Radioactive Waste Management and Decommissioning
NEA/RWM/R(2022)5
October 2026
www.oecd-nea.org

NEA Workshop on Digital Safety Case Methods and Development

Summary Report
25–26 October 2022
Berlin, Germany

OECD
BETTER POLICIES FOR BETTER LIVES

NEA
NUCLEAR ENERGY AGENCY

Nuclear Energy Agency

NEA/RWM/R(2022)5

Unclassified

English text only

26 October 2026

NUCLEAR ENERGY AGENCY

Cancels & replaces the same document of 22 October 2026

Radioactive Waste Management Committee

NEA Workshop on Digital Safety Case Methods and Development

Summary Report

25-26 October 2022
Berlin, Germany

JT03596437

This document, as well as any data and map included herein, are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT

The OECD is a unique forum where the governments of 38 democracies work together to address the economic, social and environmental challenges of globalisation. The OECD is also at the forefront of efforts to understand and to help governments respond to new developments and concerns, such as corporate governance, the information economy and the challenges of an ageing population. The Organisation provides a setting where governments can compare policy experiences, seek answers to common problems, identify good practice and work to co-ordinate domestic and international policies.

The OECD Member countries are: Australia, Austria, Belgium, Canada, Chile, Colombia, Costa Rica, Czechia, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Israel, Italy, Japan, Korea, Latvia, Lithuania, Luxembourg, Mexico, the Netherlands, New Zealand, Norway, Poland, Portugal, the Slovak Republic, Slovenia, Spain, Sweden, Switzerland, Türkiye, the United Kingdom and the United States. The European Commission takes part in the work of the OECD.

OECD Publishing disseminates widely the results of the Organisation's statistics gathering and research on economic, social and environmental issues, as well as the conventions, guidelines and standards agreed by its members.

NUCLEAR ENERGY AGENCY

The OECD Nuclear Energy Agency (NEA) was established on 1 February 1958. Current NEA membership consists of 34 countries: Argentina, Australia, Austria, Belgium, Bulgaria, Canada, Czechia, Denmark, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Japan, Korea, Luxembourg, Mexico, the Netherlands, Norway, Poland, Portugal, Romania, Russia (suspended), the Slovak Republic, Slovenia, Spain, Sweden, Switzerland, Türkiye, the United Kingdom and the United States. The European Commission and the International Atomic Energy Agency also take part in the work of the Agency.

The mission of the NEA is:

- to assist its Member countries in maintaining and further developing, through international co-operation, the scientific, technological and legal bases required for a safe, environmentally friendly and economical use of nuclear energy for peaceful purposes;
- to provide authoritative assessments and to forge common understandings on key issues, as input to government decisions on nuclear energy policy and to broader OECD policy analyses in areas such as energy and sustainable development.

Specific areas of competence of the NEA include the safety and regulation of nuclear activities, radioactive waste management and decommissioning, radiological protection, nuclear science, economic and technical analyses of the nuclear fuel cycle, nuclear law and liability, and public information. The NEA Data Bank provides nuclear data and computer program services for participating countries.

This document, as well as any data and map included herein, are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.

Corrigenda to OECD publications may be found online at: *www.oecd.org/about/publishing/corrigenda.htm*.

Please cite this report as:

NEA (2026), "NEA Workshop on Digital Safety Case Methods and Development: Summary Report", OECD Publishing, Paris,
https://doi.org/10.82155/894d-4n72.

© OECD 2026

Attribution 4.0 International (CC BY 4.0).

This work is made available under the Creative Commons Attribution 4.0 International licence. By using this work, you accept to be bound by the terms of this licence (https://creativecommons.org/licenses/by/4.0).

Attribution – you must cite the work.

Translations – you must cite the original work, identify changes to the original and add the following text: *In the event of any discrepancy between the original work and the translation, only the text of original work should be considered valid.*

Adaptations – you must cite the original work and add the following text: *This is an adaptation of an original work by the OECD. The opinions expressed and arguments employed in this adaptation should not be reported as representing the official views of the OECD or of its Member countries.*

Third-party material – the licence does not apply to third-party material in the work. If using such material, you are responsible for obtaining permission from the third party and for any claims of infringement.

You must not use the OECD logo, visual identity or cover image without express permission or suggest the OECD endorses your use of the work.

Any dispute arising under this licence shall be settled by arbitration in accordance with the Permanent Court of Arbitration (PCA) Arbitration Rules 2012. The seat of arbitration shall be Paris (France). The number of arbitrators shall be one.

Foreword

The Expert Group on a Data and Information Management Strategy for the Safety Case (EGSSC) operates under the OECD Nuclear Energy Agency (NEA) as part of the Radioactive Waste Management Committee (RWMC). Established in 2020 to address the growing complexity of safety cases, the EGSSC focusses on developing strategies for the effective management of data and information throughout the lifecycle of a geological repository safety case.

In October 2022, the EGSSC held the "NEA IDKM/EGSSC Workshop on Digital Safety Case Methods and Development" in Berlin. This event benefited from the input of the Integration Group for the Safety Case (IGSC), the Working Party on Information, Data and Knowledge Management (WP-IDKM) and the Forum on Stakeholder Confidence (FSC), in particular through the Programme Committee and through a questionnaire for the workshop. The event brought together experts from waste management organisations, regulators and other stakeholders to explore the potential of digital tools to enhance the development, review and communication of safety cases. The workshop highlighted the importance of involving all stakeholders − technical and non-technical alike − in the digital transition, ensuring that the safety case remains transparent, accessible and adaptable to evolving needs. This report was approved by the RWMC on 30 September 2025 by the written procedure.

This workshop marked a pivotal moment in the ongoing effort to digitalise safety cases, offering valuable insights that will shape future developments. It has therefore been decided to proceed with a second workshop to continue addressing the challenges of the digital safety case. The EGSSC remains dedicated to advancing best practices in data and information management, ensuring that safety cases continue to meet the highest standards of safety, sustainability and stakeholder engagement.

Acknowledgements

On behalf of all participants, the Nuclear Energy Agency (NEA) expresses its gratitude to the Bundesgesellschaft für Endlagerung (BGE) for hosting the "NEA IDKM/EGSSC Workshop on Digital Safety Case Methods and Development" in Berlin, Germany, and for their impeccable local logistical support.

For their invaluable contributions to the success of the Workshop, the NEA expresses its gratitude to the members of the Workshop Programme Committee for defining and facilitating the Workshop: Manuel Capouet (ONDRAF/NIRAS, Belgium), Alexander Carter (NWS, UK), Olivier De Clercq (FANC, Belgium), Jonathan Kindlein (BGE, Germany), Ann-Kathrin Leuz (ENSI, Switzerland), Daniel Marta (SOGIN, Italy), Shogo Nishikawa (NUMO, Japan), Luca Piciaccia (DSA, Norway) and Klaus-Jürgen Röhlig (TUC, Germany). Their expertise, guidance and engagement were instrumental in the success of the event.

In addition, the NEA greatly appreciates the valuable contributions of guest speakers, presenters, panellists and facilitators for generously sharing their time and knowledge, and acknowledges the insightful participation of all the attendees, who significantly enriched the conversation with a unique blend of cross-disciplinary perspectives.

This report is the result of a sustained collaborative effort that deserves special commendation. The NEA expresses its gratitude to the workshop rapporteur, Oliver Hall (NWS, United Kingdom), for skilfully summarising all the technical talks and panel discussions, and to all the breakout session facilitators and rapporteurs, for conducting and capturing the interesting discussions held. Finally, the NEA gives special thanks to the Workshop Programme Committee members and the members of the Expert Group on a Data and Information Management Strategy for the Safety Case (EGSSC), for reviewing these materials and for collaboratively developing the main conclusions and recommendations sections.

The efforts of all these contributors made this event a success and this report possible.

Table of contents

List of acronyms **7**

Executive summary **11**

1. Background **12**

2. Objectives and scope of the workshop **13**

2.1. Objectives 13
2.2. Scope 13
2.3. Programme description 14

3. Summary of presentations **16**

3.1. Session 1 16
3.2. Session 2 35

4. Panel discussion **60**

4.1. Q1 How can digital technologies and systems benefit knowledge and information management and help demonstrate a Golden Thread between safety claims and evidence? 60
4.2. Q2 How can digital workflows increase auditability and efficiency in the safety case production process? 63
4.3. Q3 How can digital technologies and systems help to improve understanding, presentation and communication of the safety case with both technical and non-technical interested parties? 64

5. Group discussion **66**

5.1. Opportunities and challenges of digitisation approaches 66
5.2. Implementation of the digitisation approaches 72

6. Conclusion and recommendations **77**

References **79**

Appendix A. Workshop agenda **81**

Appendix B. Questionnaire **85**

Figures

Figure 3.1. Representation of the RDF model 17
Figure 3.2. UK EA Bathing Water Application 19
Figure 3.3. SPARQL query used to answer questions on the DBpedia dataset 21
Figure 3.4. DBpedia entry for the Morsleben Waste Repository 21
Figure 3.5. Visualisation of the databus 23
Figure 3.6. Databus identifier structure and an example from Open Energy 23
Figure 3.7. The virtual reality continuum 24
Figure 3.8. Simplified definition of augmented reality 25
Figure 3.9. BGE AR system to support underground workers 27
Figure 3.10. Example of NER form the PharmKE Knowledge Extraction Platform (Jofche et al., 2023[5]) 30
Figure 3.11. Mapping of concepts recognised in NER to terms in COPD-relevant ontologies (Fu et al., 2015[6]) 31
Figure 3.12. Question-answering news chatbot 33
Figure 3.13. The NER component of the LISTENER project shows identified entities like hazards 34
Figure 3.14. CoolRep structure (NEA, 2022[8]) 38
Figure 3.15. Extract from the Conceptual Model Interaction Matrix (Bailey and Billington, 1998[10]) 41
Figure 3.16. Relationships between FEPs and safety functions visualised by SKB 42
Figure 3.17. NIREX (UK) FANFARE GUI 43
Figure 3.18. VISI Home Screen 44
Figure 3.19. Authoring environment for VISI, Oxygen XML, with specialisations that add structured content 46
Figure 3.20. Content displayed in VISI 47
Figure 3.21. Claims, arguments and evidence diagram in VISI 47
Figure 3.22. A snapshot of VISI data plotted in a graph database (Neo4J) 48
Figure 3.23. Illustrative example of a data model for a DGR safety case 49
Figure 3.24. The high-level approach for the management of radioactive waste in Sweden 50
Figure 3.25. SKB FEP database GUI 51
Figure 3.26. Select FEP categories in SKB FEP database 52
Figure 3.27. A FEP record in the SKB FEP database 52
Figure 3.28. Safety assessment flowchart for SR-PSU (SKB, 2015[19]) 53
Figure 3.29. The auto-generated interactive flowchart for FEPs and dataflows for SFR 53
Figure 3.30. Structure of the SKB site investigation database 54
Figure 3.31. Simple example of surface parameters in the SKB site investigation database 55
Figure 3.32. Typical contents of a safety case (from IAEA (2012[25])) 58

Table

Table 2.1. Examples of Topics Within Scope and Out of Scope 14

List of acronyms

ADAMS	Agencywide Documents Access and Management System (US NRC)
AGBe	Arbeitsgruppe Bevölkerung (TRANSENS Citizen Working Group, Germany)
AMF	Assessment model flowcharts
API	Application programming interface
AR	Augmented reality
AV	Augmented virtuality
BASE	Bundesamt für die Sicherheit der nuklearen Entsorgung (German Federal Office for the Safety of Nuclear Waste Management)
BERT	Bidirectional Encoder Representations from Transformers
BGE	Bundesgesellschaft für Endlagerung
BIM	Building information modelling
BPMN	Business Process Model and Notation
CAE	Claims, arguments and evidence
CAD	Computer-aided design
CAVE	Cave Automatic Virtual Environment
CNC	Computer numerical control
COPD	Chronic obstructive pulmonary disease
DGR	Deep Geological Repository
DITA	Darwin Information Typing Architecture
DSC	Digital safety case
DSSC	Disposal System Safety Case
EA	Environment Agency
EGSSC	Expert Group on a Data and Information Management Strategy for the Safety Case
EHD	Electronic hearing docket
ENSI	Swiss Federal Nuclear Safety Inspectorate
ESC	Environmental Safety Case
FANFARE	Historical NIREX safety case software/tool
FANC	Federal Agency for Nuclear Control (Belgium)
FDA	US Food and Drug Administration
FEP	Features, Events and Processes
FSC	Forum on Stakeholder Confidence

GDF	Geological disposal facility
GIS	Geographic Information System
GISADA	Geographic Information System Analysis Database
GIT	Git Version Control System
GUI	Graphical user interface
HTC	HTC Corporation (used in HTC Vive)
HTTP	HyperText Transfer Protocol
IAEA	International Atomic Energy Agency
IGSC	Integration Group for the Safety Case (NEA)
IDKM	Information, Data and Knowledge Management
InfAI	Institute of Applied Informatics
INSPIRE	Infrastructure for Spatial Information in Europe
JAEA	Japan Atomic Energy Agency
JURION	Legal Information Platform (product name)
KMS	Knowledge Management System
LIMS	Laboratory Information Management System
LISTENER	Leverage Insight from Safety Text Excerpts with Named Entity Recognition
LLW	Low-level waste
LLWR	Low-level waste repository
LSN	Licensing Support Network
MCC	Matthews Correlation Coefficient
MRC	Machine reading comprehension
NEA	Nuclear Energy Agency
NER	Named Entity Recognition
NGO	Non-governmental organisation
NIH	National Institutes of Health
NLP	Natural Language Processing
NNL	National Nuclear Laboratory (United Kingdom)
NRC	Nuclear Regulatory Commission (United States)
NWS	Nuclear Waste Services (United Kingdom)
OCR	Optical character recognition
OECD	Organisation for Economic Co-operation and Development
OKG	Open Knowledge Graph
OS	Ordnance Survey
PA	Performance assessment
PDF	Portable Document Format

PROV-O	Provenance Ontology
QA	Quality assurance
RDF	Resource Description Framework
RCM	Requirements and Configuration Management
RD&D	Research, Development and Demonstration
RK&M	Records, Knowledge and Memory
RWMC	Radioactive Waste Management Committee
RWMD	Radioactive Waste Management Division
RWMO	Radioactive Waste Management Organisation
SC	Safety Case
SDM	Site Descriptive Model
SER	Set of Essential Records
SFK	Spent fuel repository (Sweden)
SFL	Repository for Long-lived Low- and Intermediate-level Waste
SFR	Repository for Short-lived Radioactive Waste
SHACL	Shapes Constraint Language
SICADA	Site Characterisation Database
SKB	Svensk Kärnbränslehantering AB (Sweden)
SKOS	Simple Knowledge Organization System
SPARQL	SPARQL Protocol and RDF Query Language
SQL	Structured Query Language
SQEP	Suitably Qualified and Experienced Person
SRL	Safety Reference Level
SVN	Subversion Version Control System
TAP	Transdisciplinary Work Package
TUC	Technical University of Clausthal
URF	Underground research facility
URI	Uniform Resource Identifier
URL	Uniform Resource Locator
VISI	Visualisation of System Information
VR	Virtual reality
WENRA	Western European Nuclear Regulators Association
WGWD	Working Group on Waste and Decommissioning (WENRA)
WMO	Waste Management Organisation
WP-IDKM	Working Party on Information, Data and Knowledge Management (NEA)
WYSIWYG	What You See Is What You Get

XML	eXtensible Markup Language
XR	Extended reality

Executive summary

The workshop on digital safety case (DSC) held in Berlin, in 2022, highlighted the evolving landscape of safety case production within the context of radioactive waste disposal. Traditional safety cases are often structured as a suite of documents managed through conventional tools like Microsoft Office software, with many organisations relying on manual processes and digital tools to partially support their production and review. As digitalisation efforts advance, it is crucial to carefully assess the level of digital implementation based on cost-benefit analyses, the requirements and objectives of the programme stage and the stability of safety case methodologies.

The complexity of safety cases can be challenging for non-specialists to understand as well. Enhancing communication through a hierarchical reporting system with a clear, central argument and better integration of electronic documentation systems could make safety cases more accessible. Suggested key areas for which digitalisation can provide support include knowledge management, and features events and processing (FEP) scenario development. It can help ensure the comprehensiveness and traceability of discourse chains during scenario development, while also supporting the evolution of the underlying knowledge base.

The workshop also underscored the importance of learning from past tools and approaches, stressing the need to maintain existing systems rather than reinventing the wheel, ensuring knowledge transfer, and integrating siloed databases and tools. Lessons from the VISI Project in the United Kingdom highlighted the three pillars of digitalisation of safety cases: structuring, authoring and managing (including version control and change management). The success of a DSC depends on careful planning, understanding user and stakeholder needs, and linking digital tools to the overall safety and management strategy.

Insights from the safety assessment and associated databases from SKB, Sweden, demonstrated that safety assessments often involve repetitive structures in the form of flow charts and written texts in the reports. Opportunities exist to enhance efficiency, reduce quality assurance risks and ensure consistency when integrating and linking various tools and activities, as updates are needed across multiple places.

While there are numerous requirements for safety case content, there is flexibility in the choice of format and tools used for development. However, frequent interactions with regulators, local communities and other stakeholders are essential to ensure that their needs and expectations are met.

The workshop concluded that digitalisation should not be pursued as an end goal, but as a means to improve the structure, workflow and communication of safety arguments. Digital tools offer opportunities to enhance consistency, traceability and accessibility of safety requirements, ultimately strengthening the safety case. Additionally, providing access to data sets and parameters can enable independent calculations by authorities as part of their review, further supporting robust safety case development. The workshop demonstrated that collaboration between safety case experts and digital technology specialists presents significant opportunities for innovation and knowledge transfer. Continued exchange and collaboration among these experts are essential for advancing the field.

As it was recognised in the workshop that a common understanding of what constitutes a DSC has not been established, the EGSSC is collecting elements and ideas for the development of the vision for the DSC with the results of the workshop.

1. Background

Radioactive waste is produced in all phases of the nuclear fuel cycle and from the use of radioactive materials in industrial, medical, defence and research applications. Disposal in engineered repositories, located in suitable geological formations, is being developed worldwide as the reference solution for the long-term management of high-level, or long-lived radioactive waste.

The safety of a repository is evaluated and documented in a "safety case" that supports the key decisions to progress between each stage of repository development and into closure. It presents the underlying supporting evidence and methods that give confidence in the quality of scientific and institutional processes, as well as in the results of analyses on repository performance and safety.

A safety case is traditionally presented through a suite of printed documents, often with a significant page count (e.g. in excess of 10 000 pages). These documents make qualitative and quantitative arguments through the use of text, tables, diagrams and plots with references provided to underlying reports, models and similar items. However, the advent of digital technologies brings an opportunity to review existing processes for the production of safety cases, as well as approaches for the presentation of the information contained therein. In doing so, we can investigate whether additional benefits (for all stakeholders involved in the development programme for the repository) can be gained from the adoption of new technologies. Examples of envisaged benefits include the use of:

- interactive infographics, video (or other media) and the use of multiple output channels (e.g. pdf, e-books or web publishing) for a single information source, to improve the presentation of information and communication, either within the radioactive waste management organisation itself or to other stakeholders;
- resolvable hyperlinks, and full-text or link search capability to improve usability (including in the review process) and reduce repetition of information within safety case sections;
- glossaries and controlled vocabularies for terms (or metadata) used in the safety case, allowing consistency and enrichment of existing information, while potentially aiding machine-readability, thereby facilitating the long-term management, storage and archiving of the safety case and the information contained within;
- workflow processes and electronic signing to increase auditability and timeliness of information sharing and efficiency.

Digitisation options for the development of safety cases form a continuum, varying from minor enhancements to existing paper-based processes, to adopting a full digital workflow involving significant changes to existing processes, formats and methods. There is currently no common international understanding of what the latter may look like. The topic of digitisation forms the subject of a number of ongoing activities within the EGSSC, who propose to host a workshop on "Digital Safety Case Methods and Development".

2. Objectives and scope of the workshop

2.1. Objectives

The objectives of this workshop were to:

- share learning on the work currently being carried out in member states in digital safety cases for radioactive waste disposal;
- lay the groundwork to develop a common understanding of the opportunities, challenges and potential extent of a digital safety case (a "shared vision") for radioactive waste disposal;
- discuss differences in the priorities made by different organisations that progressed safety case digitisation, depending on the context under which they operate (e.g. stage of programme, quantity of existing records);
- establish general guidance for member organisations to advance digitisation within their organisations;
- inform the future work programme of EGSSC in 2023, and beyond.

2.2. Scope

The workshop focussed on all aspects related to *authoring, updating, and presenting the safety case and associated performance assessment (PA)* using digital systems and technologies, together with its approval and review.

To avoid dilution of the specialist nuclear "safety case" emphasis of the workshop, only aspects related directly to the safety case narrative, and its associated safety and performance assessment were considered; no earlier activities which this assessment relies upon were considered.

