T5

Structured Data, Data Standards, and Interoperability

Table of Content

Introduction & Learning Outcomes

In this training unit, you will learn about structured data, data standards, and interoperability, each of which plays a crucial role in the development of the Qualifications and Credentials Platform. A Qualifications and Credentials Platform, such as the African Qualifications and Credentials Platform (QCP), is a digital infrastructure designed to manage, verify, and share educational qualifications and professional credentials. Understanding structured data, adhering to data standards, and ensuring interoperability are fundamental to the platform's success. These elements enable accurate data handling, seamless integration across various systems, and reliable verification of credentials, which are essential for the credibility and usability of the QCP.

Key Topics Covered

Structured data

Structured data is defined by its organised format, which allows for easy access and analysis. It facilitates efficient data management, enhances data accuracy, and supports effective data verification processes. In the context of a QCP, structured data is vital for maintaining clear and consistent records of educational qualifications and professional credentials.

Data standards

Data standards are predefined norms or requirements that ensure data consistency, accuracy, and quality across different systems. Understanding data standards involves recognising their significance in maintaining uniformity and reliability in data handling. In developing a QCP, adhering to data standards ensures that the data remains consistent and accurate, which is crucial for the platform's integrity.

Interoperability

Interoperability refers to the ability of different systems to work together seamlessly. It encompasses various aspects, including legal, organisational, semantic, and technical dimensions. The concept of interoperability is closely connected to structured data and data standards, as it relies on consistent and accurately formatted data to function effectively. In a QCP, interoperability is critical for ensuring that data from different sources can be integrated and verified without issues.

Learning Outcomes

By the end of this training module, you will:

1. Understand Structured Data

- Define structured data and identify its key features and benefits
- Explain the importance of structured data in managing and verifying educational and professional credentials

2. Understand Data Standards

- Describe what data standards are and their significance in ensuring data consistency, accuracy, and quality
- Identify key characteristics and considerations for developing effective data standards
- Illustrate the use of specific data standards within the educational sector

3. Familiarise yourself with Interoperability

- Explain the concept of interoperability and its different aspects (legal, organisational, semantic, technical)
- Understand how structured data and data standards facilitate interoperability

Structured Data

Understanding Structured Data

What is Structured Data?

Structured data refers to information organised in a predefined, consistent format, making it easily accessible, searchable, and analysable by both people and computers. Typically, it is arranged in tabular forms such as databases, spreadsheets, XML files, and CSV files. For example, a customer database might include rows of customer records with columns for name, address, and phone number. This systematic organisation allows for efficient data handling and retrieval.

Features of Structured Data

Definable Attributes: Structured data consists of clear and consistent attributes across all entries. For example, each grading record for a student will include the same fields, like the name of the student, the subject being graded, the date of grading, and the grades.

Relational Attributes: Different structured datasets can be linked through common fields. For instance, student application records can be connected to their grading records via student ID fields. That means that both datasets will include the field for the student ID (one per student), creating a relation or bridge between the two datasets. That allows you to store different structured data in relational databases.

Quantitative Nature: Structured data is inherently numerical or categorical, supporting straightforward mathematical and statistical analysis.

Examples of Structured Data

Databases: SQL databases, which use a schema to define data structures.

Spreadsheets: Excel or Google Sheets files, where data is organised in rows and columns.

CSV Files: Text files with comma-separated values, often used for data import or export.

Example of structured data

```
Title, Author, ISBN13, Pages
Animal Farm, George Orwell, 978-0451526342,144
Brave New World, Aldous Huxley, 978-0060929879,288
Fahrenheit 451, Ray Bradbury, 978-0345342966,208
Jane Eyre, Charlotte Brontë, 978-0142437209,532
Wuthering Heights, Emily Brontë, 978-0141439556,416
Agnes Grey, Anne Brontë, 978-1593083236,256
Walden, Henry David Thoreau, 978-1420922615,156
Walden Two, B. F. Skinner, 978-0872207783,301
"Eats, Shoots & Leaves", Lynne Truss, 978-1592400874,209
```

Benefits of Structured Data

What are some of the benefits of using structured data?

Consistency and Reliability

Structured data ensures uniformity and accuracy, reducing errors and discrepancies. For instance, when all entries in a "date" field adhere to a specified format (e.g., YYYY-MM-DD), this ensures consistency. When compiling or operating with the data in the ‘date’ fields, it is also easy to distinguish which numbers correspond to days, and which numbers correspond to the month, following the specified format.

