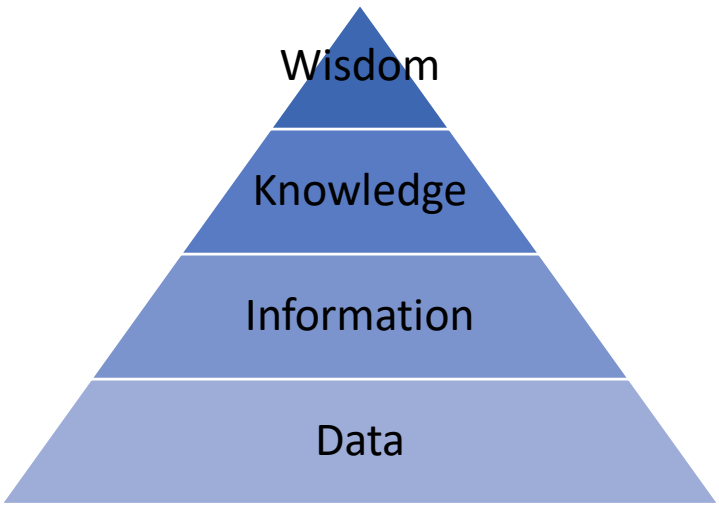

Figure 1. DIKW pyramid

3. **What makes infrastructure and civil engineering applications of digitalisation and data science challenging, in practice, compared to other industries? What are the main ideas, lessons learnt, best practices, etc. that can be incorporated in the industry from other areas? You may consider availability/size of datasets, data sharing, interpretability, expertise, uncertainty in materials or behaviour, complex environmental factors, integration with other systems, uncertainty affecting investments, etc.**

- **A lot of data collection, but uncertain benefits:**

Data exist, but how we can use it, and for which purpose? As a real example, here at the University of Cambridge, there is a building where we sense everything, but we do not know what data exist and how we can use these data. Perhaps, there was no plan at all on how to use these data and how to demonstrate their value. Here we need to include the Systems thinking approach to include, outcomes, threats, and benefits. And how we can communicate these to the people. It is managing the process of data-driven solutions that use systems thinking.

- **Learn from previous experiences:**

Lessons not to do - a document to share with all concerned: That document tells us what not to do. For example, projects where they did not use instrumentation in tunnels in the 80s. What sections of the tunnel do you want to replace, channel by channel? You need to close the train tracks in two ways. So, we've learned from that and now have some segmentation in the tunnels to allow recovery teams to minimize issues -> minimize and prevent tunnel closures.

- **We need to look at the asset as a system. Having the knowledge to predict and prevent:**

The supply chain problem, since 2012, is the use of technologies that exist in a laptop. Nobody is going to use it. There is no trust that defines the data you need to use. Many consultants have experience in the field of instrumentation monitoring but in other arenas. So, they don't know what to do in this context (i.e. physical infrastructure). Some assets are designed not to be replaced and it is very expensive to do so.

Table 3 (including flipchart) Fiorella

Fraser Perceval-Jacobs, Fiorella Dell'Olio CSIC, Mark Enzer-Motts, Tim Embley-Costain, Anne-Marie Friel-Pinsent Masons, Adam Box-Topcon, Farhad Huseynov-CSIC, Haitao Lan-CSIC, Jill Campion-UC Land Economy, Ajith Parlikad-CSIC

Fraser Perceval (Jacobs), Fiorella Dell'Olio (CSIC), Mark Enzer (Motts MacDonald) , Tim Embley (Costain) , Anne-Marie Friel (Pinsent Masons), Adam Box (Topcon) , Farhad Huseynov (CSIC) Haitao Lan (CSIC) Ajith Parlikad (CSIC)

1. Data into knowledge – actionable insight
    - The answer is basically No!
    - Market by market

Examples of collecting data and responding in real time

- E.g.: Newcastle Council responding to disaster
    - event ← need ← collect data ← to correct
    - setting a plan to collect data
- 'Cart & Horse' ← lots of data ← to what purpose? ← need to start with purpose
    - Decisions we need to make lots of data
    - How do we decide what to use?
- E.g.: Aviation ← data ← aerospace manufacturing
    - Data we might be able to get
    - Do what you can with data
    - Matchmaking data
- Open to mind around the vast of data we will have in future
- What it could be ← top down ← bottom up; decisions what data – to do […]