This was intended to provide focus on those aspects which are specific to a nuclear safety case without diluting this focus through the consideration of wider needs such as systems or databases for digital design/BIM, site characterisation, GIS, radionuclide inventories or similar. Including such activities would have meant taking the workshop to a higher level, inviting additional attendees to cover multiple specialisms and reducing the time dedicated to the core of the safety case itself, i.e. the development and presentation of the claims and arguments that support the safety objective. In addition, many of these omitted aspects are within the focus of developments driven by other industries (for example, oil and gas exploration or urban design), which provide opportunities for commercially available offerings that are adaptable to the specialist needs of geological disposal.

Note that the *integration* between these latter aspects and the safety case *was* within scope, however. This includes the means by which the safety case system interacts with such other systems (human, data files with metadata, automated protocols), and how changes cascade and are tracked.

The digitisation of the safety case may also support the objective of improving stakeholder confidence by local stakeholders in affected communities and the wider interested public. To accommodate such a non-expert audience, it would be interesting to know which additional functions are necessary. For example, are special guidance within this digital system, different levels of detail for the same information or providing different modes of

presentation adapted to the audience addressed necessary? Examples of aspects proposed to be within scope, and out of scope, for the workshop are provided in Table 2.1.

Table 2.1. Examples of topics within scope and out of scope

Within scope	How digital systems and technologies can help with:
	• recording safety case claims, arguments and evidence (the safety case "Golden Thread"), including identifying evidence gaps and RD&D planning;
	• use of interactive graphics or digital media to communicate the safety case narrative, including visualisation of input data and PA results;
	• data, model and calculation ("run") management for scenario development and performance assessment;
	• change control, configuration management (including consistency with wider systems) and data freeze for the safety case (versioning, change tracking, timely notification of status change, etc.);
	• process automation, including recording and auditing approval workflows;
	• providing SC essentials and SC-related information *consistently* to various audiences in a traceable and transparent way, including non-specialists, and allow for feedback from these actors;
	  ○ Potential for interaction with the NEA Forum for Stakeholder Confidence (FSC)
	• interactions with regulators;
	  ○ providing regulatory access to the safety case
	  ○ taking feedback from regulators on the safety case and managing review comments
	  ○ potential for interaction with the NEA Regulator's Forum
	• the use of consistent vocabulary and terminology across the SC;
	• searching for safety case content, including its underpinning knowledge base, or other knowledge management activities;
	• transfer of the safety case to company or national archives in line with policy or legislation;
	• integrating the digital safety case with supporting tools and databases (such as design, GIS) and associated interoperability requirements.

Out of scope	The following systems, for which third-party, "off-the-shelf" solutions are becoming more and more available, are proposed to be out of scope: • site characterisation, GIS, etc.; • storing inventory and waste package information; • CAD and design, "digital twins", BIM, etc.; • finance, procurement, HR or other business support functions.

2.3. Programme description

This in-person workshop, held in Berlin, Germany, over two full days, was structured around the following four sessions:

- Session 1 - Digitisation and digitalisation experiences:
  Keynote presentations and panel discussion on digital technologies and approaches which could have applicability to safety cases for radioactive waste disposal.
- Session 2 - Stakeholder perspectives and national use cases:
  Presentations and panel discussions on perspectives of potential stakeholders of digital safety cases as well as Radioactive Waste Management Organisations' (RWMOs) current work on digital safety case tools and systems.

- Session 3 - Digitisation of the safety case:
  Breakout sessions with open discussion, followed by a summary in plenary to share and develop ideas on relevant topics.
  - Session 3a: Opportunities and challenges
    - How can digital technologies and systems benefit knowledge and information management, and help demonstrate a "Golden Thread" between safety claims and evidence?
    - How can digital workflows increase auditability and efficiency in the safety case production process?
    - How can digital technologies and systems help to improve understanding, presentation and communication of the safety case with both technical and non-technical interested parties?
  - Session 3b: Implementation of digitisation
    - How can a business case support digitisation of the safety case and what are the arguments and components of such a case?
    - How do organisations migrate to a digital safety case?
    - What does a vision for an 'ideal' digital safety case look like and what future work would be needed to achieve this?
- Session 4 - Summary of workshop conclusions and proposed ongoing work

3. Summary of presentations

3.1. Session 1

The topic of Session 1 on Day 1 was digitisation and digitalisation experiences. Four speakers from different organisations gave presentations on digital technologies and approaches that are being used in many sectors today.

3.1.1. Linked data

Khalil Ahmed from Epimorphics provided an overview on the concept of linked data and its uses. Below is a summary of his overview.

Linked data is a set of principles for publishing data. It is concerned with the standards, techniques and tools that allow organisations to connect and publish structured data, primarily (but not exclusively) over the World Wide Web. These are embodied in a set of principles for publishing data on the World Wide Web. Linked data commonly overlaps with the related concepts, such as:

- Open Data - which is about restrictions on the (re)use of data;
- RDF (Resource Description Framework) - which is a technical standard used for models, which represent data with semantics; but linked data principles can be applied to restricted data and to data formats other than RDF.

Linked data challenges

Linked data seeks to address challenges associated with how to make use of data, which is distributed across many datasets, domains and organisations; challenges which are closely linked to the development of the World Wide Web in the 1990s and early 2000s. These challenges can be summarised as:

- How to provide access to data so it can be reused?
- What forms to store knowledge and data in?
- How to enable the discovery of data?
- How to represent relationships between data?
- How to integrate data from large numbers of unknown data sets?
- How to track the provenance of data?
- How to manage the growing amount of data and knowledge?
- How to cope with changes over time?

Linked data principles

To address these challenges, a set of linked data principles was provided, including:

- using Uniform Resource Identifiers (URI) as names for things;
- using HTTP URIs so that people can look up those names;
- providing useful information when someone looks up a URI;
- including links to other URIs so that they can discover more things.

The use of URIs as names of things provides the advantage that widely implemented schemas, like HTTP web addresses (URLs) can be used for URIs which then enables someone else to understand them (such as in a web browser). The specific use of HTTP(s) has the advantage that these addresses are controlled, which avoid duplications (it is bad practice to assign URIs to addresses over which you do not have control), and take advantage of widely available protocols, so there is a low threshold for adoption and use.

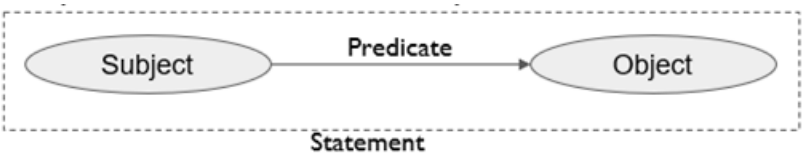
The address ought to provide information about the object in question, and this ought to use an appropriate standard, such as RDF.

Linked data elements

Linked data techniques often use certain standards for storing and querying data. Commonly used standards are RDF and SPARQL (Simple Protocol and RDF Query Language):

- **RDF** is a graph-based[1] model for representing data. The logic of this model is that a statement is composed of a subject, an object and a predicate, where the subject and object are entities, and the predicate is the relationship between them. URIs are assigned to all objects to allow them to be uniquely identified. Statements can be collected together into named graphs with their own URI. There are multiple formats for data exchange. An informal example is the statement "Bob is a friend of Alice". In this case "Bob" is the subject, "Alice" the object and "is a friend of" the predicate. A collection of a subject, object and predicate is often referred to as a triple.

Figure 3.1. Representation of the RDF model



- **SPARQL**[2] is a query language for use on RDF, enabling pattern matching against graphs constructed in an RDF format. It includes some basic mathematical and string functions. It is analogous to the use of Structured Query Language (SQL) on relational databases. It is extensible, for example, with Geographic Information System (GIS), and there are protocols for client-server interaction, allowing a distributed query to be run across multiple services.

[1] Graph in this context refers to representing relationships in a network of objects, so a collection of objects and their relationships (or nodes and edges).

[2] SPARQL is a recurring acronym standing for Simple Protocol and RDF Query Language.

Related to linked data are the information-theoretic concepts of ontologies and taxonomies:

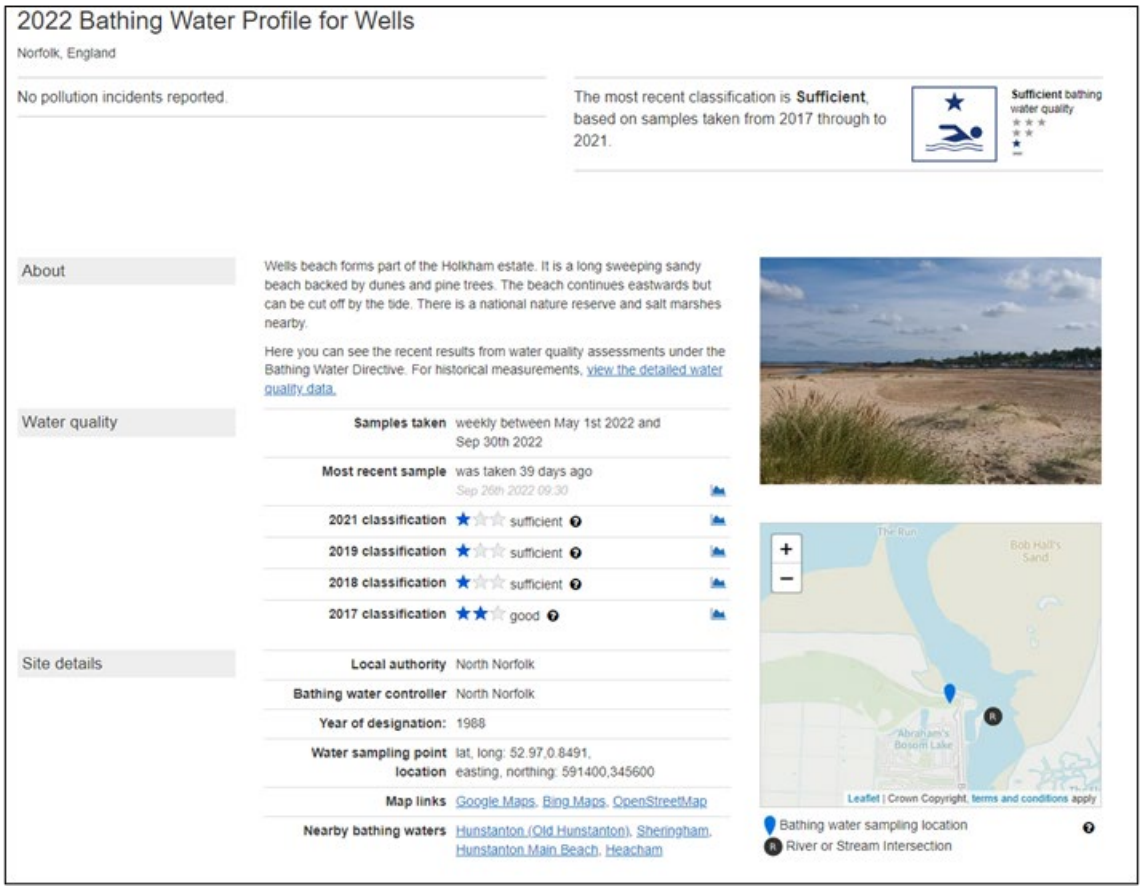
- **Ontologies** are descriptions of the type of things in a domain and how they relate to each other. These may provide common URIs for some predicates and RDF type statements and may define constraints on how predicates are used in the domain. Large numbers of predefined ontologies exist for particular domain types; for example, Simple Knowledge Organisation Systems (SKOS) for hierarchical taxonomies. Ideally, when defining a new ontology, a specialisation of an existing ontology should be made which narrows or broadens the definitions of the types and their relationships. This avoids repeating the same definitions (which can lead to inconsistency).
- **Taxonomies** are controlled lists of instances ("things") from the same ontology. For example, a list of chemical elements, where an ontology on chemical elements provides the information and relationships used to define a chemical element (so properties like atomic mass, atomic number, etc., would be examples of information, and would be isotope an example of a relationship, in this case) and the taxonomy derived from the ontology is the populated list from the periodic table. Taxonomies are useful as they provide a way of ensuring that users are talking about the same "thing". Defining and managing taxonomies is often a function of a domain expert group.

Example

The UK Environment Agency (EA) has created a bathing water web application, which provides a demonstration of how linked data principles and technologies can be implemented. The application provides data on bathing water quality collected from around the United Kingdom, such as data on bacteria like *E. coli* present in the water, with descriptions that are relevant, including the catchment area and some relevant definitions such as plankton. The EA provides this data to inform the public and allow them to make decisions on whether to bathe in a specific place. The application is shown in Figure 3.2 for Wells in Norfolk, England[3]. Information on the page is taken from an RDF source.

[3] The EA application can be accessed from https://environment.data.gov.uk/bwq/profiles/profile.html?site=ukh1305-10000

Figure 3.2. UK EA Bathing Water Application



The application provides key information on bathing water, that is linked via URIs to the relevant data. For example, the 2021 classification on water quality in Figure 3.2 is linked to a dataset on water samples. Data on sewage from an external source (the water company, Anglian water) is also linked to, as is information on the company provided by UK Companies house (the UK Government body for company registrations). Other links are provided to Ordinance Survey (OS), to show the map of the area. These aspects are good examples of the use of linked data principles (particularly the principle of including more links to discover more things). For example, in order to show the map of the area on the page, all that is required is to provide the appropriate link to the OS map information.

Benefits and challenges

The use of linked data has a number of benefits and challenges. The benefits include:

- It provides a low bar for accessing data.
- It is well-supported in tools and libraries that are provided commercially or for free.
- Shared semantics ensure users can understand what an author really means by an assertion.

- It is an open system, meaning that someone can make statements which anyone else can then reference. This enables knowledge to be built up more easily and incrementally. Users are able to produce new information that builds on and references existing knowledge and information, without disturbing or altering that information. It therefore facilitates incrementally expanding knowledge in which the flow of information is traceable.

It also has several challenges, which include:

- If data is distributed across many servers (usually on the web), this can create issues around latency and the uptime/availability of the servers.
- There is a dependency on publishers of data to maintain access to it, meaning that there is often a dependence of a third party's dataset.
- There is a large technology "stack"[4], which can be an initial hurdle for implementation.

3.1.2. Knowledge graphs

Sebastian Hellmann from DBpedia Association and Institute of Applied Informatics (InfAI) introduced DBPedia and knowledge graphs, and provided technical highlights on them. Below is a summary of his talk.

DBPedia

DBpedia is a crowdsourced community effort to extract structured data from information available in Wikipedia, Wikimedia Commons and Wikidata. DBPedia has produced an Open Knowledge Graph (OKG), which is available free over the web. It contains 20 billion "triples"[5] which are stored as graphs[6], and the dataset itself as a graph database[7]. Such a knowledge graph can then be queried with tools like SPARQL.

DBpedia extracts the semi-structured data from "infoboxes", where data is stored as sets of key-value pairs, and then transforms them into RDF representations. This newly structured data can then be used to answer complex queries on the Wikipedia dataset. For example, answers can be given to questions like: "*List all soccer players that played for a club that has a stadium with more than 30,000 seats*".

Figure 3.3 shows how this question can be translated into a SPARQL query on the data in DBpedia.

[4] A technology stack is the set of tools, libraries, programming languages, etc., that is required for a particular use case (e.g. certain websites).

[5] A triple is "fact". In a linked data sense, it is a statement composed of a subject, object and predicate. See above in the section on linked data.

[6] This is a graph in a mathematical sense, of nodes (objects and subjects) and edges (predicates) which show a relationship between nodes.

[7] A database of nodes and edges, which make up a graph.

Figure 3.3. SPARQL query used to answer questions on the DBpedia dataset

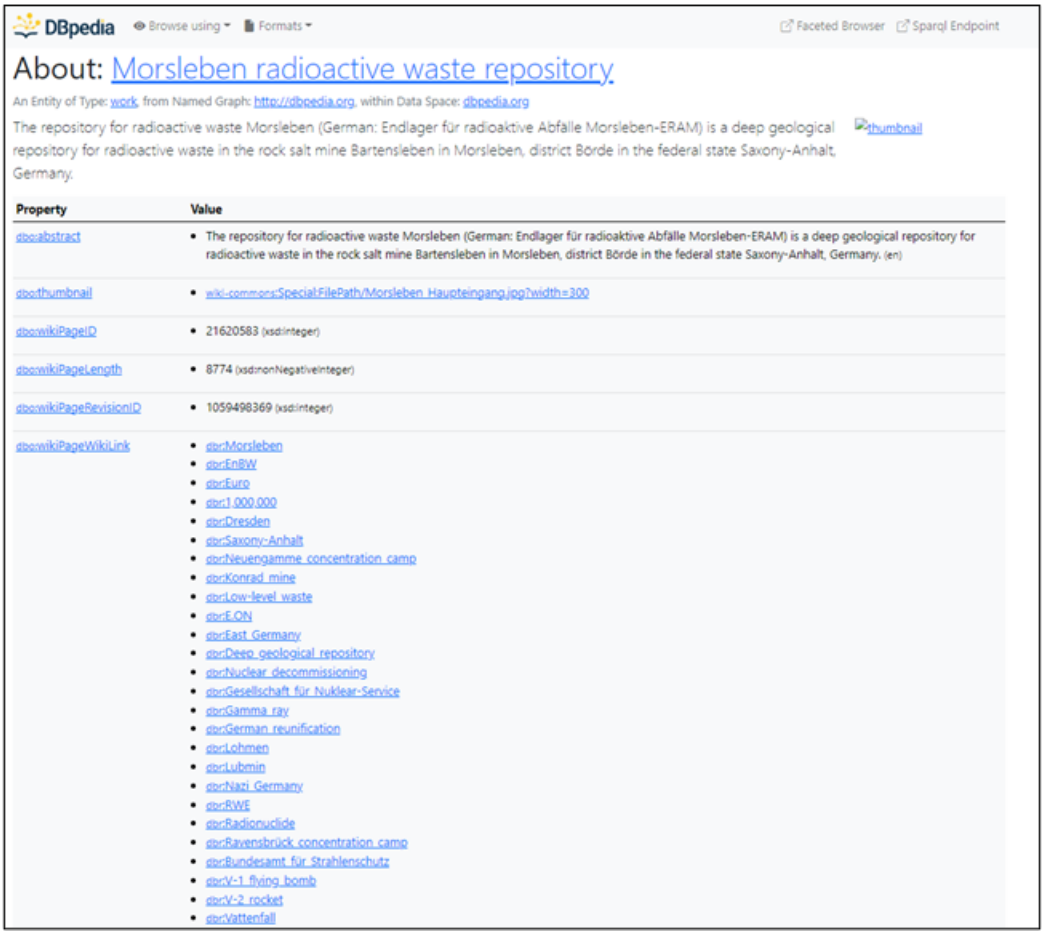
```
SELECT DISTINCT ?person ?name ?countryOfBirth ?population ?team ?stadium ?stadiumCapacity
WHERE {
    ?person a dbo:Person .
    ?person rdfs:label ?name .
    ?person dbo:birthPlace ?countryOfBirth .
    ?countryOfBirth dbo:populationTotal ?population .
    ?person dbo:team ?team .
    ?person dbo:position|dbp:position <http://dbpedia.org/resource/Goalkeeper_(association_football)> .
    ?team dbo:stadium ?stadium .
    ?stadium dbo:seatingCapacity ?stadiumCapacity .

    FILTER (langMatches( lang(?name), "EN" ) )
    FILTER (?stadiumCapacity > 30000)
    FILTER (?population > 10000000)

} ORDER BY DESC(?stadiumCapacity)
```

Entities from the DBpedia dataset contain many links to other entities. For example, the page on the Morsleben Waste Repository, shown in Figure 3.4, has links to entities representing concepts related to Morsleben, which have been extracted from the Wikipedia data (e.g. radioactive, East Germany).

Figure 3.4. DBpedia entry for the Morsleben Waste Repository



Economics of knowledge graphs

Economic output is conventionally described in terms of the factors of production:

Output=Land ×Labour ×Capital × ???(Solow Residual) where the Solow Residual (represented by the ???) is a measure of ignorance, or the proportion of economic output not attributable to the factors of production. The effect of productivity increases due to technological innovation would be a part of this and is described as the *Structural Capital*. This capital enables human capital to function better and remains when individuals leave an organisation. In other words, this encompasses the tools, routines, methods and procedures which form the organisation's body of knowledge.

The purpose of the digital safety case is to build structural capital, enabling the human capital, which composes, uses and reviews the safety case to become more productive. Knowledge graphs could be a key component of such a digital safety case. They would not necessarily use a centralised graph database, but would provide a vehicle to discover, access and integrate information that is distributed on demand. Techniques and tools like linked data, RDF and SPARQL were designed to work on top of and across existing systems and their boundaries to help achieve this. However, for productive outputs there would also be a need to invest in the quality of information and data, as well as its re-usability and interoperability that is used to build knowledge graphs.