Efficient Storage and Retrieval

The predefined format of structured data allows databases to store information efficiently and use indexing to speed up search operations, reducing the amount of data scanned during searches.

Enhanced Data Analysis and Visualisation

Structured data's organised format supports advanced analytical tools and techniques. It makes the generation of meaningful visualisations like charts and dashboards easier, helping us discover patterns and trends which can, in turn, inform our decision-making.

Examples of Structured Data in Practice

- **Academic Records:** Electronic academic records store student information in a structured format, enhancing data sharing between departments and improving administrative efficiency.
- **Student Databases:** Educational institutions can use structured student data for targeted academic support, and improved student relationship management.
- **Enrollment Management:** Structured data assists in managing student admissions, tracking demographic trends, and forecasting enrollment numbers to better plan for resource allocation and staffing.
- **Faculty Performance Analysis:** Educational institutions can collect structured data on teaching performance, student feedback, and peer evaluations to enhance teaching quality and professional development for educators.
- **Library Management:** Educational libraries use structured data to catalogue books, track borrowings, and manage digital resources, enhancing access to learning materials for students and faculty.
- **Web Analytics:** Structured data tracks and analyses website performance metrics like page views and user behaviour, aiding in optimising online presence.

To consider: Can you think of other use cases for structured data?

To consider: Can you think of examples where using structured data would not be useful?

Structured Data in the Development of a Qualifications and Credentials Platform

In developing a Qualifications and Credentials Platform, structured data plays a crucial role. The platform can efficiently manage and verify educational qualifications and professional credentials by using structured data. For example, it can store and retrieve standardised information about institutions, courses, grades, and certifications. This structured approach ensures data integrity, supports complex queries, and facilitates seamless integration with other systems, ultimately enhancing the platform's reliability and user experience.

Implementation examples:

European Qualifications Framework (EQF): The EQF[1] is a common European reference framework that links different countries' qualification systems and frameworks together. Structured data underpins the EQF by providing a standardised format for describing qualifications. This helps in making national qualifications more readable and understandable across Europe, thus promoting workers' and learners' mobility between countries.

Europass: Europass[2] is another EU initiative that leverages structured data to enhance the transparency of skills and qualifications across Europe.

- One example is the **European Digital Credentials Infrastructure (EDCI):** The EDCI[3] uses structured data to issue, share, and verify digital credentials. This ensures that qualifications are easily accessible and verifiable across Europe, reducing fraud and increasing trust in credentials.
- **Standardised Formats:** Europass documents, including the Europass CV and Language

[1] *The European Qualifications Framework (EQF) | Europass.* https://europass.europa.eu/en/europass-digital-tools/european-qualifications-framework. Accessed 30 July 2024.
[2] *Home | Europass*. https://europass.europa.eu/en. Accessed 30 July 2024.ualifications-framework. Accessed 30 July 2024.
[3] *European Digital Credentials for Learning- Introduction to Digital Credentials | Europass.*. https://europass.europa.eu/en/stakeholders/european-digital-credentials. Accessed 30 July 2024.

Passport, use standardised formats to record personal skills and qualifications, making it easier for employers and educational institutions to understand and compare credentials from different countries.

Credential Engine (USA): Credential Engine[4] is a nonprofit organisation in the United States that aims to create an open, transparent credentialing system. Their credential registry uses structured data to store detailed information about credentials, including the skills, competencies, and learning outcomes associated with each credential. This allows for better comparability and understanding of credentials.

Linked Data

Linked data refers to a method of publishing structured data in a way that allows it to be interlinked and become more useful. By using standard web technologies such as HTTP, RDF (Resource Description Framework), and URIs (Uniform Resource Identifiers), linked data enables educational resources to be connected and shared across the web, creating a more interconnected and accessible web of educational data.

Benefits of Linked Data in Education

- **Interoperability:** Linked data facilitates seamless data exchange and integration between different educational institutions, learning management systems, and digital libraries.
- **Discoverability:** Interlinked educational resources can be more easily discovered and accessed by students, educators, and researchers, leading to enhanced learning and research opportunities.
- **Data Enrichment:** Linking educational resources from various sources can enrich the original data set, providing more context and depth to the information.
- **Reusability:** Standardised data formats and structures make it easier to reuse educational resources for different purposes, enhancing their value over time.