- We will find we can do more and request for more data
- Data is a 'learning process'
- We will have more capacity for data – insight + knowledge = wisdom
- Need to focus on skills and learning around data
- Starting with 0 data? If we start with some data we will try to make our decisions based on some data. We will be able to identify the gaps. Otherwise everything is the gap.
- There's a challenge of having access to the data, especially in big organizations. There's a Data button somewhere in the folder, which only few people have access to. So data sharing sources in order to apply these data science tools, advanced tools in order to use data for better data informed decisions.

2. Data
    - Data quality is a massive issue: when we talk about data quality data we need to relate it to the purpose ← what can data be used for?
    - 'Data as radioactive gold': do not use it for the wrong purpose.
    - Security overrides everything ← decisions i.e. △ flow of motorway ← impact
        - How free? Who is using data?
        - Films provide examples of what could happen
    - 'Data for the Public Good' – commercialise data
    - Machine Learning ← needs big data sets. It is already happening for sharing projects in the health system: Farmer AI buy from public sector ← what's the model?
    - Long term collaborative relationship where the public provides data and they eventually get something in return.
    - Models ← accountable ← public good should be part of transaction ← cyber security
    - Farmer ← access to patient data ← what will you give back?
    - Data sharing – interoperability: if you can't share data you will inevitably build restrictions into the system
    - It is important to start putting a proper value on data, even at the expense of getting off balance sheets. We need to get to the point where People Boards genuinely care about data. Often, people in boardrooms only start to care about things when there is a downside, so that there's a problem.
    - Data as an asset in particular during risk management processes. Data can talk about the potential risks, and it is important to know how to interpret those data. Risk is uncertainty and the information that data can provide will overcome that uncertainty.
    - If you're valuing information, and you're recognizing the risks associated with it, these are directly connected things. More information means less risk.
    - Data silo can be a real challenge: Huge software companies like Google, they are only interested in owning that particular environment and own the customer. They are mainly interested in commercial gain.
    - The problem of the legacy of data that can no longer give you the system it was generated on.
    - public good v commercial gain ←exploitation of data
    - Political power ← covid passport ← Google ← do it for free V BoE £1 bn
    - Trust & culture ← how do the upper management trusts these models

- Banking industry trusting the algorithms of 2 billion transactions
- How should organizations change? Bank ← AI ← exchange
- Do we have the right skills? Do we have the skills to develop the right system?
- AI presents its problems
- Black book ← don't know how it works Lots of people will receive an answer to an AI question and trust the response
- Barriers ← human and org factors: One of the biggest barriers/challenges and maybe also enabler will be human and organizational factors - the technical stuff is relatively easy- getting organizations to actually adopt it is another thing. Socio-technical challenge: getting the culture of an organization changed, so that people who make decisions will make decisions based on the data given to them rather than based on their gut instinct.
- Many organizations won't have the skill to build their own train their own models create their own system.
- Data is placed in the middle between AI and human analysts. And the question is what kind of education we need for the next generation? More AI schools?

3. Org change examples
    - Covid – remote working has driven to an online environment.
    - Time sheets ← power BI ← time booking automation.
    - Resourcing of people on projects is driven by advantage← we can limit our clients
    - Power BI ← aggregation ← Digital Twin
    - Complexity of data: DT Aggregation Future
    - Looking at the big picture ← industry has not changed compared to other industries i.e., sport; aerospace. There must be a reason for this. The main one is the way it is structured: It is fragmented.
    - If there's going to be an effective digital transformation in the industry, it needs to be more joined-up. It can't be that everyone does their own digital transformation in their own way. As an example, manufacturing is fragmented
    - Civil do 'not have master' ← project-based industry
    - Consulting engineers ← all projects different ← industry / geography
    - Diversity of location not an issue
    - Requirements ← learn what each client is asking for ← consultancy ← £ by hour

4. Interview questions
    - (1) Who is target audience?
    - (2) It looks like a range people
    - (3) Data science / specialists
    - (4) Legal ← data management team ← technical aspects ← need to discuss with people who use the data
    - (5) What is the purpose of questions?
        - It looks like data into detail
        - Digital PMO
        - Not just data
        - Bigger picture
        - Where is this going
    - (6) Need to check bigger research questions