Semantically enhanced quality assurance and the JURION business case

JURION is a legal information platform developed by the German company Wolter Kluwer. It merges and interlinks over one million documents of content from diverse sources including national and European legislation, and court judgements. It also includes extensive internally written data, as well as local data from customers, social media and web data. In collecting and managing this data, all stages of the Data Lifecycle are present - extraction, storage, authoring, interlinking, enrichment, quality analysis, repair and publication. JURION has made use of linked data and knowledge graph technologies to do this. The development of JURION led to a massive increase in digital products for the company.

Test-driven data development

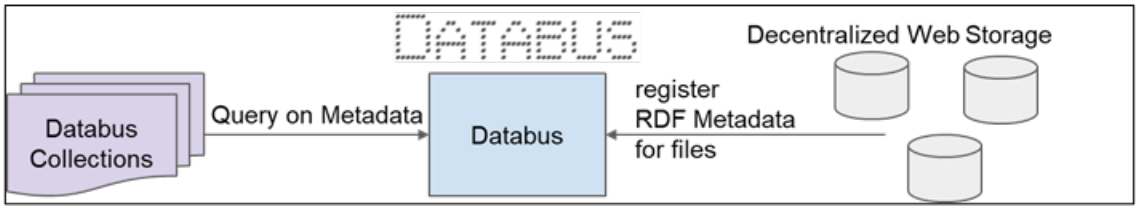
A key challenge associated with JURION was how to continuously incorporate significant amounts of data into the platform. This data is often in a semi-structured form, which needs to be incorporated in a structured form (like that of triples) adhering to data schemas within the software. These schemas are routinely updated and enhanced. This can mean that ensuring data quality becomes challenging, and is a manual task. Automating this task can therefore improve efficiency across many areas. This is addressed by implementing test-driven data development, where there is automatic confirmation that data adheres to a schema. This also has the advantage that it brings the data update process closer to that of the software one. This has been implemented in the RDFUnit (AKSW Research Group, n.d.[1]) software framework for structured data stored in RDF, where unit tests are applied to data. Analogous to unit tests used for software development, these unit tests are automatically derived from schemas and test the adherence of an RDF dataset to the schema as new data is ingested by a system.

Complementary to this, RDF graphs can be validated against a set of conditions by using the Shapes Constraint Language (SHACL). This allows a user to write a query to check whether the RDF graph meets a set of conditions they expect, and, in principle, it allows a user to write a database query to check that there are no errors in the database.

Databus

There are many valuable datasets (or "data treasures") that have been built up online, but which are often in silos, and beyond the reach of publishers and consumers who do not know how to find or retrieve them. The DBPedia Databus is designed as a platform for republishing these files and datasets in a more accessible way. This is modelled after a hardware bus in computer architecture, which transfers data between components inside the computer; in contrast, the DBPedia Databus (DBpedia Association, n.d.[2]) is a virtual bus that stores the address and very basic metadata on files and datasets added to it, while the datasets themselves remain in their decentralised web storage. Linked data principles are applied, using standards like RDF, SPARQL and SHACL to manage files, their content and any additional metadata added to them. The databus allows efficient retrieval and reuse of these datasets as well as structured querying to enable users to find datasets. The system aims at maximum automation, as for the JURION case, and is able to manage variant and version updates to datasets by incorporating some features of software version control systems (e.g. GIT).

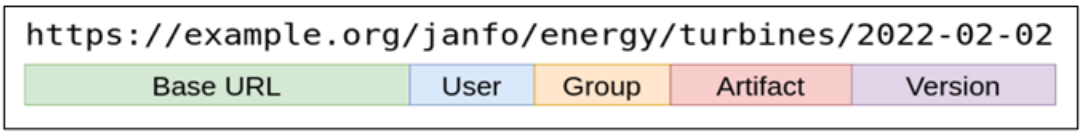
Figure 3.5. Visualisation of the databus



The databus has been developed for the DBpedia dataset but can be applied to other domains. One example is the Open Energy databus (Open Energy Platform, n.d.[3]), part of the Open Energy platform which is designed to facilitate search, discovery and accessibility of data in the energy sector. This includes a published ontology, a database tutorial and factsheet, as well as application programming interfaces (APIs) for interaction with languages, such as python.

The databus makes use of linked-data inspired identifiers that enable the quick retrieval of data. These offer a clean structure, as shown in Figure 3.6, with metadata such as groups or versions within the identifier. This allows an almost automatic generation of metadata (as its embedded in the identifier) and a powerful way of retrieving and accessing files. For example, files for the same user or group can very quickly be retrieved, which is analogous to how files are found and accessed on a hard drive. Data lifecycles of files are also followed here, with versions and users of a file distinguished.

Figure 3.6. Databus identifier structure and an example from Open Energy



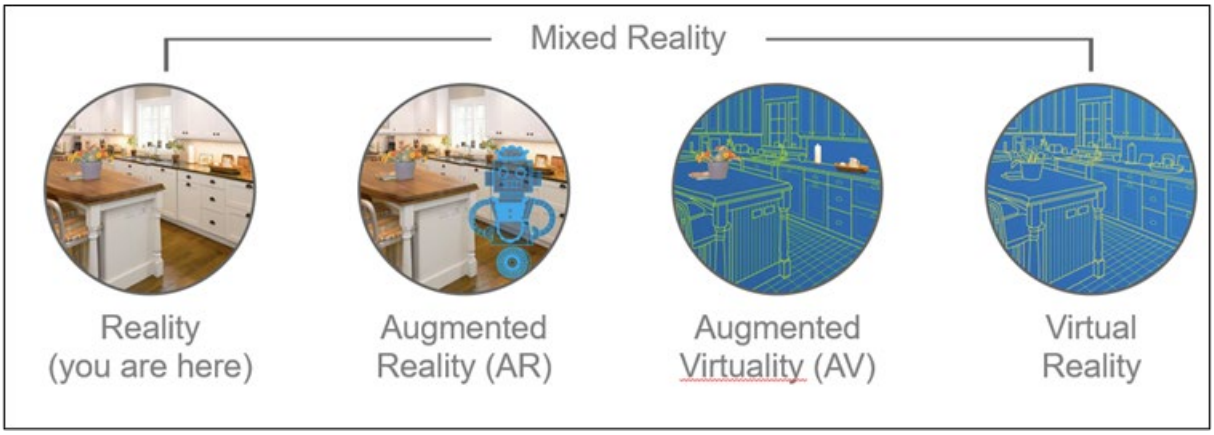
3.1.3. Virtual and augmented reality – applications in production engineering

Philipp Klimant from Chemnitz University of Technology and Fraunhofer Institute IWU 4 presented about virtual and augmented reality and examples of its application. Below is a summary of his presentation.

Virtual reality can be said to exist in a continuum, with one end being real and the other end being virtual. In between these endpoints, a set of mixed realities exists, which involve real and virtual parts, as can be seen in Figure 3.7. These states are:

- *Reality* – where you currently are;
- *Augmented reality (AR)* – reality is augmented to include some virtual components, for example displaying a virtual seat in a real room;
- *Augmented virtuality (AV)* – a virtual work is augmented to include some real components; for example, showing a real seat in a virtual 3D environment;
- *Virtual reality* – a completely virtual 3D environment.

Figure 3.7. The virtual reality continuum



Virtual reality

"Virtual" and "virtual reality" are not quite the same thing:

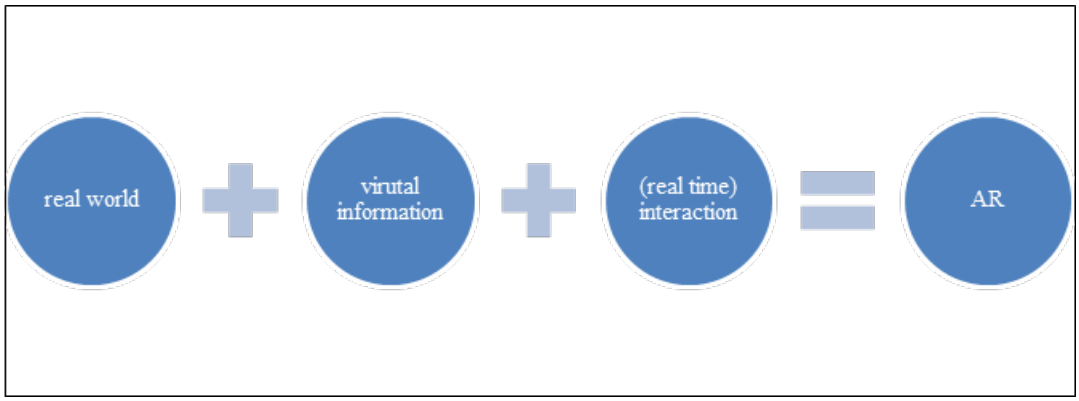
- *Virtual* – a three-dimensional computer-generated world, with objects, characteristics and processes that are digitally illustrated using computers, focussing on shape, effect and functionality;
- *Virtual reality* – virtual (above), plus a degree of immersion, which includes real-time interaction at a 1:1 scale.

Virtual reality (VR) systems can take a few forms; the most commonly used are projective systems and those which use head mounted displays. Projective VR systems, such as CAVE (Cave Automatic Virtual Environment) or Powerwall, use projective screens (up to 5 in the case of CAVE, or one mobile one for Powerwall) which can be either front- or back-projected to offer immersion. These systems can offer different degrees of immersion and are generally more expensive than head mounted displays. They enable group working, since interactions with people are possible. By contrast, head mounted displays offer the highest degree of immersion, with examples including HTC Vive and Oculus Quest. These are less expensive, easier and faster to set up than projective systems. However, interaction with other people is more difficult. Additional feedback can be added in the form of robots providing acoustic or tactile feedback to the user.

Augmented reality

Augmented reality has a simplified definition of 'the real world, plus some virtual information and interaction'. The real world is usually captured by hardware such as a camera and is then visualised. Virtual information is then added through software and can be in the form of 3D-data or additional information, such as text or symbols. There is a need for tracking of the environment so that the system can respond in real time to changes and provide the contextual visualisation expected (Figure 3.8).

Figure 3.8. Simplified definition of augmented reality



AR systems can take a few forms. Handheld smartphones or tablets displaying the "real world" with partly 3D recognition are commonly used. Data glasses are a hands-free system, with partial 3D recognition, but with some limitations regarding interaction (which is hard with a hands-free system), runtime, resolution and performance. A good example of this is Microsoft HoloLens. Projective AR systems are also available, which do not have additional components at the body. These can be used by several users but can have contrast problems and are generally low mobility systems. There are a number of examples of projective AR systems, including:

- virtual X-ray vision of a vehicle – mobile projection system visualises hidden structures of the vehicle;
- the augmented piano – projection of notes on a piano to support learning;

- welding assistance – step-by-step projection of welding instructions on the work piece;
- design presentations – projections of different design variants on a blank shoe.

Applications and research areas

There a several key areas of application of and research into VR and AR systems, including machine and production simulations, digital twins, human-machine interactions and eXtended Reality (XR) technologies.

Machine and production simulations

Machine and production simulation is about virtually testing manufacturing methods or machines. This allows for testing of components before the physical system is produced. This can also reduce the need for physical testing of such systems. Both cases reduce costs. These simulations can take a few forms, such as *virtual commissioning* where a real component is connected with virtual models. A couple of examples of this include a real controller for a milling machine that is tested with a virtual machine, or a car in virtual reality that is used to programme and test control software for the production of the car, to check how it may work on the real car when it is manufactured.

A specific example of machine simulation can be provided by a material removal system. In this case the simulated removal of material is combined with a real control device (computer numerical control – CNC/NC) to perform the subsequent (virtual) removal. The simulations model kinematic movements and associated material removal, which are then fed into the real control device. Additional features can be added to the model including energy consumption. This simulation seeks to guarantee that an item can be manufactured using the device and it enables easy testing for errors.

Digital twins

A digital twin is a combination of design, planning and simulation data, with collected data from the real product, machine and/or process. It connects real, measured data with planning and simulation data, to give a more integrated and accurate picture of the object. This can be combined with 3D geometric information to form a *virtual twin*. A virtual twin combines the data-orientated approach of the digital twin with localisation on a 3D geometry. It is more time- and model-orientated than the digital twin, allowing for a more user-oriented experience.

An application of digital twins is *functional visualisation*, in which simulated data of an application is visualised. An example of this is the application of functional visualisation to press hardening[8]. In this case the visualisation of simulation data occurs, allowing for:

- optimisation and error detection when planning complex processes;
- verification of correct interactions of individual elements within complex process chains before their assembly;
- support of the (re)design process with this as a basis for communication.

[8] Also known as hot stamping, this is a process where ultra-high strength steels are formed into complex shapes. This is commonly used in the automotive industry.

Other applications of digital twins concern *Product and Process Visualisation*, for example:

- a walk into a virtual factory with colleagues to see how it works;
- a walk into places that cannot be accessed for reasons of safety, e.g. dangerous places in manufacturing facilities;
- visualisations of new housing.

Some applications can be built into real-time 3D webpages.

Human-machine interactions

Human-machine interactions concern the application of augmented reality to support humans. This is often in the form of extended reality (XR) assistance for training or learning systems. In XR assistance, the aim is to support the human. Such system can have a number of goals including increasing efficiency, optimising processes, improving error detection and enabling the transfer of knowledge. These XR assistance systems can meet such goals by, for example, providing current or process-relevant information to workers directly on a site, and/or providing individual or group tailored information. Often these use user-friendly mobile devices.

Examples of such tools include: augmented reality systems for monitoring machine tools, where a tablet system can visualise live values of the control system (such as position and speed); and AR tools to support maintenance on tools, which support technicians- they can lay a phone or tablet over a component to bring up relevant work instructions.

A specific example comes from BGE (the German Federal agency for radioactive waste disposal) where underground workers use mobile devices to service or operate control cabinets. The workers view the control through their device, which presents an overlay with 3D visualisations to help with complex operations, as shown in Figure 3.9.

Figure 3.9. BGE AR system to support underground workers

Other examples include:

- VR systems to test human-robot co-operation, designed to test how robots will interact in an environment with humans, using virtual robots with real scenario data. This enables significantly safer testing of robots, allowing the discovery and addressing of potential errors or issues before the real systems are built.
- VR systems for training, designed to reduce mistakes via a virtual training session, before they can happen in real life. This means people can learn from mistakes without the potential safety or cost issues associated with real-life mistakes.

XR Technologies

Extended reality (XR) technologies are systems that involve interaction and collaboration, using XR systems, as well as recognition of semantic content in the real world. A good example of this is LichtAR (Arlight and LLC, 2026[4]), a light simulation designed to visualise the impact of lighting a real room based on virtual light sources. This makes use of 3D reconstructions using AR technology to display and calculate the possible lighting in a room. This is part of a smartphone application for an online lamp shop to improve the experience for customers.

3.1.4. Unlocking knowledge from textual data using natural language processing

Riza Batista-Navarro, from the University of Manchester, provided an overview of natural language processing (NLP) technology and its proof of concept in application to technical reports in the nuclear field. Below is a summary of her overview.

The practice of text mining seeks to extract information from text, e.g. within documents like research papers or official reports. However, there are barriers to extracting information from digital libraries using common methods, such as conventional text searching. This may be addressed by using NLP, which can analyse text in its raw prose format to generate semantic metadata[9] for use in techniques such as text classification, named entity recognition and entity linking. These can unlock knowledge with textual data by enabling targeted searches underpinned by natural language[10] question answering. A proof of concept to apply this process to safety documents has been undertaken by the University of Manchester.

Several digital libraries that implement some of these techniques have been developed. These include:

- PubMed – a digital archive for biomedical literature for the United States National Institutes of Health (NIH). This service enables advanced querying over a large library and includes a full-text search on abstracts of all papers referenced in the archive.

[9] Semantic metadata is metadata (data about data) that describes the meaning of data or content.

[10] Natural language refers to naturally developed human languages (as opposed to those used for computers) such as English.

- Biodiversity Heritage Library – a library that stores digitised legacy literature on biodiversity and currently holds 286 000+ volumes, corresponding to 60M+ pages (PDFs and optical character recognition (OCR) generated text). The content was contributed by a consortium of botanical and natural history libraries. Advanced search is also keyword-based, except for the functionality for finding scientific names, which uses named entity recognition and entity linking. It supports searching through metadata and full-text documents using basic keyword matching.

Barriers to search and knowledge discovery

The conventional way of finding information in textual data is to use keyword-based searching. However, there are some issues associated with such searches, which can be a barrier to discovering knowledge. These issues can be summarised as:

- *The inability to disambiguate* - for example: "California Bay" can refer to a tree or a location, and a keyword-based system does not know how to distinguish between different meanings or contexts.
- *The inability to account for variants or synonyms* - for example: "lion" and "Panthera leo" may refer to the same animal, the latter being the scientific name for a species and the former the common English name.

The implications of the former issue are getting less precise search results, where documents irrelevant to the intended purpose can be returned. In the case of the latter, the effect will be a more limited coverage in searches, where some information may be overlooked in returned results.

NLP for the generation of semantic metadata

To address these issues, NLP can be applied to automatically generate semantic metadata for each document. This is provided in two forms, of which the first is *text classification*. A text classification system will generate classifications for some text and assign a label from a predefined list or dictionary. This can apply to long pieces of text like a full article, or a shorter text like an abstract or even a very short-form text, like a tweet. This classification is based on the content or topic of the document or text. The second form is named *entity recognition* (NER) and entity linking (c.f. linked data). This is about classifying what entities or concepts are discussed in the text or document, rather than what is the exact textual content. For example, a document about wildlife may discuss lions in India, so the text classification of the document might be "wildlife" while NER may give "Lion" and "India" as entities discussed in the document.

Text classification

The aim of text classification can be stated as: "given a document *d*, assign a class label *c* from a predefined set of categories". In simpler terms, for a passage of text (be it a document or paragraph, etc.) pick (a) label(s) from a defined list (like a dictionary) to describe what it is about.

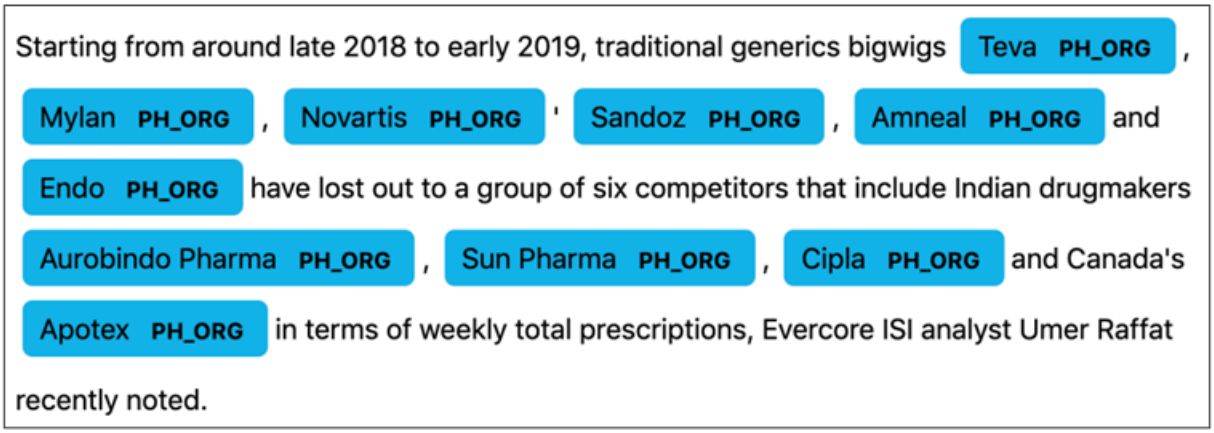
These passages of text can be long-form, such as a full article or report, or short-form, such as a comment or social media post. The classification can be formulated such that either only one or many class labels are assigned to a document; the latter is known as multi-label classification. The classification can be according to genres (such as academics papers, reports or personal communications) or topics (politics, sports, etc.), or it can be binary (simply whether the document belongs to a certain class or not). An example of binary classification is whether promotional material for a pharmaceutical product is FDA[11] compliant or not.

The standard approach to text classification is to represent text as a vector of words. The system then simply counts the occurrence of a particular word from a (pre-defined) dictionary in the vector, noting that the word order and associated meaning are lost when using this approach. Features can be identified by the presence or frequency of certain words or sequences of words. Traditional machine learning techniques, such as logistic regression or support vector machines, as well as deep learning techniques, such as recurrent neural networks, can be used to learn and discover features in data that will allow for classification. In such a machine learning pipeline, labelled text (that is text where a classification is set) is used to train a model, which learns feature representations with which to classify the text. This model is then evaluated with another set of data for testing purposes. The trained model is then used to classify data, assuming it performs well on test data.

Named entity recognition

The aim of NER can be stated as: "given a piece text, locate the word (or tokens) that correspond to entities of interest and assign an entity type". In simpler terms, given a passage of text, pick some entities that are related to it. Entity types are domain dependent; for example, generic (e.g. person, organisation, location) or biomedical (e.g. gene, protein, chemical, disease). An example is recognising the names or drugs, diseases and pharmaceutical organisations in a piece of text, as shown in Figure 3.10.