Consider a scenario where linked data is used to connect educational resources across different universities. Each course, textbook, and research paper has a unique URI, and its metadata (such as authors, topics, and prerequisites) is linked across various educational platforms. This creates a comprehensive and accessible educational resource network that can be queried and analysed to improve curriculum development, student learning paths, and academic research.

FIND OUT MORE

Learn more about how linked data was formed and the idea behind it from the inventor of the World Wide Web in the video presented.

Tim Berners-Lee:
The next Web of open, Linked data[5]

[4] *Home | Credential Engine.* https://credentialengine.org/. Accesed 11 July 2024.
[5] TED. *Tim Berners-Lee: The next Web of Open, Linked Data.* 2009. YouTube, https://www.youtube.com/watch?v=OM6XIICm_qo.

INDIVIDUAL & GROUP ACTIVITY

Training Activity

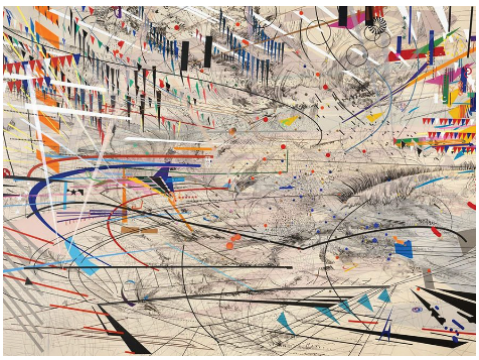
Objective of **Developing a Data Model for the Qualifications and Credentials Platform**: Understanding structured data and creating a structured data model for managing educational qualifications.

Part 1 (10-15 min)

Below you have three pieces of data. Describe each piece of data.

Piece of Data A[6]

Piece of Data B[8]



Julie Mehretu – Stadia I (2004)

Piece of Data C[7]

The Peace of Wild Things

Wendell Berry

When despair for the world grows in me
and I wake in the night at the least sound
in fear of what my life and my children's lives may be,
I go and lie down where the wood drake
rests in his beauty on the water, and the great heron feeds.
I come into the peace of wild things
who do not tax their lives with forethought
of grief. I come into the presence of still water.
And I feel above me the day-blind stars
waiting with their light. For a time
I rest in the grace of the world, and am free.

After describing each piece of data, think about the types of information you have collected about the different pieces of data. Have you been describing the size, age, language, authorship, medium of expression? See if there are any common categories across the different pieces of data. How easy or difficult is it to find common categories to connect or sort the descriptions of the three different pieces of data?

If there is time, report back to the group and discuss with peers.

[6] 'Cat'. *Wikipedia, 29 July 2024. https://en.wikipedia.org/w/index.php?title=Cat&oldid=1237407687.*
[7] 'The Peace of Wild Things by Wendell Berry'. *Scottish Poetry Library,* https://www.scottishpoetrylibrary.org.uk/poem/peace-wild-things-0/. Accessed 30 July 2024.
[8] 'Stadia I'. *SFMOMA*, https://www.sfmoma.org/artwork/2006.7/. Accessed 30 July 2024.

Part 2 (30 minutes + presentation time)

Form groups of 4-6, and before starting the activity select someone who will in the end present your results to the other groups. Try to form a group with learners coming from a different educational context from your own. This can mean they represent a different country, or they engage with education in a different capacity (e.g., policymakers vs representatives of academic institutions).

Within the group, brainstorm and outline the **key categories of data or data elements** you would need for a Qualifications and Credentials Platform, considering your institutional context (e.g., institutional information, course details, date of issuing credentials).

In your group, discuss which of these elements should be mandatory to have for a functional QCP, and which should be additional 'nice to have' elements. Consider, for example, which elements can be used across the different educational contexts your team members are coming from, and which are specific to the educational context you represent.

Consider and discuss with your group where the data for each of the key elements would be coming from, e.g. how would you collect the data? Would any of the data have to be treated confidentially? Who in your institution would be allowed to see confidential data? How does your institution store different data at the moment?

Prepare the selected categories, including examples of each of them in a structured format, using an appropriate format (look at the learning materials for examples of appropriate formats).

Present your model to other groups, sharing which elements you think should be mandatory to collect, and where the data for these elements would be coming from.