- (7) Product Delivery
- (8) System of assets
- (9) Organisation ← digital transition
    - [lawyers] / investors
    - big ask
- (10) Organizations striving to reach maturity: time and opportunity to build maturity
- (11) Tried with BIM ← accreditation on 1 project
- (12) Have you got money in budget to deal with BIM ← longer time for digital design
- (13) Tendering ← increase cost of BIM
    - BIM should reduce costs
    - Price for first phase ← detailed design before construction
    - Demand designers to modeling
    - Output ← digital rep of the building on delivery of the building
    - Do engineers know what the right specifications are?
    - No ← how can you ask for digital model if you can't specify
- (14) The loop of design operation.
    - Government clients who are not ready to specify BIM are not ready for data in operation.

Table 4 (including flipchart)

Carlos Laguna Sanchez-Motts, Shelley Arora-Tailby CSIC, John Pelton-Costain, Olly Wright-Aviva, Sharon Duffy-Thames Water, Keith Bowers-COWI, John Allum-UK Parliament, Manu Sasidharan-CSIC, Jennifer Schooling CSIC

Given the increasing presence of data in infrastructure applications, in your opinion, to what extent is the industry transforming this data into knowledge – providing actionable insights and informing decision-making?

SD -depends on where you are on digital maturity scale of the organisation. TFL less digital mature than my current company.

Depends on the infrastructure sector. Water companies are collecting large amounts of data. There is disparity across the sectors.

CLS – data ownership is an issue in the water industry.

JP – Rolls Royce have sensors in all their engines around the worlds and control them from Derby. Very digital mature

SD – need to install monitors. From an asset management point of view. It is not consistent across the industry.

Also depends upon the type of infrastructure – legacy vs new.

KB – Nat. Highways issues dealing with old legacy infrastructure. User behaviours hard to anticipate. System only as good as the weakest part.

At the present time, what are the practical constraints preventing the civil engineering industry from achieving better solutions that are driven by data, given the current maturity of sensing technologies and data science methods? Do the current main challenges relate to the amount or quality of dat

being collected, the data ownership/availability, or the methods to make use of these data? Is there a useful framework for measuring the level of maturity in terms of digitalisation, not for individual assets but from the systems perspective (e.g., a framework that resembles the L1-L5levels of autonomous vehicles)?

Challenges for the industry to derive insights from data. Bridging the gap between data and knowledge.

JP – constraint in the question – don't know what the better solutions are.

JS – in order to get better solution need long term view. Clients delegate innovation to the supply chain. But they do not have their skin not in the long game – they do what they are delegated. Innovation needs ownership at the client level to drive the supply chain.

Value of the innovation needs to link societal benefits to the business cases for collecting/analysing/using the data driven solutions. It is a regulatory challenge.

OW – there needs to be a link between the people collecting the data and then the person with the purse strings. Those spending need to have greater buy in to the data.

Cultural/organisational aspects contributing to the gaps in (i) skills appreciation (ii) data analysis skills.

KB – people with skills are not driving the culture. Skills appreciation is not currently there but it is necessary.

JP – who pays for the cost of lack of resilience? There is no funding available to do a lot of this stuff. How do you get the right people to understand the value of the data in the system and the increase/decrease in resilience that accrues from such an approach/strategy.

JS – need to change our way of thinking about physical assets to being digital assets. Need to appreciate and consider the 'sensors' are also assets and it is not just the infrastructure that is an asset. Data is an asset in its own right which needs to be managed.

JS - Thames tideway – they claimed they were designing an asset that requires no maintenance. They did not put in monitoring sensors. They are retrospectively putting in monitoring sensors. There is no such thing as a maintenance free asset, unless it is disposable.

SD – data is an asset in its own right which needs monitoring and management which changes with increase in configuration management. Links back to culture and people skills.

JS – if you know you need data that can be specified early on and the systems can be specified.

OW – cyber-security. System needs to self- check that it is not being manipulated and highlight if there are issues and create multiple check-points within the system. Risk management.

JS-Needs firebreaks built in.