Figure 3.10. Example of NER form the PharmKE Knowledge Extraction Platform (Jofche et al., 2023[5])



[11] United States Federal Food and Drug Administration

The approaches taken for NER include:

- Dictionary based: this involves matching tokens with entities in the text. This can be limited by the inability to disambiguate between entity types. (Is JFK an airport or a person?)
- Rule based: this is based on using a set of rules to check for patterns. This requires the handcrafting of rules, which need to be updated regularly.
- Machine learning: each token is represented as a feature vector, and a model is trained to label each token in the data. This requires many labelled data for training purposes.

Entity linking

The aim of entity linking (also known as entity normalisation) can be stated as: "given a named entity, map to a corresponding concept in the knowledge base (or ontology)". In other words, take the "things" found in NER and link them to items in a knowledge base. These knowledge bases are domain dependent and can include a general domain (for example, Wikipedia) or a specific domain (for example, Biomedical). An example for the biomedical domain is the identification of COPD (chronic obstructive pulmonary disease) phenotypes identified using a NER system to terms defined in an ontology. This is shown in more detail in Figure 3.11.

Figure 3.11. Mapping of concepts recognised in NER to terms in COPD-relevant ontologies (Fu et al., 2015[6])

Automatically recognized underlying concepts	Automatically linked ontological concepts
chronic airways obstruction	*chronic (PATO:0001863) respiratory airway (UBERON:0001005) obstructed (PATO:0000648)*
parenchymal destruction	*parenchyma (UBERON:0000353) damaged (PATO:0001167)*
decrease in rate lung function	*decreased rate (PATO:0000911) lung (UBERON:0002048) function (PATO:0000173)*
chronic bronchitis	*chronic bronchitis (DOID:6132)*
myocardial infarction	*myocardial infarction (DOID:5844)*
enhanced response to corticosteroids	*enhanced (PATO:0001589) response to (PATO:0000077) corticosteroid (ChEBI:50858)*
FEV1	*Forced Expiratory Volume 1 Test (NCIT:C38084)*
alpha1 antitrypsin deficiency	*alpha-1-antitrypsin (PR:000014678) decreased amount (PATO:0001997)*

The approaches taken for entity linking include:

- Dictionary based: this involves using string similarity, between the mention of a token and the concept. However, this is limited to similarity at the surface (lexical) level.
- Machine learning: each mention and concept is represented as a feature in a vector, and a model detects a similarity between the mention of a token and the concept.

Exploiting semantic metadata for search

These techniques can be used to enhance the ability to search documents. Labels generated by *text classification* can be used to quickly filter out documents, which are irrelevant to a topic or category of interest. Similarly, labels generated by NER and entity linking can be used to help a search engine to return more precise results. A good example is the searching of a system containing data on the history of medicine for 'the common cold'. In conventional text searching, "cold" could return results about a medical condition or descriptions of temperature but using NER this search could recognise that the search is for 'cold' as a disease and so it removes irrelevant search results (Thompson et al., 2016[7]).

Targeted search with question answering

Instead of searching via keywords, searches can be made using natural language, these are known as *natural language queries*. This is implemented in some search engines, such as Google, with relevant search results to natural language questions such as "When is the coronation of King Charles III?".

Natural language queries are underpinned by machine reading comprehension (MRC) models. The aim of these models is that, given a context, for example, a passage written in natural language, and a question (also in natural language), the model should provide answers to a question. The formulation of this task takes two forms:

- Extractive – this is where the answer to the question is part of the passage and can be taken literally;
- Abstractive – this is where the generated answer is not necessarily taken from the passage literally.

MRC models learn to read text using a technique called *transfer learning*. This is where a pre-trained model that has been trained to learn a language, using lots of unlabelled text, is then fine-tuned using labelled text relevant to the domain. An example of this is a "news chatbot", in which an MRC model trained on current affairs can answer questions about the news. This can be seen in Figure 3.12.

Figure 3.12. Question-answering news chatbot

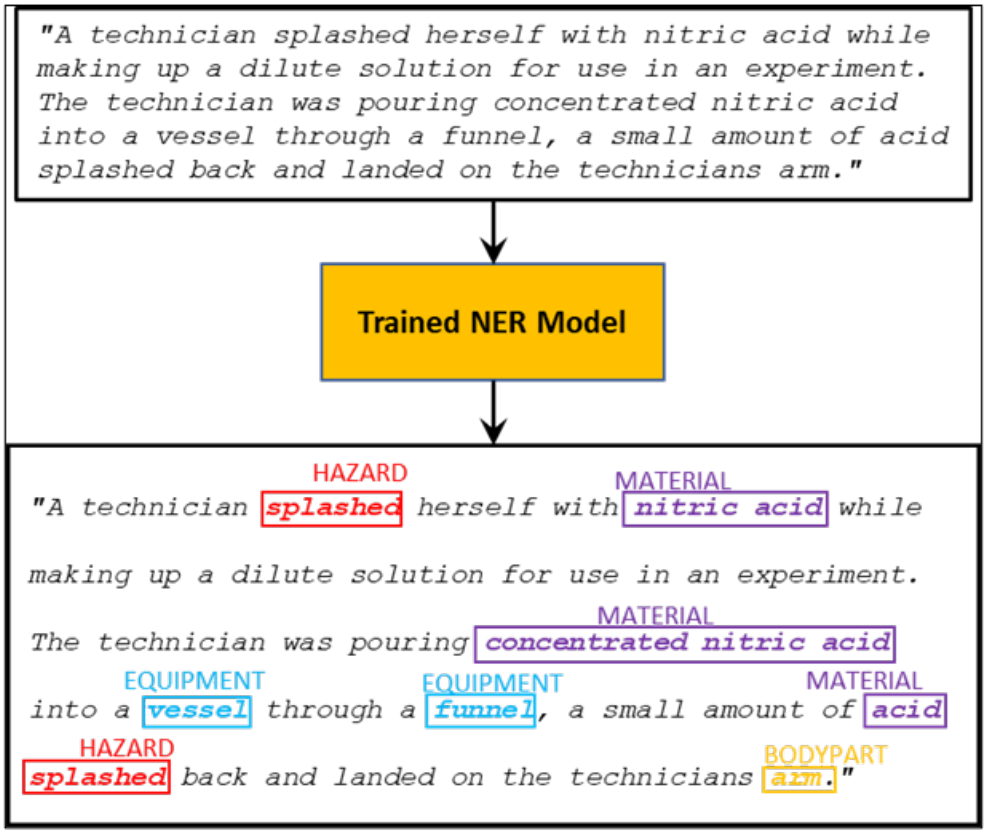
Proof of concept – NLP applied to safety reports

A relevant example of such techniques is the application of NLP to Sellafield[12] safety reports, as part of the Sellafield Game Changers Challenge. The LISTENER project (Leverage Insight from Safety Text Excerpts with Named Entity Recognition) was a feasibility study exploring whether machine learning could be applied to safety documents. The end goal was to identify automatically useful trends in order to recognise risks. It used a set of 11 000 event and 13 000 observation reports from the UK National Nuclear Laboratory (NNL), and 100 observation reports form Sellafield.

The first task for this project was to automatically label a given observation report as either *positive* or *negative*. The second task was to automatically label a given event report as one of eight event types (examples include: hazard report, injurious non-contamination, injurious personal contamination and security report). A further task, using NER, was to automatically highlight names of Materials, Facilities, Artefacts/Equipment, Body Parts, Organisations and Event triggers (Figure 3.13).

[12] Sellafield is a major nuclear site in the United Kingdom.

Figure 3.13. The NER component of the LISTENER project shows identified entities like hazards



Text classification was undertaken by dividing labelled NNL data into training and test sets, with the model being trained on the training set using a deep learning model (namely BERT). The results of the model were judged using the test set and given a score using the Matthews Correlation Coefficient (MCC)[13]. For observation reports, this score was 0.67 and for event reports the score was 0.7, where a score of 1 demonstrates complete agreement between the model and the test data. Such scores are not ideal, but this is possibly driven by the lack of available nuclear safety datasets, which is driving poor performance. Thus, the language model is not optimal for nuclear safety.

Conclusions

Keyword-based searches in digital libraries have limitations, particularly their inability to disambiguate and capture semantic variations. NLP can be used to automatically generate semantic metadata to accomplish text classification, to identify the categories of document, and NER (and entity linking) can be used to identify the concepts and entities discussed in documents. Such metadata can then be exploited by search engines to make results more

[13] MCC is a statistical tool used for model evaluation for binary classifications. It describes a measure of association between two binary variables. A score of 1 indicates perfect agreement between the two variables, a score of zero indicates only a random correlation and -1 perfect disagreement.

precise and allow for queries using natural language. NLP can be feasibly applied to data from the nuclear sector although some labelled data is needed to train models.

3.2. Session 2

The overarching topics of Session 2 on Day 1 were stakeholder perspectives and national use cases. Five speakers gave presentations on perspectives of potential stakeholders of digital safety cases as well as RWMO's current work on digital safety case systems.

3.2.1. Main takeaways from the Digital Safety Case Questionnaire

Olivier De Clercq, from the Federal Agency for Nuclear Control (FANC) in Belgium, a Bureau member of EGSSC, shared the results of a questionnaire on digital safety cases in RWMOs.

Prior to the workshop date, the NEA Integration Group for the Safety Case (IGSC), the Forum for Stakeholder Confidence (FSC), and the Working Party on Information, Data and Knowledge Management (WP-IDKM) issued a questionnaire on digital safety cases to participating organisations. The template questionnaire can be found in 0. The NEA received responses from 14 organisations on the IGSC/IDKM questionnaire: nine of these where waste management organisations or implementers, four were from regulatory bodies, and one from a research organisation. The set of organisations that responded, the country of origin, and the type of organisation is shown in Table 3.1

Table 3.1. The answers received from organisations ; country of origin, and the type of organisation

Organisation	Country	Organisation type
FANC	Belgium	Regulator
ONDRAF-NIRAS	Belgium	WMO/ Implementer
NWMO	Canada	WMO/ Implementer
SÚRAO	Czech Republic	WMO/ Implementer
ANDRA	France	WMO/ Implementer
BASE	Germany	Regulator
Sogin	Italy	WMO/ Implementer
JAEA	Japan	Research Organisation
NUMO	Japan	WMO/ Implementer
SSM	Sweden	Regulator
SKB	Sweden	WMO/ Implementer
NAGRA	Switzerland	WMO/ Implementer
US Nuclear Waste Technical Review Board / US NRC	United States	Regulator
US DOE	United States	WMO/ Implementer

The key points expressed in the response were that, in many cases, safety cases are planned as a suite of documents and supporting material written in and managed by traditional Microsoft Office tools (Word, Excel, SharePoint, etc.). The indications are that many organisations use manual processes and digital tools to support only parts of the safety case production/review processes. Furthermore, less than half of respondents raised ideas under "interesting topics", which might indicate that a clear vision on how to pursue further drastic digitalisation is currently missing. Additionally, there was also no detailed answer

on how consistency of data/information/knowledge in the safety case is assured. Consistency assurance gets more difficult as programmes progress and the amount of data/information/knowledge to be dealt with increases. Furthermore, the view was expressed that the level of detail of digital implementation should be evaluated according to a costs/benefits ratio, and according to the requirements and objectives of the programme stage. Additionally, before a full digitalisation is undertaken, the safety case methodology and argumentation structure should be stable.

Responses to specific questions

B1 Modelling of information flow

Highlights of organisations response on how information flows are modelled are:

- **NAGRA** uses a handbook for the handling of uncertainty, including uncertainty propagation along the process modelling chain.
- **US NRC** has an all-electronic licensing system, comprised of:
  - ADAMS: "Agencywide Documents Access and Management System" compiling all information related to licensing, which is searchable;
  - LSN: "Licensing Support Network";
  - EHD: "Electronic Hearing Docket";
  - SharePoint;
  - Separate NRC network at the implementer office holding output GoldSim files from performance assessments.

B2 Use of metadata and management of information

Highlights of responses on the use of metadata in information management included:

- **NAGRA** uses metadata for version control and keywords to search archives. A Digital Requirements and Configuration Management (RCM) system is in development.
- The **US NRC** uses an all-electronic licensing system (ADAMS, EHD, LSN and SharePoint), ensuring version control. Version control of each document within SharePoint is done by a document control manager. Access to metadata is limited to assigned staff reviewers, whereby staff and managers are responsible for approval.
- **Sogin** published the National Map of Potentially Suitable Areas exploiting the GIS technology, and furnished it with shapefiles (georeferenced file) and metadata on the ground of INSPIRE Directive 2007/2/EC.

B3 Use of glossaries and controlled vocabularies

For the use of glossaries, **NAGRA** has a two-step process for fast prototyping and adding terminology to an organisation-wide glossary. This promotes uniform terminology across teams. The collection of terms is bilingual, given the Swiss context.

B5 Structure of the safety case

Highlights of organisations' responses regarding the structure of the safety case included:

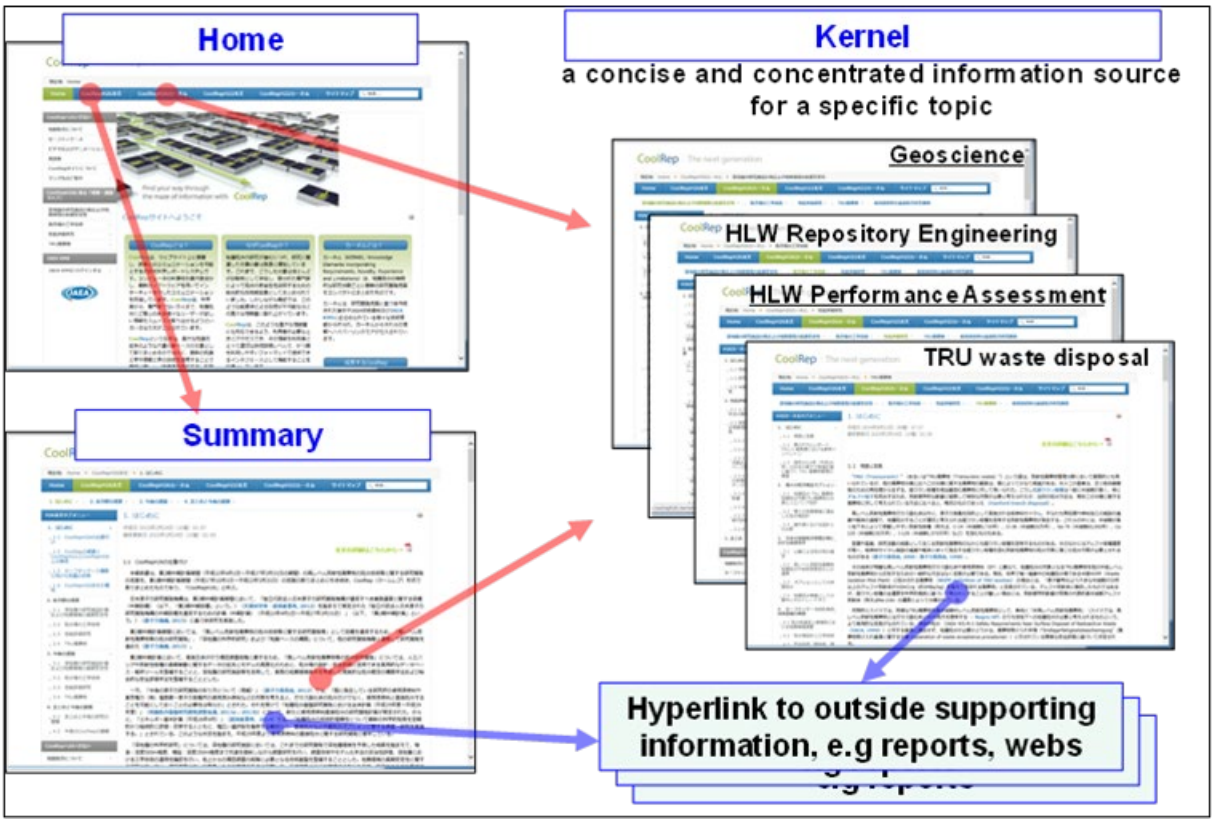
- **NAGRA** has an argumentation structure that is independent from actual arguments used in the safety case.
- **ONDRAF-NIRAS** are structuring their safety case according to a hierarchy of "safety statements" supporting the key statement that "the disposal system is safe, can be constructed and operated safely". It provides a global picture of the safety case in progress. The digital implementation and maintenance of such a structure appeared to become heavy quickly, particularly when the safety case methodology is still under development and the argumentation structure is subject to regular changes. Today, tracking is done in spreadsheets.
- **US WIPP** stated that besides the overarching QA programme, there is no true WIPP safety case logical hierarchy. The regulatory body provides guidance to prepare documentation by chapter to match the individual regulatory sections.

C1 Content management systems

Highlights of responses on content management systems included:

- **NAGRA** report a heavy dependence on typical Microsoft Office tools (Word, Excel, Access, Visio, SharePoint). In addition, they reported the development of a graphics management system that handles and co-ordinates all illustrations and graphics at the organisational level.
- **ONDRAF-NIRAS** reported their knowledge management engine (KMS), which is developed essentially for the management of the documents. There are no direct links between the models, the database and the reports. The main attention in the KMS is thus on proper referencing of all information and good traceability of the objects (reports, database and models).
- **JAEA** described CoolRep, a tool for R&D content management. CoolRep provides a portal function to manage access to other knowledge contents, such as visual support material (including graphics, videos and animations), relevant external websites, knowledge management tools and knowledge bases, etc. Three versions of the CoolRep (CoolRepH22 opened in 2010, CoolRepH26 opened in 2014 and CoolRepR4 opened in 2022) are currently available in Japanese (https://kms1.jaea.go.jp/CoolRep/index.html). An overview of this tool is displayed in Figure 3.14.

Figure 3.14. CoolRep structure (NEA, 2022[8])



C2 Use of interactive charts or infographics

Responses on using interactive charts or infographics included the following highlights:

- **JAEA** reported that there is an interactive scheme of the Underground Research Facility present in their Information Management System, where one can drill down to the detailed information on specific experiments (layout, instrumentation, etc.).
- **SÚRAO** reported the use of interactive charts for demonstrating the results of the experiments in the Bukov Underground Research Facility (URF). Integration of this experience into the safety case is under study.

C3 Management of input data, models, scenarios and performance assessment results

Responses on the management of data, scenarios and PA results included the following highlights:

- **ONDRAF-NIRAS** uses databases for all input data to be used in safety assessment models. Links are included to reference documents, bridging the science reports with the safety assessment reports. However, there is no direct link between model results and a report in which they are used (e.g. no tag associated to a result figure to insure its traceability and uniqueness).
    - A tool has been developed to transfer the data from this database to the safety assessment model to minimise manual manipulation.
    - A model management database is kept up to date. It lists all models developed so far and the applications for which they were used.
- **SKB** uses many databases to support post-closure safety assessment, including the FEP database, the activities database and others. This is fully summarised in the SKB talk on "*Use of databases for FEP, activities and input data in SKB's post-closure safety assessments*" that was given during Session 2.

C4 Tracking review and stakeholder comments

Responses on tracking and reviewing stakeholder comments included the following highlights:

- **NAGRA** reported using central tracking of comments in Excel spreadsheets, logging and archiving of meetings (recording of notes).
- **SOGIN** does not foresee automation, but rather a systematic procedure to collect observations and technical proposals by a wide range of stakeholders.
- **US NRC** has a formalised hearing process (U.S. Nuclear Regulatory Commission, 2025[9]) with limited stakeholders. The NRC's hearing process makes it possible for the public to get a full and fair hearing on civilian nuclear matters. It uses the EHD (Electronic Hearing Docket) for participants in the hearing.

3.2.2. Improving safety case content and communication using electronic tools – Inputs from outside the safety case community

Professor Dr Klaus-Jürgen Röhlig from TU Clausthal presented selected outcomes of the TRANSENS project: "Transdisciplinary Research on the Management of High-Level Radioactive Waste in Germany" which was of relevance for the workshop. Highlights from his presentation are listed below.

TRANSENS is the first large-scale transdisciplinary project on nuclear waste management in Germany, and it involves research partners beyond academia, with the explicit aim of accessing non-specialists outside the nuclear waste management community. TRANSENS research is split into four transdisciplinary work packages (TAP). The TAP SAFE is concerned with the safety case.

The safety case is an established tool to support decisions in a repository programme, based, *inter alia*, on a variety of safety assessments. However, actors can be, and sometimes are, critical of the relevant approaches, procedures and standards, which are important for the German DGR site selection process. The TAP SAFE work package aims to answer the research question: *Can a transdisciplinary approach contribute to the optimisation of the safety case?*

To address this question, the TAP SAFE has started with a focus group involving actors with experience in safety cases to identify issues, strengths and challenges, with the goal being to extract issues to work with communities on. Then, it proceeded with workshops involving the TRANSENS citizen working group (AGBe), which includes 16 citizens from different regions of Germany. The members were selected via a process developed by social scientists considering diversity in age, gender, professional background and other characteristics, with stakeholders being excluded from the group. The participants are highly motivated and interested in the topics at stake and the group co-operates with scientists over the entire project timeline. The aim is to develop familiarity with the concept of the safety case and, consequently, to identify issues of particular interest, and elicit input and ideas on proposals or concerns about the safety case. The ultimate objective is to investigate how two-way communication between various actors and safety case experts can work and, ideally, result in proposals for optimising the safety case.