Data Standards

Defining Data Standards

Consider a student who is moving from a school in one country to a school in another country which follows a different school system. The school records of the student need to be shared and accessed by the new school to determine where to place the student in this new educational environment. If each of the schools is storing school records of the students in a different way, student data might be difficult to translate to another system, some information might be lost when translating the student's record to a new system, and data might be missing which is wanted in the new school but has not been recorded in the student's old school. The new school might therefore lack the full learning history of a student needed to make the correct decision of where to place the student. School staff might need to spend more time to bring all the data together manually, and interpret the data provided in different formats. This might delay the decision for the student and affect their learning journey. In this case, data standards could provide a common framework for recording, exchanging, and interpreting student data making communication between the two schools smoother and faster.

But what are data standards? Data standards are agreed-upon conventions and guidelines for formatting and handling data. Some well-known examples include XML, JSON, CEDS (Common Education Data Standards), SWIFT and SCORM

Benefits of Data Standards

Interoperability
Different systems and applications can understand and use data consistently. This facilitates better data exchange, integration, and communication across platforms, organisations, and systems.

Consistency and Accuracy
Standardised data formats and definitions ensure that data is captured, stored, and processed uniformly. This reduces errors in data handling, ensuring accuracy and reliability.

Data Quality and Integrity
Data standards provide guidelines for data validation, ensuring that the data meets defined criteria. This enhances the quality and integrity of the data, making it more reliable and trustworthy.

Characteristics of Effective Data Standards

The benefits might sound similar to how we were describing structured data, however, data standards have some specific characteristics:

Widely Used: A standard must be widely adopted and used consistently to be effective. For example, barcodes and QR codes are scanned millions of times daily worldwide, illustrating their widespread use. If a data standard is not used much, it is also not very useful. The two schools in our example need to both use the same data standards to be able to communicate with each other effectively.

Supported by a Competitive Marketplace: Effective standards are maintained by organisations that ensure their accessibility and low costs. This support encourages innovation and keeps the cost of implementation and usage low.

Open and Non-proprietary: Standards should be freely available to encourage rapid market adoption. Open standards are safer and more reliable for regulatory adoption as they are not tied to commercial interests.

Machine-Readable: Data should be in a format that computers can process without human intervention. This ensures accuracy and efficiency in data handling, reducing the need for manual data entry.

Taxonomy and Ontology: Taxonomy classifies information or sub-categories, providing a hierarchical structure. Ontology defines concepts and relationships within a domain, ensuring structured and consistent data representation. Together, they establish an organised system that clarifies definitions and relationships among data.

Considerations when Building Data Standards

- **Modelling:** Modelling involves transforming real-life examples into structured data records. This process includes making decisions about what details to include and ensuring that different contexts are captured. This is best done by having a diverse and representative group of stakeholders collaborate on providing real-life examples.
- **Representation and Interoperability:** Representation involves deciding on file formats, date formats, and how to store images, considering who will use the data standards. The goal is that standardised data can be exchanged between different systems, maintaining consistency and usability.

Data Standards in Education

Examples of data standards used in education include:

- **CEDS (Common Education Data Standards)[9]:** Many US states use CEDS to align their data systems. This alignment enables consistent data reporting to federal agencies, ensuring educational data is uniform and accurate.
- **SCORM (Sharable Content Object Reference Model)[10]:** SCORM is a set of technical standards for e-learning software products. It tells programmers how to write their code so that it works with other e-learning software, ensuring compatibility and interoperability.
- **ISCED 2011[11]:** ISCED 2011 is a UNESCO-maintained standard that describes educational levels. It allows for the comparison of educational attainment across different countries.
- **ISCED-F 2013[12]:** ISCED-F 2013 is a UNESCO-maintained standard that defines subject matters. This standard allows for the comparison of education elements internationally.
- **MLO (Metadata for Learning Objects - Advertising)[13]:** MLO is a European standard that specifies the characteristics of the electronic representation of learning opportunities. It makes it easier for prospective learners to find educational opportunities.
- **EuroLMAI (European Learner Mobility - Achievement Information)[14]:** EuroLMAI is an EU standard that specifies the model for recording and exchanging achievement information. It supports the mobility of students and the recognition of their achievements across different institutions.

To consider: Can you think of other examples of data standards within education or other fields you are familiar with?