JP – cant have single point of failure systems in place

JS – use your systems to monitor things rather than operate things. Levels of digitisations.

KB – safety case built into assets is basic stuff now. Very attractive on one hand. But it is very tricky to do where different organisations operate different companies.

SD – it's about regulation – mandatory requirements necessary with a fail-safe. A safety case is necessary for the system where there are digital inter-dependencies between the multiple systems that interact with each other. E.g., self-driving autonomous vehicles and smart motorways.

JP – Data protection – two sides of the coin. Making some data from 'social agents' available for infrastructure monitoring or operational management. Although GDPR issues in using data from mobile phones.

CLS – ML data models work as a black box. There is a trust constraint.

Explanations of the ML model to be improved, this would improve trust.

MS – Can we augment existing ML models to make better insights?

JS – background data can bias the algorithms.

JS – a consistent set of data architecture is required across models and algorithms. This is one of the worries of ML.

JS - Governance and government need to be separate to remove bias.

What makes infrastructure and civil engineering applications of digitalisation and data science challenging, in practice, compared to other industries? What are the main ideas, lessons learnt, best practices, etc that can be incorporated in the industry from other areas? You may consider availability/size of dataset, data sharing, interpretability, expertise, uncertainty in materials or behaviour, complex environmental factors, integration with other systems, uncertainty affecting investments, etc.

General thought from table - Be mindful not to lead the interviewee with bullet points in the question.

SD - Question 3 interesting.

JS – difficult to recruit and **retain** data scientists

JP – aging work force - majority in industry over 50…may be holding back the effort.

Table 5 (including flipchart)

David Simavorian-Accenture, Dee Dee Frawley CSIC, Peter Hewitt-LOR, Mehdi Alhaddad-TfL, Alejandra Masia-BP, Nicky De Battista-Epsimon/CSIC, Chiho Jeon-CSIC/CAU, Paul Fidler-CSIC, Sakthy Selvakumaran-CSIC, Scott Steadman, BSI, Janet Lees, CSIC

With infrastructure or buildings, half the half the battle is just trying to figure out which bits you're supposed to be looking at. Where do you monitor? What data do you need? For example, when instrumenting a new bridge, how do you identify which bit of the bridge is actually going to fail and where knowing that the failure is unlikely to occur for 50 or 100 years? You can end up collecting a lot of data you are never going to use.

- Network Rail have had incredible results relating satellite data used to measure soil moisture with the number of slope failures that have happened over the last 10 years. There no specific answer but using a statistical approach, decisions can be made on where to focus efforts based on the

probability of failure in a specific region. This can also help focus operational decisions such as deciding to slow down trains is areas of high soil moisture. We are as an industry starting to make use of this data intelligently, but the example above at Network Rail hasn't been rolled out to the operators to actually apply it.

- There are also many examples of bad practice. Often sensors are included and data is collected but there's no follow up. No plan to get the most value out of the data during construction or after, no plan to properly store the data. Those at the coalface often do not know why data is being collected, what it will be used for, nor do they understand the value of the data. Often in civil engineering these questions are not well designed or understood so operatives collecting the data are then less likely to value it themselves. Also, it's hard to convince the client to pay for something when they don't know what value it's going to deliver.

- Industries like oil and gas and if we go even up higher up in the in the maturity levels, automotive and aerospace, are much more advanced. The difference is that in those industries the questions are very well defined - it's much more constrained. Whereas in the civil infrastructure industry, the challenge is much bigger because it's always a unique structure etcetera. Also, there is an issue in the construction industry of a lack of willingness or much desire to do things better except in small areas like Network Rail for landslides because of accidents like the Stonehaven derailment.

- These large InSAR data sets are available only through paying satellite companies for the data. Large numbers of data sets will need to be available so that people can start doing interesting things like the example above of looking at soil moisture versus failures. Is the Crossrail data publicly available? A lot of research is stifled because lack of transparency and availability of these large data sets .

So what are kind of the key constraints preventing us from actually better results using?

Technology is not a constraint. It's the use of the technology and the issue above about discoverability and availability of data.