Selected interim results

The interim results from the safety case specialist focus group suggest that there are complex communication issues surrounding the safety case, which can make it inaccessible to laypersons. Ideas to address this issue include a multi-level/hierarchical reporting system with a comprehensible, central chain of argumentation that laypersons can follow, with better use of electronic documentation systems and better integration of specialised knowledge into safety cases. These aspects could be addressed by the development of digital safety cases.

Research with the AGBe citizen working group focusses its interest on the use of Features, Events and Processes (FEP) and post-closure scenario development. The FEP and scenario tools that were presented were well understood by the group, which noted the human subjectivity involved and the diverse expertise needed in producing these scenarios.

A recurring theme was knowledge management, which is another area that could be addressed by digital safety cases. In particular, AGBe was interested in FEP dependencies, their intensity, their likelihood and their evolution in time, as well as the source and evolution of knowledge, and the comprehensiveness and traceability of "discourse chains" when developing scenarios. Some members had their own suggestions about visualisation and interaction in the scope of FEP processing and scenario development, which could be implemented in electronic information management systems.

The preliminary conclusions of the exercise suggested that safety case specialists believe that, *inter alia*, there is a need for better integration of safety case components and their specialist knowledge bases. The non-specialists particularly addressed FEP processing and scenario development. Their suggestions went in two directions for which digitalisation could assist: knowledge management, and guiding the processes when working with FEPs and scenarios.

Knowledge management

Currently, FEP databases are state-of-the-art. They could become the basis of an information management system with a graphical user interface (GUI) in which

- the evolution of knowledge concerning each FEP can be documented and traced over project time; and
- the relationship among FEPs and their impact on safety functions can be documented and visualised.

The latter could include cause-effect chains and information on the likelihood of occurrence of such effects or their intensity. It could also allow for the possibility of completeness checks for FEPs. Existing tools have already implemented aspects of such a GUI, e.g. (Bailey and Billington, 1998[10]) for interactions between FEPs (Figure 3.15), and (SKB, 2006[11]) for interactions between FEPs and safety functions (Figure 3.16). However, not all of these tools are still being used – access to some of the information may be lost due to changing technologies and much has never been documented in public fora; and some were apparently either abandoned in later work, either due to changing preferences or perhaps simply due to gaps in the knowledge transfer processes.

Figure 3.15. Extract from the Conceptual Model Interaction Matrix (Bailey and Billington, 1998[10])

Shows impact intensity, and potential impact / likelihood and evolution of FEPs.

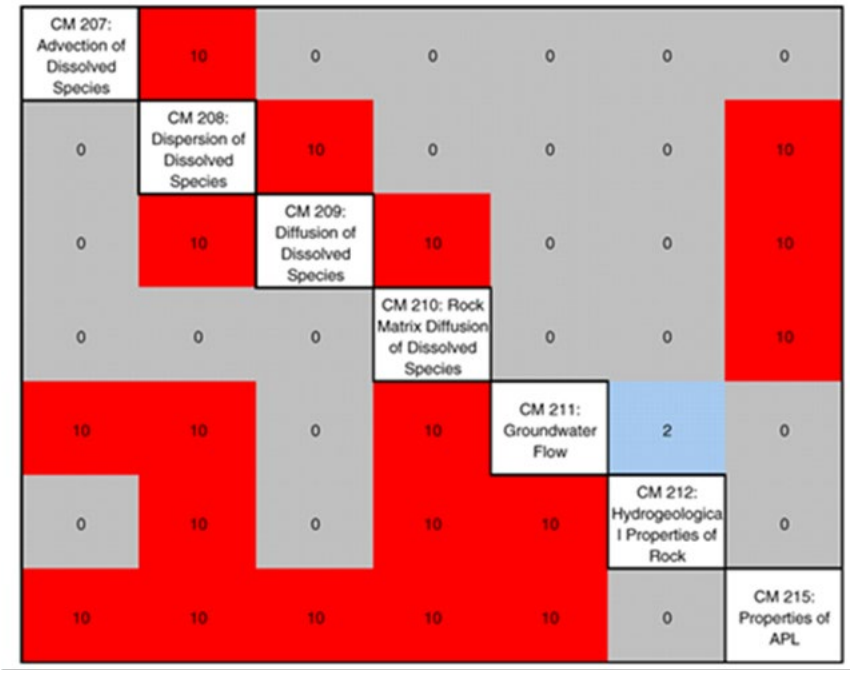


Figure 3.16. Relationships between FEPs and safety functions visualised by SKB

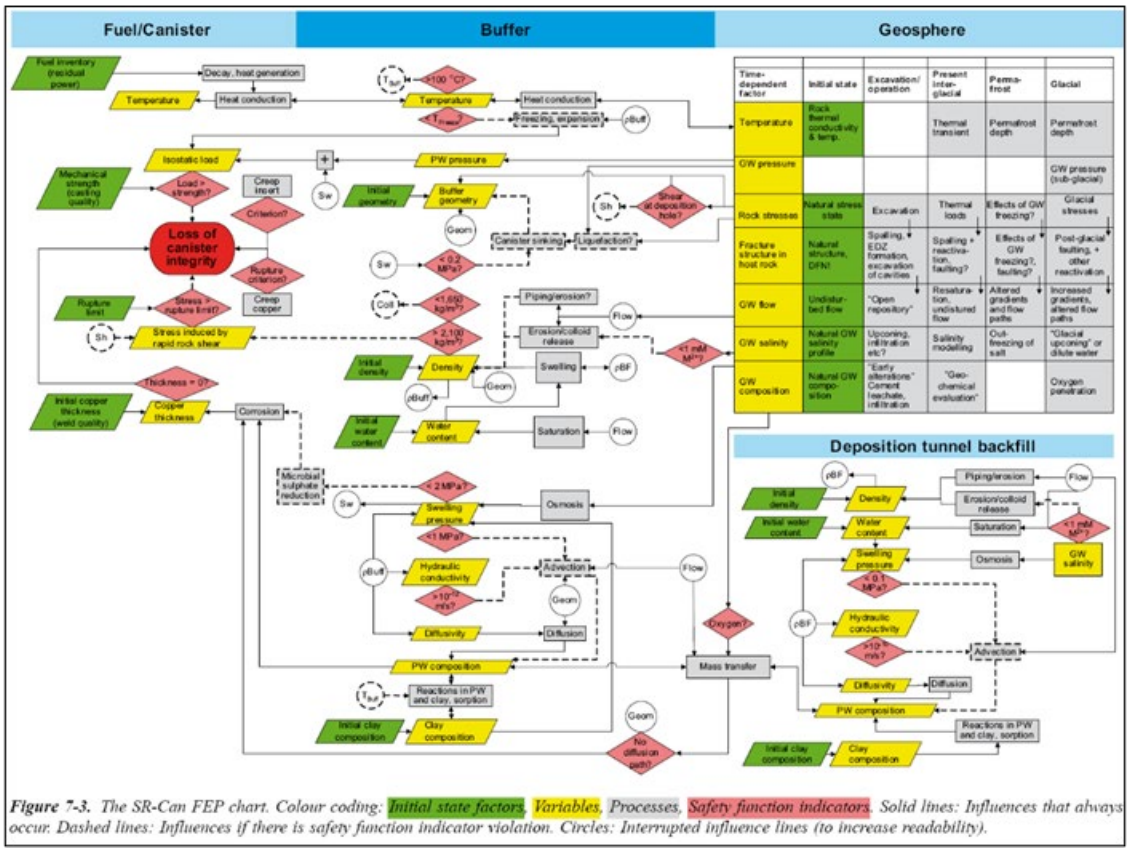
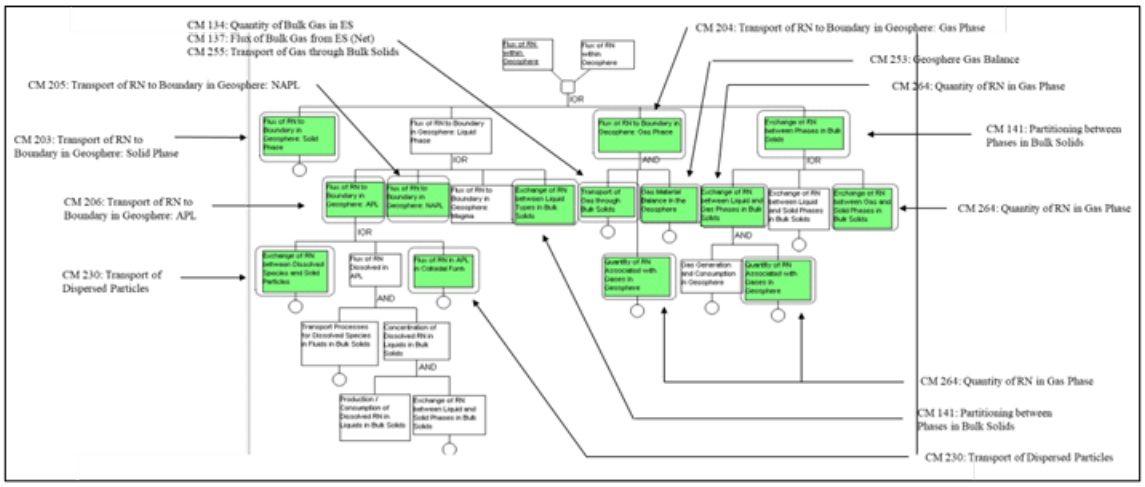


Figure 7-3. The SR-Can FEP chart. Colour coding: Initial state factors, Variables, Processes, Safety function indicators. Solid lines: Influences that always occur. Dashed lines: Influences if there is safety function indicator violation. Circles: Interrupted influence lines (to increase readability).

These tools could, however, be further developed and refined, with these visualisations generated semi-automatically from an appropriate knowledge database using an appropriate GUI, which would, in turn, serve the database and could be examined interactively by both scientists developing them and by interested audiences. Such an interactive tool could aid scenario development by guiding the technical discourse as well as recording decisions and relevant conclusions. This would ensure traceability of the scenario development process and ensure that nothing (no thought, idea or relationship) would be forgotten. A past tool from NIREX (United Kingdom), called FANFARE, already provided some of these capabilities (Figure 3.17).

Figure 3.17. NIREX (UK) FANFARE GUI



Conclusions

The transdisciplinary research presented here is ongoing, and the conclusions are preliminary. However, many issues concerning safety case production, documentation and communication that were raised during the project could be addressed by using electronic tools.

There are several important electronic tools for safety cases, as have been shown. Nevertheless, when considering the advancement of digital safety cases, there is a need to take a holistic view on safety case production and communication, and to establish the link to up-to-date technologies. Starting from there, these developments might also result in innovations around safety case production, knowledge management and communication, and might give an indication on how to develop the vision for an "ideal" safety case. For such a vision, it would be important to account for the interplay of:

- actors and their activities in a disposal programme;
- the needs and requirements from a safety case perspective;
- the potential of digitalisation (IT perspective).

3.2.3. VISI safety case management system

This chapter focusses on the work carried out by Nuclear Waste Services[14] in the United Kingdom to develop a digital tool called VISI (Visualisation of System Information) to store and present its GDF Environmental Safety Case (ESC, Figure 3.18), which was presented by Alexander Carter from NWS. The ideas within VISI and presented in this chapter are general, and therefore applicable beyond the ESC and geological disposal.

[14] Nuclear Waste Services (NWS) was formed in January 2022 by merging Radioactive Waste Management (RWM), which was planning for a UK geological disposal facility (GDF) under a community partnership-based siting process, and the Low-level Waste Repository (LLWR), which was operating a near-surface disposal facility for low-level solid radioactive waste in West Cumbria. Besides these, NWS has taken over a number of tasks related to waste treatment and management.

VISI is part of a larger vision for digital transformation within NWS. Any such system needs to be considered as a *component* of a wider digital tool set within a Radioactive Waste Management Organisation (RWMO). Such components can include requirements managements system(s), databases for site characterisation, design management tools, and waste package record systems. This chapter focusses on how VISI is used to manage the safety case narrative, in addition to its development.

Introduction to VISI

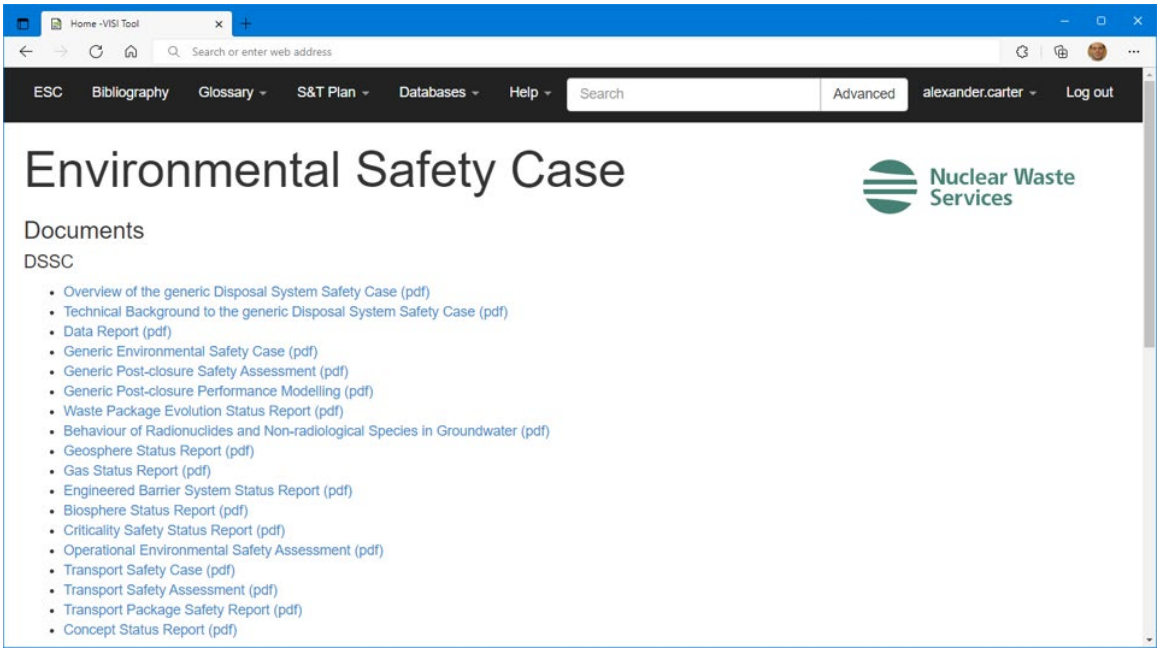
Over a number of years, NWS has advanced the development of its environmental safety case narrative through the VISI Project. The project has helped to:

- author its safety case narrative in a digital, standards compliant, format which allows content to be utilised in multiple output channels;
- structure its safety case narrative so that safety case claims are clear, and that each claim is substantiated through suitable safety arguments and evidence; and
- manage its safety case narrative, by providing a system that allows the safety case, and its underlying knowledge base, to be developed in a suitably controlled fashion.

Information in a modern context may be found on digital devices, such as tablets and smart phones, through search engines finding linked information stored on the web. However, in the nuclear industry context, most information is in older, less accessible formats like "dusty old books".

VISI is a tool developed to store the NWS safety case narrative (underlying digital "documents" and associated diagrams that show the structure of the safety case). Its home screen is shown in Figure 3.18. All documents and diagrams are searchable and hyperlinked but can still be accessed via PDF files if needed. VISI is currently on its second major version, but is still undergoing active development. The aim is that better information management should make it more efficient to update the safety case with the possibility of a future "live" electronic safety case.

Figure 3.18. VISI Home Screen



Information management

Improving data and information management often relies upon improving the ***structure*** of information and understanding the ***relationships*** between different pieces of information. These aspects require consideration both in advance of the information being created[15] and at the time when it is being created. A key concept in this is metadata, though the dividing line between data and metadata is not always clear (NEA, 2018[12]).

Metadata

Metadata can be added to the structured content within VISI, which uses an XML format called DITA (OASIS, 2012[13]), including references and keywords. In outputs used for VISI, such as webpages, this structured information and metadata is used to build links, for example to glossary keywords. Much structured content could be added to the VISI files including:

- Links to FEPs, to enable an audit that all FEPs have been considered;
- Links to glossary items or controlled vocabularies;
- Links to regulations to audit that all regulatory requirements have been met;
- Links to reference documents, to show what underlying knowledge the safety case utilises;
- Links to models / data, to show what computer models and datasets the safety case relies on;
- Links to waste packages or waste package records;
- Metadata to aid searching, archiving or helping define the Set of Essential Records (SER)[16] or about language, document publication or approval and change history.

VISI content and single source publishing

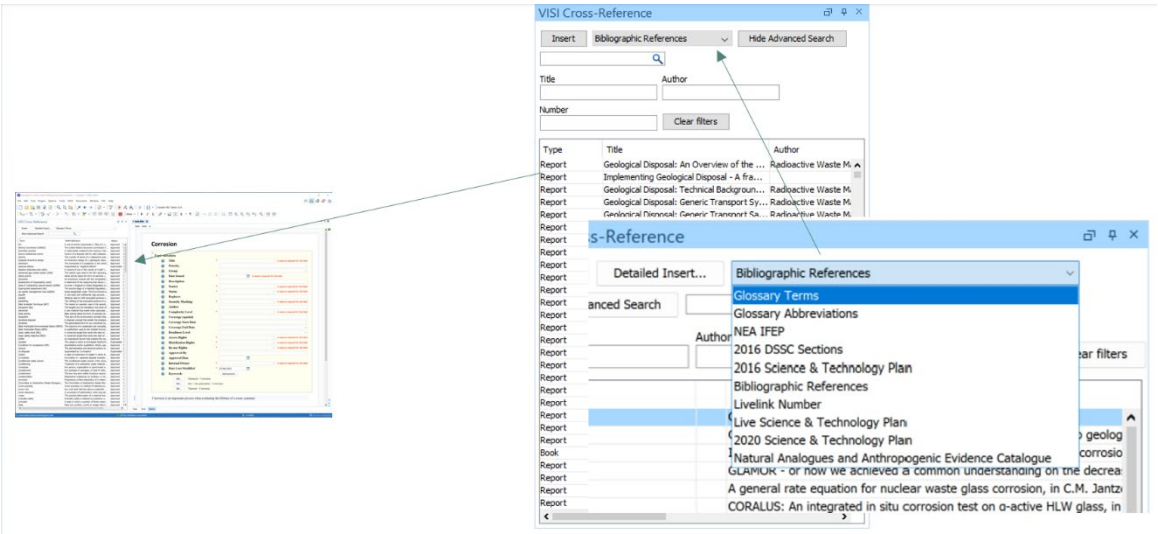
Single source publishing is about separating the content of documents from the output. Its central idea is to write once and then publish to many outputs. This approach is widely used in publishing houses, such as those responsible for producing journals. It also avoids problems with prior digital tools, such as FANFARE (Kelly and Billington, 1997[14]), in which the content and the system are tightly intertwined so when the software is discontinued, the content is lost, too. In this approach, the content written for VISI can be used for different outputs and therefore retained beyond the life of the tool.