[9] *Common Education Data Standards (CEDS).* https://ceds.ed.gov/. Accessed 30 July 2024.
[10] 'SCORM.Com'. *SCORM.Com*, https://scorm.com/. Accessed 30 July 2024.
[11] UNESCO Institute for Statistics. International Standard Classification of Education: ISCED 2011. UNESCO Institute for Statistics, 2012.
[12] UNESCO Institute for Statistics. International Standard Classification of Education: Fields of Education and Training 2013 (ISCED-F 2013) Detailed Field Descriptions. UNESCO Institute for Statistics, 2015. DOI.org (Crossref), https://doi.org/10.15220/978-92-9189-179-5-en.
[13] *Metadata for Learning Opportunities (MLO) - Advertising | Joinup.* 1 Dec. 2008, https://joinup.ec.europa.eu/collection/european-committee-standardization-cen/solution/metadata-learning-opportunities-mlo-advertising.
[14] *European Learner Mobility Achievement Information (EuroLMAI) | Joinup.* 1 June 2010, https://joinup.ec.europa.eu/collection/european-committee-standardization-cen/solution/european-learner-mobility-achievement-information-eurolmai

Data Standards in the Development of a Qualifications and Credentials Platform

A Qualifications and Credentials Platform such as the African QCP can streamline the management, verification, and sharing of educational achievements, skills, and professional credentials. Data standards play a crucial role in the development and functionality of such a platform:

- **Interoperability and Integration:** Data standards ensure that different educational institutions, certification bodies, and employers can share and verify credentials seamlessly. This capability enables the platform to integrate with various learning management systems (LMS), student information systems (SIS), and other relevant databases.
- **Consistency and Accuracy:** Standardised formats for credentials ensure that information is captured and presented uniformly. This uniformity reduces discrepancies and errors, ensuring that credentials are accurately represented and recognised across different institutions and employers.
- **Data Quality and Integrity:** Data standards provide validation rules and criteria that ensure the quality and authenticity of credentials. These standards help prevent fraud and misrepresentation of qualifications, maintaining the platform's credibility.
- **Scalability and Efficiency:** Standardised data formats enable the platform to handle a large volume of credentials efficiently. This scalability facilitates the rapid processing and verification of credentials, enhancing the user experience.
- **Global Recognition:** International data standards, such as ISCED, ensure that credentials are recognised and comparable across different countries. This supports student mobility and international employment opportunities by providing a common framework for credential evaluation.
- **Security and Privacy:** Data standards include guidelines for secure data transmission and storage. These guidelines ensure that sensitive information, such as personal identification and academic records, is protected from unauthorised access and breaches.

By leveraging data standards, a Qualifications and Credentials Platform can provide a reliable, efficient, and secure way to manage and verify educational and professional achievements. This approach benefits learners, educational institutions, and employers alike, fostering trust and facilitating career development and educational opportunities.

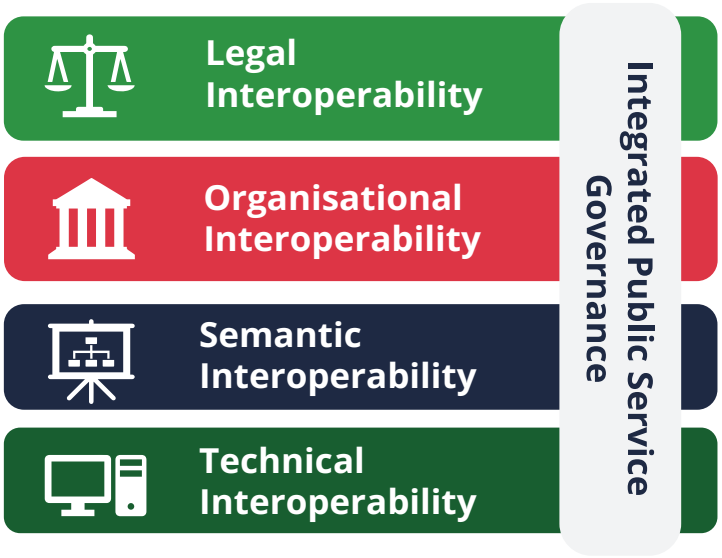
Interoperability

Introduction to Interoperability

In the field of education, interoperability refers to the ability of different educational systems, technologies, and institutions to work together seamlessly.