The breakthroughs come when trying to combine different sets of data together and at the moment often we don't know what those what those connections are. There's absolutely no chance that one can predict accurately what's likely to happen, maybe two years ahead of where we are, much less 10 years. (P Hewitt LOR) Combining different types of data and using data that perhaps wasn't meant for that purpose and using clever data processing methodologies like at the Alan Turing Institute you can get much more from than the sum of the total.

Procurement and risk is a constraint. (P Hewitt) We find that we have to rebuild the 3D model on most projects because it is an architect model with trees but not with concrete tolerances etc. Consultants won't carry any responsibility for the model they have generated which is another part of the reason we have to redo it. Where the risk lies and how the project is procured to manage the risk is key.

Engineers can end up using data for decision making without really understanding where it's coming from, how it's being processed etc. From the design side in a lot of cases you have massive software capability and you are facing kind of black boxes where a lot is going on inside but you only get the result. Unless you have a really have a good understanding, you are not able to detect any error bugs within that block.

Some AI driven models really are black boxes - inside it's just a bunch of coefficients to that matrix into that matrix. If someone had to stand up in court and explain to a judge why a particular thing was not recognised as a crack, they would struggle to do so with any confidence.

The questions seems to be for someone who works exclusively in data science? A structural engineer might find these hard to answer. The questions are too leading, open them more. Suggest changing to:

Q1: Start with 'What does 'data science' mean to you and in your line of work?'

What is your background in data?

Q2: Start with 'Are you using machine learning?'

Q3: - This is only relevant to people involved in recruitment.

But the skills piece is really interesting. Are engineers getting training in data science or like many of us just having to pick it up when you need to?

Main question Category - Q1 does not belong in this category as it's not relevant to 'ownership and governance'

Main question Category 3: Q2: The suggestion is 'either/or' , do you think data science might be more business in the future in the near future?

AI will definitely disrupt the labor market. Lots of jobs will be replaced by AI, but lots of other jobs will be created requiring new skills.

- So perhaps the questions would be 'How do you think data science will affect the labour market?'

- Generally the questions probably could be more open and less leading.

Q1, How do you think data science will affect your business? Your business might not be disrupted or change but it might be affected in a different way.

Ending Question:
The obvious answer to that question is that there is going to be more of a multidisciplinary focus in the future, so it's good to have a solid grounding in both.

Table 6- Flipchart

Lizzy Moyce-Arup, Wei Bi CSIC, Xiaomin Xu CSIC, Zaid Rawi-BP, Chris Barker-Arup, Paul Campion-TRL, Mike Spencer-IMIA, Nevena Vajdic, DRF, Kristen McAskill-CSIC, Jason Sun-CSIC

- Most of the digitalisation and data-driven work goes to the new infrastructures. How about the retrofit of the ageing ones?
- At the moment, the digitalisation and data-driven work are still at the very beginning stage. Standards and guidelines are needed to make sure this is delivered. There is a lack of handover from project planning to operations. For example, National Highways has built a few digital models at the planning stage, however these were thrown away when they came into the operation.
- There are still quite a number of problems in sharing data, e.g. data quality, data source, cost of data, how to handle data, data security… And there is a lack of leadership in this aspect. We need industry leadership to set up the market, and make the business model. For example, to share data with civil engineer, the clients would be very cautious and difficult, typically through a third-party to filter some of the information.

- In the academic environment, sharing data is also very difficult. People are scared of sharing some of the data, which might need to involve with liability or lawyer.
- The value of data. Data needs to be transformed into knowledge, otherwise you will just lose data and also the knowledge. In most cases, people obtain data and manage data to enhance their competitive commercial edge. And this requires people to understand the value of their data, to provide additional service…and may eventually change their culture, e.g. pay by mile.
- At the moment, there are some ongoing specifications which we can refer to: Capital facility information handover specification (CFIHS) and data exchange in the process industry (DEXPI DEXPI – Data Exchange in the Process Industry). Other example like aircraft accident, people learn from the lesson and experience and share the data extensively.
- Other issue about sharing data is that this needs to be between organizations, not just people, as people are moving from project to project. And that also make the sharing the data more difficult. Not everyone stay in one organization for 40 years.
- Data needs to inform the commercial activities and to maximize the business impact. For example, linking monitoring to insurance… And this also involves long-term data storage and management.