[15] Principle 2 at the Bucharest safety case workshop: IDKM starts before the content is created. (Event website: www.oecd-nea.org/jcms/pl_32286)

[16] Principle 1 at the Bucharest safety case workshop: IDKM needs to consider the full lifecycle of the information. (Event website: www.oecd-nea.org/jcms/pl_32286)

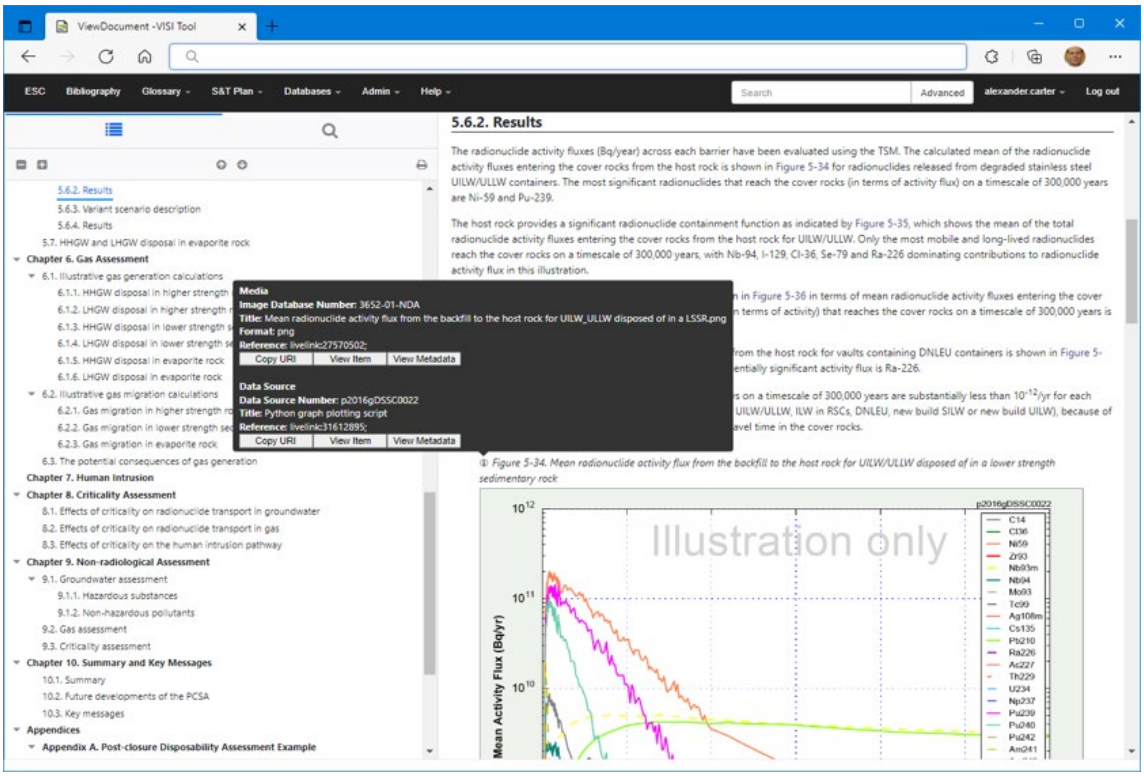
VISI content is written in the DITA, a standards compliant format based on XML, in which the single source paradigm is implemented. Content for VISI is written using the Oxygen XML editor (Syncro Soft SRL, 2002-2026[15]), this is a 'what you see is what you get' (WYSIWYG) editor for editing XML formats (including DITA) and provides a familiar environment for users used to MS Word (Figure 3.19). The editor can be installed on staff laptops, so content can be edited both online and offline. This format allows structured content to be added to documents, with metadata such as keywords added via a specialised toolbar written for use with the VISI system. This structured content includes items such as a large number of bibliographic references and glossary items, which have been complied as part of the project.

Figure 3.19. Authoring environment for VISI, Oxygen XML, with specialisations that add structured content



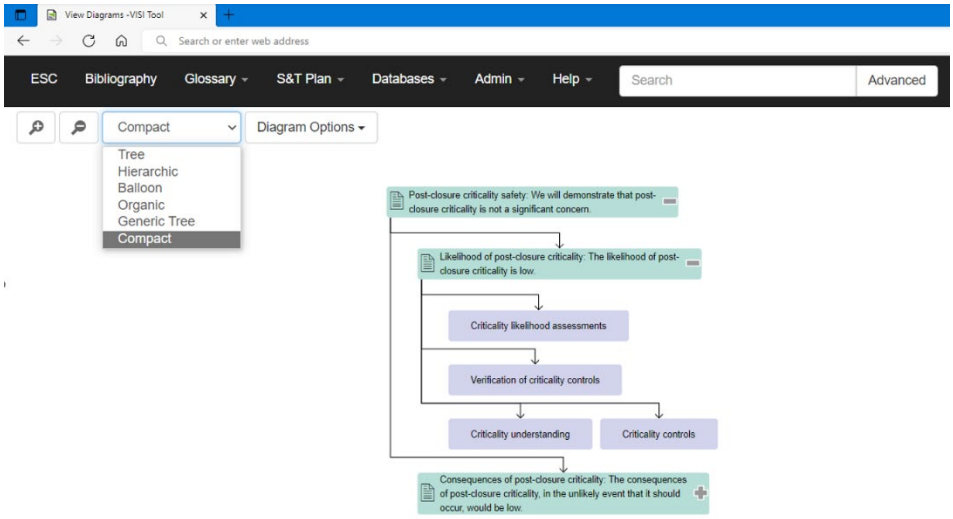
In the VISI system, this content is transformed into webpages accessible from the home screen. The structured content added, such as glossary items or bibliographic references, can be used to populate reference links and popups in the webpages, as can be seen in Figure 3.20.

Figure 3.20. Content displayed in VISI



NWS is currently exploring the use of Claims, Argument and Evidence (CAE) to structure its safety case (Figure 3.21). VISI can plot diagrams of CAE to structure each "node" linked to appropriate underlying content and is accessible by clicking. The engine which plots these diagrams does so dynamically and could be used to visualise any type of relationship.

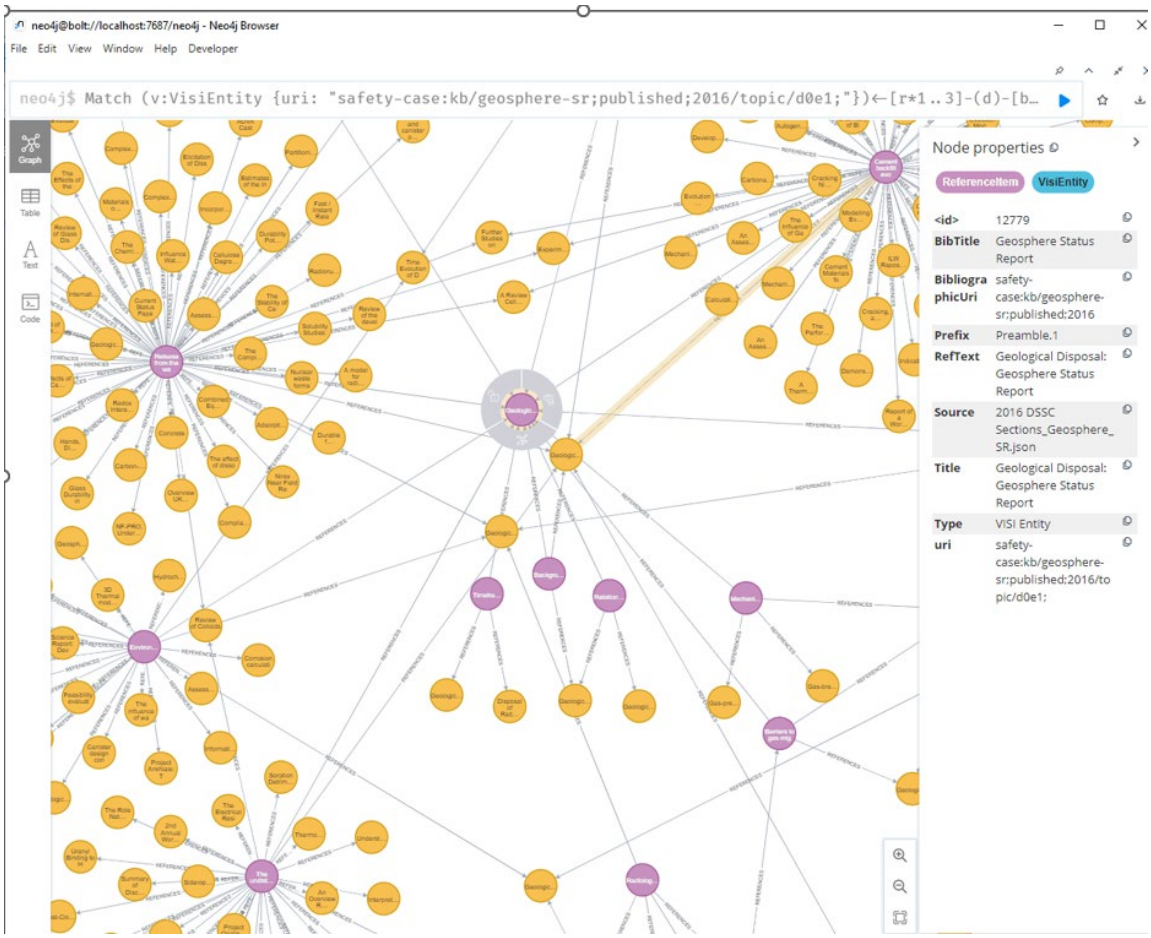
Figure 3.21. Claims, arguments and evidence diagram in VISI



Potential future developments

The VISI system is still under development. Potential future developments include version controlled file storage with a workflow engine and electronic signing; features that may be realised through commercially available tools coupled into the existing system. Further development will allow VISI to connect to other digital tools and vice versa, via appropriate APIs. This, together with the wider digital tool set, would allow VISI to handle key safety case processes, such as change control, approval, and publication, in an efficient and auditable way[17]. Another key potential development for VISI is the transfer of the VISI data graph into a graph database in order to utilise the associated toolset, such as the technologies discussed in the *Linked Data* and *Knowledge Graph* presentations, to search and visualise VISI's web of knowledge. These are significant undertakings which will require detailed thought and planning.

Figure 3.22. A snapshot of VISI data plotted in a graph database (Neo4J)



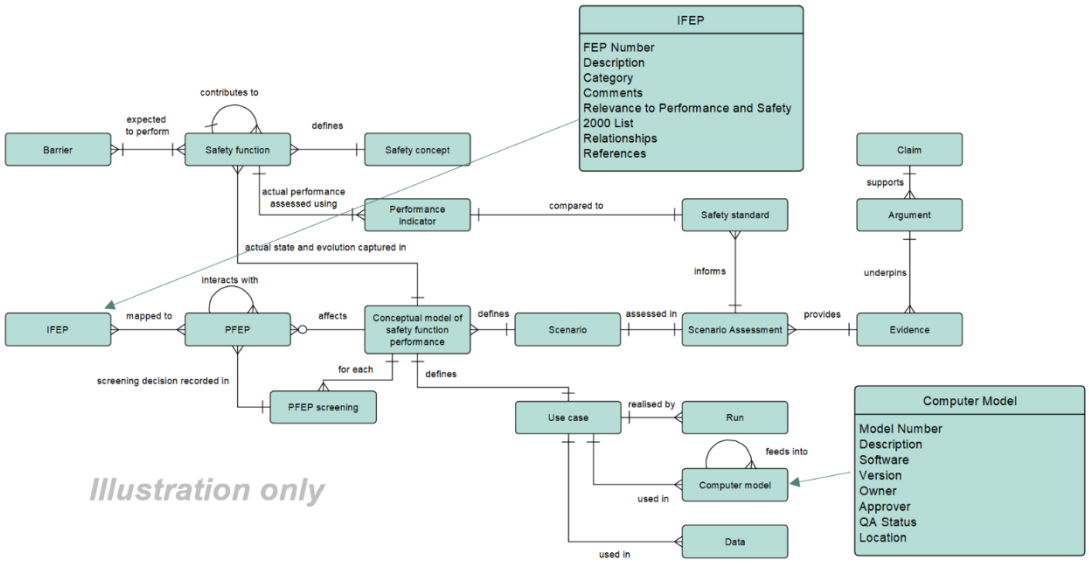
[17] Principle 4 at the Bucharest safety case workshop: IDKM works best when people, process and tools work together. (Event website: www.oecd-nea.org/jcms/pl_32286)

Thoughts and reflections on digital safety cases

A key question on the development of digital safety case systems concerns users and use cases for such systems now and in the future. Key users of such systems include a diverse array of actors including safety case professionals, subject matter experts, managers and key decision makers in radioactive waste management organisations, staff in technical specialist organisations supporting RWMOs, regulators, members of the public and archivists. At the highest level, the use cases and requirements for all these users could be developed holistically in international fora, such as the EGSSC or IGSC.

The benefits of digital tools are fully utilised when they are being considered simultaneously with methodology development. For example, the use of CAE to visualise the safety case and ensure completeness can also allow the reuse of content (such as evidence) in multiple places. This points to a need to develop an appropriate data model for a safety case, which can be used to inform the types of cross references, enables validation, and underpins creation and search of knowledge graphs. Safety case data models can be used to support the tasks undertaken by related working groups, such as the NEA IGSC follow-up project to MeSA (NEA, 2012[16]). A simple, illustrative example of a data model, based on the UK Deep Geological Repository (DGR) programme, is shown in Figure 3.23.

Figure 3.23. Illustrative example of a data model for a DGR safety case



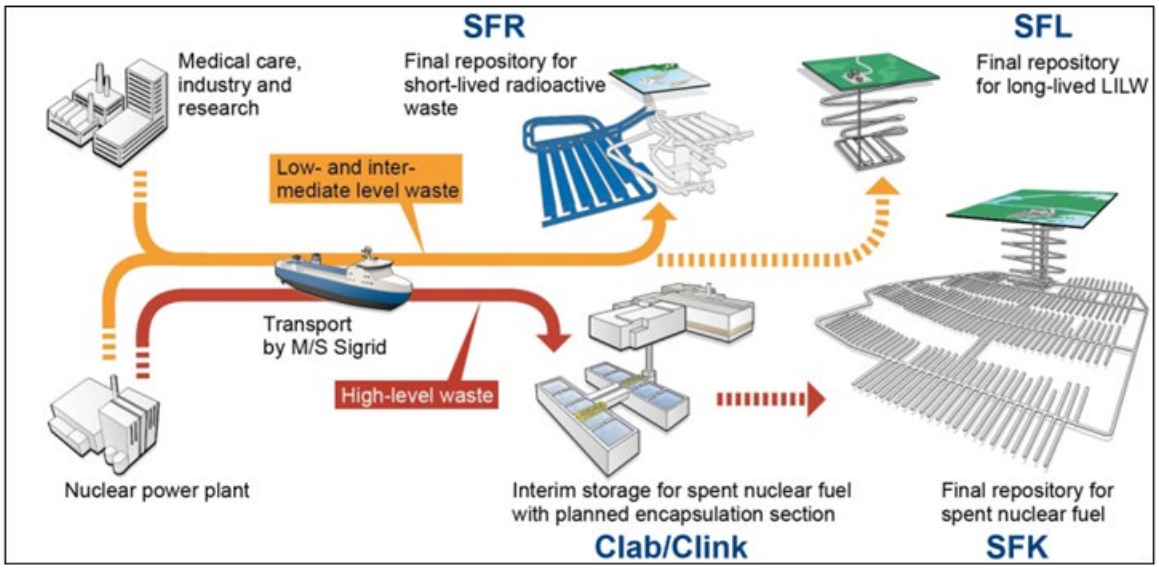
3.2.4. Use of databases for FEP, activities and dataflows in SKB's post-closure safety assessments

The Swedish system

Niko Marsic, from SKB in Sweden, provided an overview of their databases for FEP, activities and input data, in their post-closure safety assessments.

The high-level approach for the management of radioactive waste in Sweden is summarised in Figure 3.24. Waste from various sources, such as medicine and nuclear power, is disposed of in repositories for short-lived waste (SFR), long-lived low- and intermediate-level waste (SFL), and spent nuclear fuel (SFK), respectively. The licence applications for both the SFR and SFK repositories have been approved.

Figure 3.24. The high-level approach for the management of radioactive waste in Sweden



SKB databases

SKB uses several databases for different purposes. These are:

- FEP database;
- Activities and dataflows database;
- Investigation database;
- GISADA (formerly SICADA) – used for primary spatial data storage;
- Labware LIMS – used for primary non-spatial data storage;
- Model database (formerly SIMON) – used for model storage;
- Subversion (SVN) (version control software) – used for input data and models for radionuclide transport and dose calculations;
- SKBdoc – used for documents and data storage (version controlled).

The above databases are generally not coupled or integrated.

FEPs - Features, events and processes

In the SKB methodology for safety analysis, a systematic identification of all factors and interactions of importance for the post-closure safety of a repository for radioactive waste is required; these are the features, events, and processes (FEPs). Identification of FEPs is the first step in this methodology and as such has a central role in post-closure safety assessments. FEPs are repository specific with a catalogue developed for each one. The catalogue and its methodology and processes are documented in a FEP report and summarised in the post-closure safety report. The FEP catalogues are included in SKBs FEP database. For each update of a safety assessment, the FEP catalogue is updated, as is the FEP report documentation and the FEP database.

SKB FEP database

The SKB FEP database contains the FEP catalogues for current and historic SKB safety reports, such as SR-97 (SKB, 1999[17]), SR Can (SKB, 2006[11]), SR Site (SKB, 2011[18]), SR PSU (SKB, 2015[19]), and SE-SFL (SKB, 2019[20]) FEP catalogues. It also includes project specific FEP lists from the NEA and the outcomes of FEP audits against the NEA FEP lists, including mapping and interaction matrices. The FEP database also records couplings of FEPs to system components (such as processes and variables); this includes influence tables, process diagrams, interaction matrices and assessment model flowcharts. The FEP database is developed and maintained in FileMaker Pro Advanced (Claris International Inc., 2026[21]), with runtime versions produced for SR Site (SFK – Spent Fuel) and SR-PSU (SFR – short-lived LLW) submissions.

The FEP database is a conventional relational database system; it has had a recent update to its graphical user interface (GUI) for the latest catalogue. It includes information from previous catalogues for past submissions, including some information on data flows from older FEPs and mapping to the more recent catalogues, although this has not included some of the oldest FEP catalogues.

The user is able to select the relevant FEP catalogue (Figure 3.25) along with mapping and interaction matrices they may be interested in. On selecting an FEP catalogue, they can select categories of FEPs which they wish review (Figure 3.26, left-hand side). Individual FEPs can then be selected (Figure 3.26, right-hand side), bringing up FEP records (Figure 3.27). These records show fields collected for each FEP. They also contain records for screening judgements and quote the report the FEP is used in, which is the only part of the database that provides such cross-linking.

Figure 3.25. SKB FEP database GUI

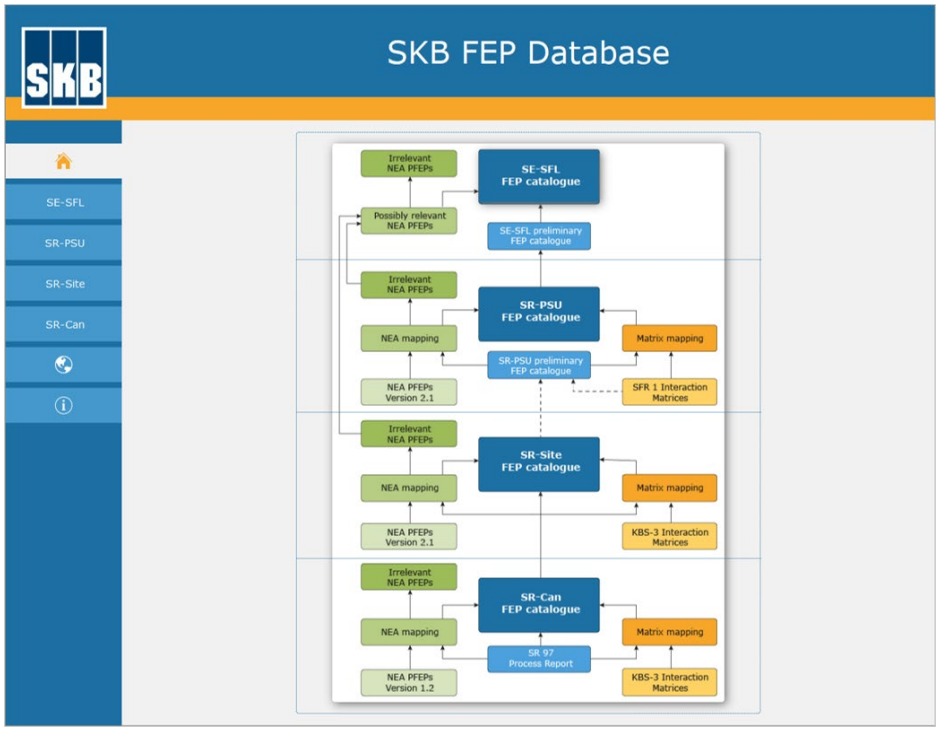


Figure 3.26. Select FEP categories in SKB FEP database

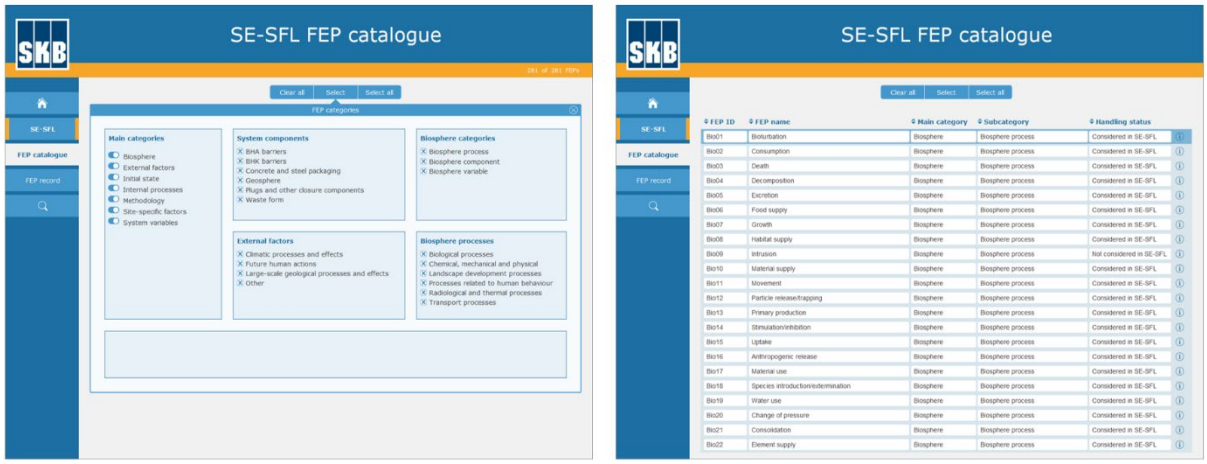


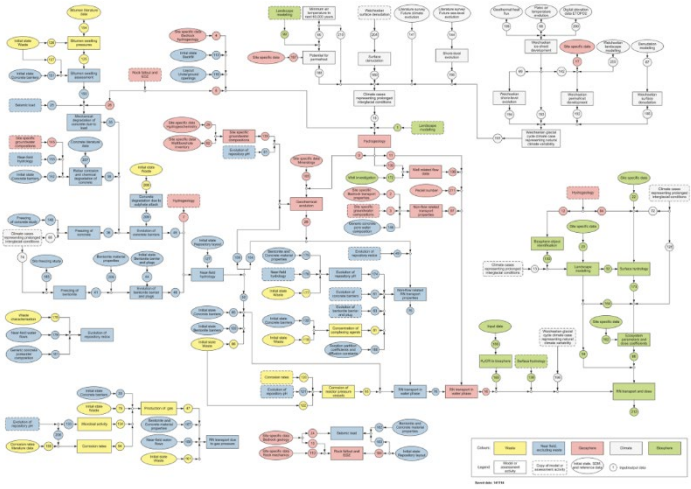
Figure 3.27. A FEP record in the SKB FEP database

Mapping of activities and dataflows in post-closure safety assessments[18]

Safety assessments are accompanied by a report which aims to show the flow of data in the assessment. The objective is to map input data and the flow of data between different activities, within the post-closure safety assessment. Flow charts have been produced to give an overview of the models used, the dependencies/interactions between them, and the data used in the modelling in the previous safety assessments for SFR and SFK. In these reports, the activities and dataflows are also described in written text. This leads to repetitive structures with the same information in several places, in the graphical form of flow charts and written prose. This creates issues with QA, as there is a higher risk of inconsistencies, and a higher cost of making changes and updates as changes must be made in several places. An example of the previous assessment flowcharts is shown in Figure 3.28 for SR-PSU, the previous safety submission for SFR.