Interoperability of different systems is a rather complex issue with several relevant dimensions that need consideration when wanting to ensure the exchange of data. When looking for these dimensions, the European Interoperability Framework is a great reference, as it offers a helpful division of interoperability into four main aspects: legal interoperability, organisational interoperability, semantic interoperability, and technical interoperability.

Legal interoperability describes the policies and legal texts that support cross-border cooperation, the free flow of information, and innovation between different actors across states. **Organisational interoperability** involves the processes put in place to allow specific organisations to work as a network. Semantic interoperability pertains to vocabularies and ontologies that enable concepts and terms to be described in a standardised manner. **Technical interoperability** concerns the technical services, such as software and hardware, that are implemented to support interoperability[15].



[15] European Commission. Directorate General for Informatics. New European Interoperability Framework: Promoting Seamless Services and Data Flows for European Public Administrations. Publications Office, 2017. DOI.org (CSL JSON), https://data.europa.eu/doi/10.2799/78681.

Connection to Structured Data and Data Standards

Interoperability is closely tied to structured data and data standards. As you have learned above, structured data refers to data that is organised and formatted consistently, making it easily searchable and usable. Data standards provide the agreed-upon formats and protocols for structuring data, ensuring that different systems can understand and exchange information seamlessly. Interoperability relies on these standards to facilitate the accurate and efficient exchange of data between disparate systems.

Example

In the educational sector, the Systems Interoperability Framework (SIF)[16] is a widely adopted data-sharing open specification that links data systems within schools in the United States, Canada, the UK, Australia, and New Zealand. SIF uses a common data model designed specifically for educational needs, enabling different applications and systems to exchange data efficiently and securely. By adhering to such standards, educational institutions can ensure that their systems are interoperable, promoting better data mobility and accessibility.

Importance in Developing a Qualifications and Credentials Platform

Interoperability is crucial in the process of developing a qualifications and credentials platform. A platform that manages qualifications and credentials must integrate data from various educational institutions, each potentially using different systems and data formats. Without interoperability, this integration would be cumbersome, inefficient, and prone to errors. Interoperability ensures that data from different sources can be standardised and seamlessly exchanged, facilitating the recognition and verification of qualifications and credentials across institutions and borders.

Furthermore, interoperability enhances the user experience by allowing students, educators, and employers to access and share information easily and securely. For students, it enables the seamless transfer of their academic records and credentials between institutions, supporting mobility and lifelong learning. For educators, it allows for the efficient sharing of resources and best practices, fostering collaboration and innovation. For employers, it provides a transparent and reliable way to verify the qualifications and skills of job candidates

[16] *SIF Implementation Specification (Australia) 3.6.2*. http://specification.sifassociation.org/Implementation/AU/3.6.2/. Accessed 30 July 2024.

GROUP ACTIVITY

Training Activity

Objective of **Tackling Challenges to Interoperability with Data Standards**: To help learners understand the importance and application of data standards; To discuss and propose solutions to common interoperability challenges in educational systems.

Instructions (50-60 min + presentation time)

Form groups of 4-6 learners, and before starting the activity select someone who will in the end present your results to the other groups. Try to form a group with learners coming from a different educational context from your own. This can mean they represent a different country, or they engage with education in a different capacity (e.g., policy makers vs. representatives of academic institutions).

Spend a few minutes individually thinking of common interoperability challenges faced by the educational system in your country.

Then re-group and list all the challenges you came up with in your group.

Together select one challenge to focus on (this can be the most interesting one that comes up, or the one most frequently named interoperability challenge).

Together think of a data standard that could tackle the interoperability challenge you have selected. What is the focus and topic of the data standard? What elements would it need to cover? You can help yourself by searching for existing data standards that could address your interoperability challenge, and take inspiration from there, but do consider your specific context. For example, if the interoperability challenge you have selected is about the different grading systems across educational institutions in neighbouring countries, could you help yourself with existing data standards which support interoperability of grading systems? Is there anything that makes it difficult to create a data standard for the challenge you are tackling?

Create a presentation summarising the key points of your data standard - the focus of the standard and the problem it addresses, the key elements, the main difficulties you have encountered during the exercise etc.

Present your findings to other groups.

If time allows, open a discussion on the feasibility and impact of the proposed solutions within the context the participants are familiar with.