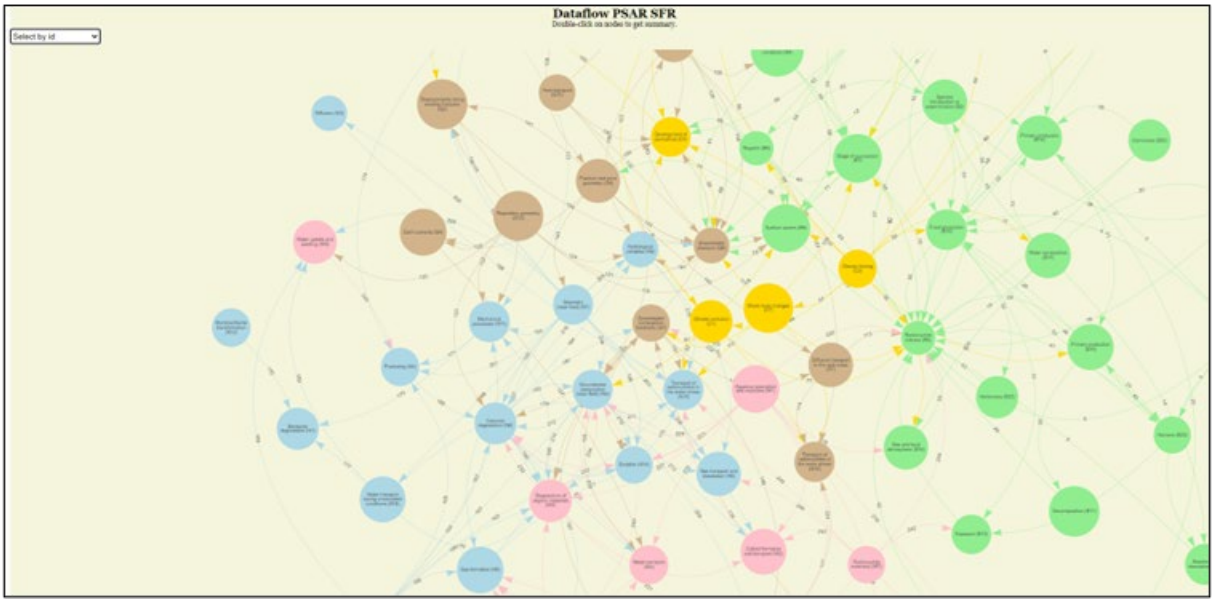
[18] With thanks to Albin Nordström

Figure 3.28. Safety assessment flowchart for SR-PSU (SKB, 2015[19])



To address this issue in ongoing work for the next SFR safety submission, a database is being used, in conjunction with flowcharts of data flows and activities, to automatically generate the safety assessment flowchart. Information is entered once, directly in a database, and not restated. These then provide an overview of the safety assessment activities with close-up views of details, allowing for traceability and informing planning activities. This includes new visualisation capabilities with the optional ability to automatically generate the (dataflow) report.

Figure 3.29. The auto-generated interactive flowchart for FEPs and dataflows for SFR

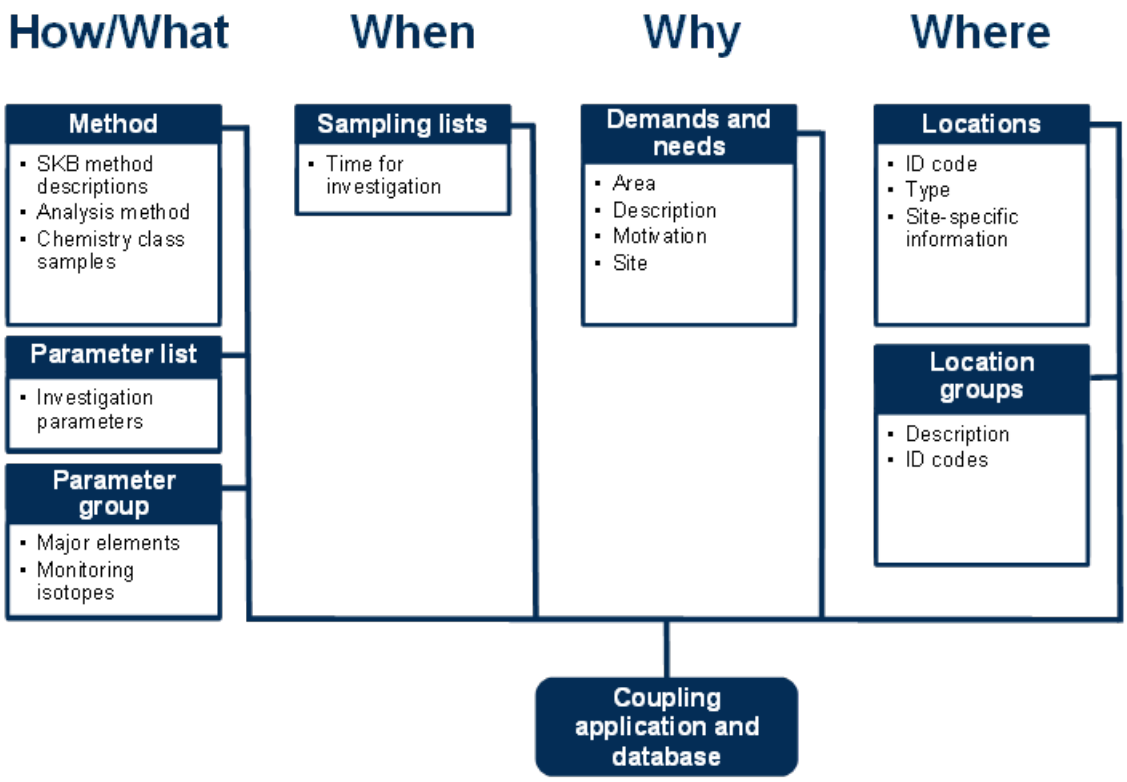


The interactive flowchart is in the form of a nested webpage with the components of the whole system visible at the top-level page (Figure 3.29). The nodes present clickable links to level 2 reports on subsystems (such as climate), which are linked to level 3 activities (such as climate evolution), which are displayed in “pop up” tables.

Investigation database[19]

The investigation database stores the information on data needs and requirements from investigations; this includes requirements from the Site Descriptive Model (SDM), safety assessments and engineering. The database system stores the "How, What, When, Why and Where" of these needs, the requests for data, and their delivery by site investigations and monitoring. This product is compiled as metadata in a MS Access database, the structure of which is shown in Figure 3.30.

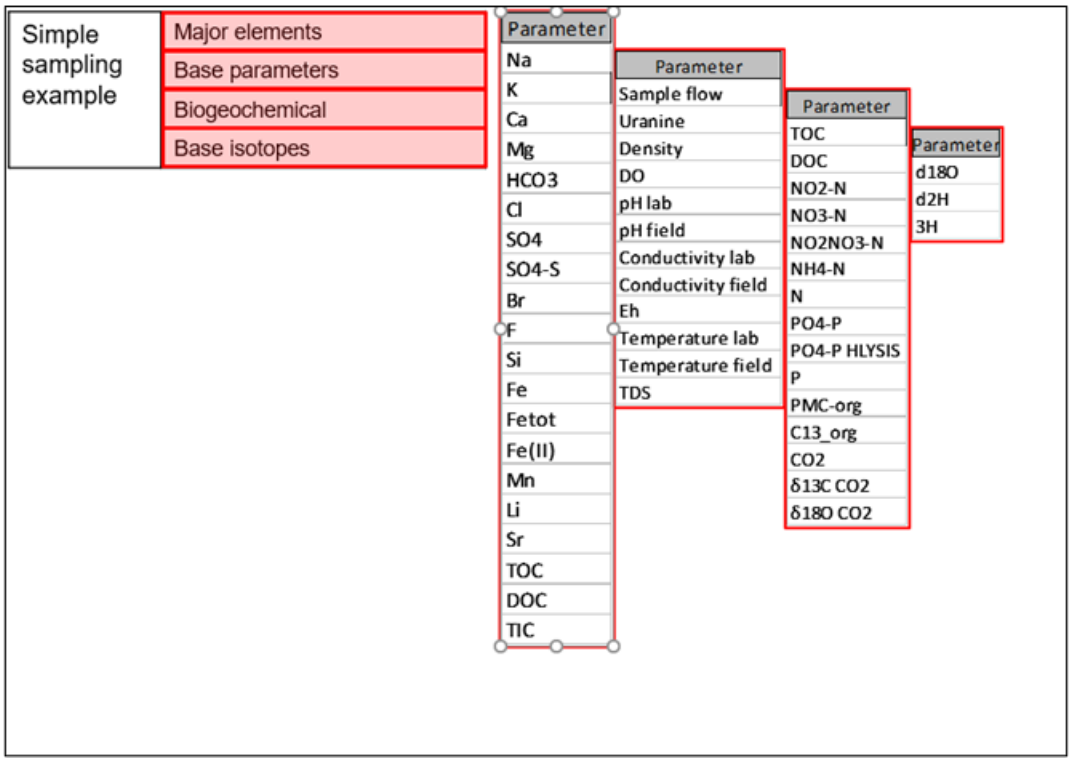
Figure 3.30. Structure of the SKB site investigation database



This system has the benefits of allowing traceability for site investigation needs, and allows for better-structured collaboration, flexibility and knowledge transfer. A simple example for use of the database is shown in Figure 3.31, for the sampling of the surface system parameters. This includes a list of parameters as part of the surface monitoring, for instance a collection of redox measurements, and sampling locations, such as stationary boreholes or lakes in a candidate area. The database does not store the monitoring data but is used to produce activity plans and sampling programmes to collect the relevant data. These activity plans are then fed into SKBs administrative system.

[19] With thanks to Lino Nilsson

Figure 3.31. Simple example of surface parameters in the SKB site investigation database



Future work

Future work on the SKB database will focus on the continuing development of individual system features as well as increasing the coupling between databases. This includes:

- a redesigned interface for the SKB FEP Database and updated FEP catalogues;
- the coupling of the FEP with the activities and dataflows databases;
- coupling of the databases to the reports on processes;
- creating a web-based solution for wider distribution.

The updated methodologies presented here and used for the SFR update can be applied to other projects and reports.

3.2.5. Regulatory perspectives on the use of Digital Safety Cases – requirements, expectations and independent Safety Case review

Olivier De Clercq from Belgium's Federal Agency for Nuclear Control (FANC) presented the viewpoint of a regulator concerning digitalisation of the safety case. He started by quoting the general purpose of writing safety cases (Why is it done? What do they demonstrate?), how digitalisation approaches can facilitate reaching these purposes and which kinds of regulations currently exist that might restrict the use of digitalisation approaches. The conclusion was that, although many regulatory requirements exist for the content of the safety case, few strict requirements on the format (structure) of the safety case exist. As a consequence of this, an implementing organisation has freedom in its choice of digital tools and digitalisation approaches.

The importance of having frequent interactions between an implementer and a regulator during the pre-licensing phase was stressed.

During the presentation there were examples given from the development process of the surface disposal facility for short-lived low- and intermediate-level waste in Dessel, Belgium. He also gave some examples of lessons learnt from FANC's participation in different international projects, such as NEA IDKM or SITEX.

The presentation ended with some comments on digitalisation of safety cases from Dr Felix Altorfer, Director of the Waste Disposal Division of the Swiss Federal Nuclear Safety Inspectorate (ENSI) and co-chair of NEA Regulators' Forum.

Introduction: General purpose of writing a Safety Case

A key study from Dr Thomas Kelly (Kelly, 1998[22]) was quoted that describes the challenges of safety case development. The study encompassed many fields or industries requiring safety cases to demonstrate the safety of their systems (including, for example, medical devices, military equipment, aircrafts and the chemical industry). Kelly identified the key problem areas for the development of robust safety cases as:

- Presentation of clear safety arguments
    - A safety case should present a clear safety argument but in practice this is rendered difficult due to the size and complexity of the safety case; the difficulty in ensuring coherency when integrating and presenting results from multiple sources of evidence and contextual material; and the use of free-format text that often allows for unclear, ambiguous and misleading arguments and obscuring the line-of-reasoning for the central safety argument.
- Incremental safety case development
    - Best practice is to start with the development of a safety case in parallel with system design, not as an after-the-fact safety demonstration exercise (which in many fields is often the case).
- Through-life safety case maintenance
    - A typical regulatory requirement in many fields is to update the safety case "whenever appropriate", but this requirement is not always clear. Therefore, there exists a risk of not documenting important changes or modifications that are made to the system itself.
- Supporting trustworthy safety case re-use
    - Although the details differ in each case, there is often commonality between the types of arguments that are used in different safety cases. It would be desirable to have an approach that supports and facilitates the reuse of common safety arguments, while minimising the risk of fallacious arguments due to inappropriate reuse or a lack of traceability.

Kelly had said that safety cases have traditionally been thought of as written documents since their introduction in many sectors. For example, the UK Ministry of Defence's Ship Safety Management Handbook JSP 430 (UK Ministry of Defence, 1996[23]) states that "A safety case is a comprehensive and structured set of safety documentation…".

Kelly argues that this should have stated "a safety case is commonly presented using, and communicated through, a structured set of safety documentation" (Kelly, 2008[24]). This distinction is important. As the demonstration that the system is safe does not happen just by having a large collection of paper documents, word processor files or HTML documents; rather, safety is demonstrated through a structured, compelling argument with relevant and comprehensive supporting evidence (Kelly, 2008[24]). Furthermore, according to Kelly: "The safety case approach is not working if the goal is simply the production of a document*"* (Kelly, 2008[24]). This is of clear importance when considering a digital safety case, as the safety case is not simply the production of a document.

When considering digital safety cases, safety case producers should not pursue digitalisation as an end goal in and of itself. Concretely, the end goal is to produce a safety case, i.e. the compelling argument that the system is safe, backed up with a comprehensive evidence base. The digitalisation effort should be done with the end goal of improving the structure of safety arguments within it, and their presentation and communication, both within the implementer's organisation and to stakeholders. As such, digitalisation offers a number of opportunities to enhance safety. These include ensuring consistency and traceability throughout the safety case and facilitating modifications to it. Additionally, digitalisation should facilitate ease of access to the existing safety requirements and show that these have been met, with appropriate justifications.

Regulatory requirements on digitalisation of the safety case

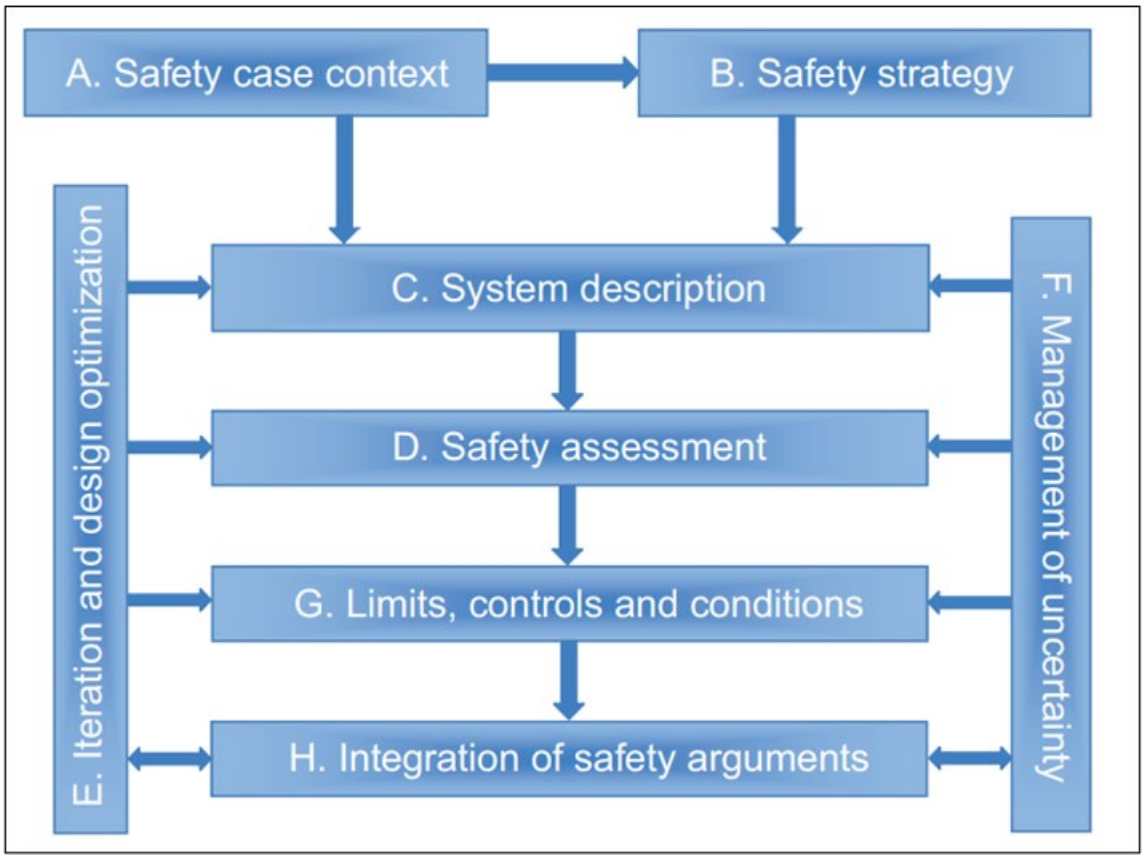
In general, existing regulatory requirements prescribe that disposal activities need to be justified (justification principle) and supported by a safety case (principle of demonstration of safety).

The safety case encompasses a wide range of topics, including:

- Legal framework of the activity;
- Safety optimisation – as justification of decisions and choices against alternative options;
- Management system of the implementing organisation;
- Handling of safety issues through:
  - The operational and post-closure repository phases;
  - Safety assessment and management of uncertainties;
- Characterisation and monitoring approaches.

The IAEA Safety Guide SSG-23 provides guidance and recommendations on meeting the safety requirements in respect of the safety case, and supporting safety assessment for the disposal of radioactive waste. It is the basis for many of the existing national regulations regarding safety cases for geological disposal facilities. The SSG-23 describes the general components that a safety case should contain and the overall structured framework for these components (see figure below):

Figure 3.32. Typical contents of a safety case (from IAEA (2012[25]))



The Western European Nuclear Regulators Association (WENRA) was established in 1999 to develop a harmonised approach to nuclear safety and to provide an independent capability to examine nuclear safety in member countries, as well as to exchange experience and discuss significant safety issues. In 2014, WENRA's Working Group on Waste and Decommissioning (WGWD) published a report on safety reference levels (SRLs) for radioactive waste disposal facilities that are intended to provide a basis for future harmonisation on a European level (WENRA, 2014[26]).

WENRA SRLs have some requirements around record keeping and archiving, including DI-20 and DI-21 below.

> **DI-20**: "The licensee shall ensure that, during design, construction, commissioning, operation, closure and post‐closure until termination of the licence, knowledge and records important to safety are available and updated as appropriate for current activities, safety assessment and long-term record keeping on:
>
> - Characterisation of the site
> - Design basis
> - Design
> - As built construction of the disposal facility

- Operation including any operational occurrences and accidents
- Inventory and emplacement of the waste
- State of the disposal system after closure
- All documents relating to the safety case"

DI-21: "The licensee shall ensure that, at termination of the licence, records on the site and on the as closed state of the disposal facility are available for continuing preservation, in accordance with the national legal and regulatory framework." (WENRA, 2014[26])

In general, although many requirements on the content of a safety case are imposed, few requirements exist on the choice of format for a safety case. There is thus an inherent flexibility for the implementer in the choice of format and tools for the development of the safety case. However, it is important that implementers have frequent interactions with regulators and other stakeholders such as NGOs, local communities or the wider public in order to make sure that the needs and expectations of these parties are considered during the development of the (digital) safety case.

Pre-licensing phase and regulator-licensee dialogue

FANC mentioned the importance of having a strong and continuous dialogue between the regulator and implementer in order to get agreement on the desired structure of a final (digital) safety case in the absence of strict regulations.

Comments from the Co-chair of the NEA Regulators' Forum

For the Swiss case, there are no specific regulatory requirements[20] regarding the delivery of documentation in a digital form. Currently, the Swiss implementer (NAGRA) provides data in a digital form to facilitate the regulators' reviews. All documents are made public in the form of PDF files. Regulatory expectations are for the safety case to be presented in a clear and concise way and while there is no requirement that this is part of a digital safety case, the regulator expects that the data and findings that support the central arguments in a safety case can quickly be found in the documentation. Hence the need for digitally available reports.

In Switzerland, regulators perform independent calculations as part of their review of the safety case. Data sets (input parameters, dimensions of a repository, etc.) should be provided by the implementer to facilitate the regulator's review process. The Swiss implementer (NAGRA) provides digital means to present data, which is helpful for the regulator when reviewing the quality of big datasets such as the results from 3-D-seismic measurements.

Regulators may adopt digital approaches themselves in order to review safety cases. For example, the Swiss regulator stores documents in an archive so that they are accessible and searchable. It also has future plans to ensure that digital twins developed during the construction and operation of the repository should be accessible to the regulator to facilitate the monitoring of ongoing operations and work.

[20] Swiss regulatory requirements are set out in Guideline ENSI0GO3 (ENSI 2020)

4. Panel discussion

The panel discussion session included speakers from the previous sessions exchanging diverse viewpoints on key prepared questions. The following sections summarise the personal opinions and insights shared by the panellists during the panel discussion. These views do not represent a collective opinion or consensus of the workshop participants or the EGSSC.

4.1. Q1 How can digital technologies and systems benefit knowledge and information management and help demonstrate a Golden Thread between safety claims and evidence?

4.1.1. Knowledge management in the safety case

When considering knowledge management, there is a need to be mindful of the distinction between data, information and knowledge. Data is processed to get information, and information is processed to get knowledge. Systems can be used to help this process, by using existing tools to analyse data and then to capture the knowledge and insights within (other) digital tools. At the start of the process there is a need to think about the question "How do we not lose that knowledge?", and then think about how to map that knowledge back to the data, and how an information system might be used to do so; for example, by representing that mapping as a graph. The chains of reasoning can then be clearly traced to show how we get the insights we do. There is considerable power in taking these insights and not just passively ingesting the information, but actively testing it by tracking knowledge back to the processes and data from which they were derived. This enables preservation, traceability and testing.

An example from Nuclear Waste Services (NWS), United Kingdom, shows how useful this capability could be. Regulatory feedback on the 2016 generic Disposal System Safety Case (DSSC) showed that regulators struggled to see how some claims made in the safety case where underpinned by the knowledge base, with an example given for the generation of gas. When this was looked into, references were made only to entire Status Reports (which describe the state of knowledge of a particular area, like gas), which made it difficult to trace exactly where specific statements in the DSSC had come from. Another example included the citation of a NAGRA report, in which different assumptions were used than in the DSSC. It was unclear to what extent this cited evidence applied to the NWS case.

A digital system could link to specific areas, topics, sentences, etc., from different sources, which allows for better traceability and for humans to jump to areas of relevance efficiently. It would mean a safety case would be quicker to check and enable a greater degree of consistency. Another, more recent, example given by SKB was about the references made in the FEP report, where it was found that cited information was missing in the source document, although it was assumed to be there. The lesson is that it is important to have strict enforcement on this point, otherwise a consistent traceability system cannot help.

4.1.2. The structure of data

It is important to distinguish between the structure and the content of data. Most data are not structured, and safety cases are mostly long documents not structured in a machine-readable manner, which projects like NWS' VISI have attempted to add structure to. Modern technology allows the automated extraction of structure from unstructured data. For example, a controlled vocabulary can be added as structured data and knowledge graphs of linked data can be developed with semantic metadata. There is indeed a lot of technology available to help put structured data into safety case documents.

Some caution is warranted in the above discussion, as the data referred to is assumed to be something existing presently. This is not necessarily the case in long repository programmes, as some data does not yet exist and is merely planned. How does a safety case manager plan to use these techniques in advance, for data they will produce? To address this question, they need to think about where the controlled terms come from, which would be tagged as part of a structured data approach. The method would be to identify a vocabulary initially, for instance, based on glossaries they refer to, and then consistently tag data with the controlled vocabulary when writing it. This needs consistency, as if users tag differently, problems will arise as some documents will be excluded from connections and search results when they should not be. Training would be required to address this; producers of safety case content are not archivists by background, and the process of tagging would require them to be trained and experienced. This would be in addition to work to define ontologies on which controlled vocabularies for tagging are built. These activities would require safety case specialists to work with knowledge management and archiving experts. They would also require interactions with many stakeholders to understand their needs, and use cases would need to be put together for digital safety case systems. This work has not yet been done in many organisations.

4.1.3. The Golden Thread

In the context of a safety case, which is a large document, there is the chance to engineer quality into it to reveal the logical line of claims, arguments and evidence, which was in the workshop was named the safety case "Golden Thread." At the beginning of the safety case production process, the "Golden Thread" needs to be clear. If a purely bottom-up approach to the "Golden Thread" is adopted, then it may only end up as a golden roof, where everything only *appears* to be connected but is not. This is when you lose the golden thread and may attempt to repair it with a superficial "golden roof".

Technology could be used to address this by using process engineering, modelling the human effort as supported by technology. The concept of "structural capital" was brought up and its importance stressed. Structural capital consists of the supportive, non-physical infrastructure, processes and databases of an organisation that enable the people within that organisation to fulfil their roles. Specifically, since the full safety case may be too big for any single individual to be fully understood, the implementing organisation needs the structural capital to keep an overview of the safety case, to be able to maintain it. Individuals within the implementing organisation need to actively engage with the safety case, not just store it.

FEP databases are an example of the databases implementing organisations are using for safety case production. Ultimately, they are structured but require a lot of preparatory laboratory work and modelling by specialists before they can be compiled. During their development, efforts are needed by both safety case generalists and specialists in technical fields, but there is difficulty in bringing these roles together as the specialist may not see the general overview, which should be provided by the safety case expert. Managing and maintaining the adoption of ontologies and controlled vocabularies will place an extra burden on the safety case expert.

An additional aspect that needs to be understood is the process that goes before the safety case production. A safety assessment does *not* make the repository safe, it is about assessing it as safe; both top down and bottom-up approaches are needed to achieve this.

4.1.4. Business case for a digital safety case

The safety case generalist picture for digital safety cases is to use the technology to make the overview of the safety case more tangible. There may be difficulty in convincing the scientific specialists of the need for such a tool, as the benefits of the digital tool are for the safety case authors and reviewers. This may be addressed by adopting digital tools incrementally, using business cases to make the need for such systems, and then developing them. The adoption can be front loaded, and lead to opposition, but this points to the need to clearly demonstrate the benefits in the business case for such a system so the organisation can decide the right balance for adoption. Contrasting approaches can be seen in the United Kingdom, where the Low-Level Waste Repository (LLWR Ltd.) wrote their safety case in a word processor and then had it transformed into a webpage, while NWS can generate the safety case in webpages automatically using VISI content, but in order to be able to do this, it requires more upfront effort. There is a need to understand the benefits and total costs of these approaches such that the business can adopt intelligently.

4.1.5. Further comments and issues

- It is unlikely that there will be a single digital safety case system, although there may be one platform with linked data and annotations. And there is a risk that different systems may not be compatible with each other. To ensure they are, an exchange format is needed.
- Digitisation risks removing intuition regarding how a safety case may be impacted by certain issues. In a traditional setting, experienced professionals can often sense these impacts, but in a digital environment, this intuition may be diminished. Structured digitalisation risks this flexibility. The rigid frameworks that can accompany digital systems can hinder the ability to quickly and intuitively understand the impacts of certain issues.
- When developing digital safety case systems considering how the needs of various stakeholders would be supported by such a system is necessary. Generally, users just want to see relevant information or the public topics that interest them. This may end up with more content produced for topics that are commented on, which needs to be considered. Focus needs to be placed on for whom the system is being produced; the safety case expert, the technical specialist or the public.
- Digital safety case systems need to make clear that some claims made in a safety case are judgements based on evidence, and that evidence may exist that is for as well as against the claim made, and the system needs to be transparent in showing this.

- The "tagging" of safety case content with a controlled vocabulary that is derived from ontologies is frozen at the point in which this tagging occurs. However, the language used in the safety case is likely to expand as a repository programme progresses. As a consequence, it might become difficult to retrieve desired content which is based on the "old tagging" from earlier programme phases.
- SKOSs (Simple Knowledge Organisation System) may provide a way of communicating knowledge and enabling a common knowledge representation as well as the dissemination of knowledge to the desired level. SKOSs might also allow for automated tagging.

4.2. Q2 How can digital workflows increase auditability and efficiency in the safety case production process?

The discussion started by considering the case of SKB, where there are already good tools for the use of FEPs combined with tools like assessment model flowcharts (AMF), but there is room for improvement, particularly given that most tools are in form of uncoupled databases. NWS was felt to be in a similar situation, it has good tools, but those tools are not connected.

It is important to distinguish between data models, which describe how things relate to one another and which are static, and organisational processes, which are time dependent. Data models and process workflows are therefore separate. It is possible today to buy workflow management systems off the shelf, with standards (e.g. BPMM) that enable such systems to follow written processes. Benefits of doing this include, for example, a user being able to quickly see where something is in a process, with an audit trail; such a workflow management system may also make links between items like FEPs and quality assurance comments. However, it must be ensured that the system does not get in the way of the workflow itself; i.e. that the system does not impose how the process works. It is important that the system is not too prescriptive but allows enough flexibility for the user to carry out their work.

How do these digital workflows deal with a "proponent – opponent issue"? For example, if the implementer wishes to use bentonite in the repository system and it is assumed to be homogenous (to a certain extent), if a laboratory result contradicts this, how does the finding then get used? In this case a specialist has found a process that they will state is important, but a safety case expert will have to judge whether it is important for safety. If it is neglected, there is a need to ensure this is considered and/or that the finding has been neglected for a good reason. Many of the issues described here are inherently part of the safety case process, digital or otherwise, and a digital system may not aid this, but it will potentially provide other benefits.

One existing ontology, the Provenance ontology (PROV-O in W3C standards) (World Wide Web Consortium, n.d.[27]), already models workflow-like activities. This could be used, or extended, to cover the issue above, i.e. inputs from a laboratory result and a decision on its inclusion in the safety case. There is a question as to how far this should be taken for safety cases. Potentially such methods could be used to trace data from its source in the knowledge base to data that goes into assessment models, through the chain of data elicitations and detailed models, thereby bringing structure to this process.

There is the potential to increase efficiency using workflows. For example, in writing content for a digital tool like the existing ViSI system, there is a need to add in references to underpin a claim made. There is the possibility that as part of a workflow, a tool could be added to support finding similar documents within (or outside of) the document

repository. Natural language processing methods can give some automation to suggest which documents are related to the claim that is made.

There are two main ways of increasing efficiency in producing an information product. One of these is how fast humans can feed information into a system to process or reuse it. This requires a good interface which, while hard to do well, is possible. The second is the reuse of existing content, which could be a key way to improve efficiency. However, more sharing of data (between organisations) could be needed to facilitate this. Sharing of existing safety case content may be of limited use for this purpose, since published safety cases tend to be in the form of unstructured PDFs.

Workflows may contain different types of processes within them. Some stages of the workflow might use automated tools, and other stages may be very manual. There could be manual stages of tagging content with an NLP-informed automated check on the relationships later. Automated and manual techniques could be combined to improve the process.

There is also a trade-off between the efficiency of producing the safety case and the efficiency of using the safety case when it exists. Digital systems are largely developed for end users. Key challenges for defending a safety case are consistency and traceability, but there are others. What does the safety case need from digital systems? There is a need to gather requirements from end users, so technology professionals can understand what the digital systems need to do.

4.3. Q3 How can digital technologies and systems help to improve understanding, presentation and communication of the safety case with both technical and non-technical interested parties?

Discussion in this session focussed initially on the use of visualisation systems, in particular the potential to use virtual or augmented reality systems, to communicate aspects related to repository safety cases. Communicating or presenting using virtual reality systems, or three-dimensional representations of processes, has the great advantage that they can be optimised for and tailored to specific audiences. Visualisation particularly facilitates discussions. It does not reduce the complexity of decision making but does reduce the risk of misunderstanding relevant safety-related processes.

Due to radioactive waste being a controversial topic, it was discussed that if communication is seen to be "too smooth", it can have an opposite effect to what is intended, i.e. it could decrease trust from stakeholders. To avoid this, the focus should be on content and not on the technology or presentation. VR and AR systems and models need to be prepared to properly address this. Augmented reality systems, for example, could be used to show lots of different elements related to safety cases, but WMOs need to think carefully about how to use the technology in any specific case. On the other hand, as these tools continue to grow in capability, there might be a growing expectation from stakeholders that such technologies could be used for communication purposes. A response that "Nuclear doesn't do that" is not likely to be helpful and in this case the implementer risks being seen as backwards thinking or stuck in the past. Therefore, when adopting VR or 3D systems for communications purposes, there is a balance needs to be struck.

The Finnish disposal programme has achieved a high acceptance rate, which is seen as the result of 30-40 years of trustworthy communication. Use of technology can receive backlash that it is "too glossy". The Finnish example shows that the important point in implementers dealing with the public is to show that they are dedicated to safety, that the organisation is trustworthy and that implementation plans are transparently communicated. Digital tools are not the key to this; rather, it is the attitude of the implementer and building confidence in their competence that is key. Tools may be helpful but are not the crucial factor in good stakeholder communication.

Virtual or augmented reality technologies could be used for communication within a technical context as well. For example, in a modelling context, AR could be used to display fracture systems or visualise chemical distributions, to show what models might mean in reality. Such systems would need relevant data, but tools should not be too prescriptive or restrictive on the modelling process. There is also a risk of ending up with systems that are unusable and have to be fixed, with lots of time spent to do this. Another use in a technical context could be to use AR with a waste package records system, to help visualise records for packages.

Technologies like Microsoft HoloLens may be mainstream soon and repository implementers are likely to use it. VR-type technology is really a discussion platform; it needs to be used with explanations to achieve correct outcomes- it is not just about demonstrating it. This might be useful later in repository programmes to use with experts and stakeholders.

These technologies are often used as part of training exercises in other industries. For example, Lufthansa used a VR system to train 5 000 flight attendants. VR could be used for training operators of a facility and may be helpful as part of the operational safety case.

The safety case includes the engineering and operation of the repository as well. Models that were built during the development processes could be compared to what is built and occurs when the repository is operating.

5. Group discussion

5.1. Opportunities and challenges of digitisation approaches

5.1.1. Q1: How can digital technologies and systems benefit knowledge management and help demonstrate a 'Golden Thread' between safety claims and evidence?

Traceability

All participants underlined the strong, fundamental need for traceability of data and information throughout the disposal programme and the safety case development process. Ensuring the traceability of data and information entails:

- The clear identification of data and its source (report, study, type of experiment, etc.).
- The preservation of the context of the data that is used in the safety case, by use of appropriate metadata (e.g. What were the limits of the experiments that are the source of the data? What was the scope of the assessment where the data was used?).

How to ensure the maintenance of the data and metadata within an organisation in the long-term (i.e. over decades) remains an open question, and participants expressed the importance of putting continued effort into this by setting up appropriate support structures and processes within the organisation.

Use of a claims, arguments & evidence approach

It was discussed that the methodology of Claims, Arguments & Evidence (CAE) provides a good high-level structure for the safety case, but it should not be the only way of developing a digital safety case. The safety case should also contain the "story" of the how the system works and how we understand it. Participants found it important that the safety case contains a detailed system description, based on solid phenomenological understanding, and that this knowledge is maintained over the long-term. In this way, the arguments and evidence that directly respond to the safety claims made via the CAE structure can be augmented with more general information.

These two, i.e. CAE and "story telling" (and other) approaches are seen as complementary. As such, digitalisation efforts could allow for different ways of structuring and presenting the safety case, and by taking them together, the robustness of the safety arguments can be enhanced.

It was discussed that a digital safety case based on CAE would provide an opportunity for developing multiple argument tree structures that show and highlight specific perspectives on the safety of the system, depending on end-user needs.

Proceeding with a digital CAE would allow for interactive querying (also by external stakeholders) and an easier navigation of the relationships between the underlying data and knowledge base, and the safety arguments made in the tree.

A digital safety case based on CAE provides an opportunity for interactively drilling down the argument tree, depending on the level of interest of the safety case end-user. However, there remains an open question of what should be the required level of detail for such a CAE tree; i.e. should it go all the way down to the individual data point level, or is there an intermediate end-point, where working out further detail on subclaims does not outweigh the effort needed to do so?

"Golden Thread" vs. Multiple lines of reasoning

Participants highlighted the importance of having multiple lines of reasoning running through the safety case. As such, participants found that one should not talk about a single "Golden Thread" for the safety case narrative, as if the safety of the disposal system would only be based on a single argument. One could rather describe the collection of arguments and the multiple lines of reasoning as a sort of "Golden Net" or "Golden Jacket". This illustrates that the number and width of the arguments themselves enhance confidence in the system's safety.

Need for international collaboration

Participants expressed a strong need for international collaboration in order to progress with digitalisation of the safety case. Strong collaboration would:

- allow the sharing of resources between organisations to solve problems faster and more efficiently;
- enhance the probability of long-term preservation and maintenance of developed methods and tools, or at least lower the risk of loss of knowledge of these (see also the examples of lost safety case tools that are given in Section 6.2);
- enable less mature programmes to use and build upon work that was done previously by more mature programmes.

Starting with digitalisation efforts

Finally, how best to start the development of a digital safety case was discussed. The following recommendations arose from the discussion:

- At first, keep digitalisation projects small and simple, with well-defined targets for success. Later, use the acquired return-on-experience to proceed with future projects. Keep flexibility when expanding incrementally, e.g. take inspiration from agile approaches in the field of software development.
- Start development based on well-known, existing systems and approaches, e.g. Wiki structures, for which much experience and support exist.
- Decouple metadata extraction and storage for use in the digital safety case from the used information-data-&-knowledge repositories, that are in many cases pre-existing within an organisation. This would allow for building upon existing infrastructure, e.g. existing but not-yet-integrated databases or assessment tools.

5.1.2. Q2: How can digital workflows increase auditability and efficiency in the safety case production process?

Discussions on this question began by considering what auditability and efficiency mean in terms of a digital safety case. For auditability, issues around demonstrating compliance with standards and processes, as well as tracing the audit trail surrounding updates to the safety case and relevant decision making, were raised. Furthermore, traceability of the technical content was seen as a potential big advantage for auditability in a digital safety case. This raised the question of for whom a digital safety case system is for, e.g. not only an implementer but also a regulator. It was discussed that many of the use cases for auditability are of most benefit to regulators.

For efficiency in a digital safety case system, participants tended to mention the reuse of content and information, and avoiding duplication as key points. Also mentioned were the potential to distribute tasks among experts and resolving conflicting assumptions within the safety case. Participants also raised a warning that digital safety case tools should not be too painful to use.

The session then discussed five aspects identified as relevant to auditability and efficiency, related to safety case production in digital systems. These were: (1) process modelling, (2) change and configuration control, (3) electronic signing, (4) workflow engines and status tracking, and (5) access to archives.

Process modelling

The discussion on process modelling was broad ranging but tended to focus on existing safety case flowcharts and workflows. These existing flowcharts are static snapshots of an organisation's mode of working at a particular time, but these processes change with time. They currently exist largely graphically but could be drawn in tools such as MS Visio or other existing diagramming tools. The explanation of these flowcharts is captured in accompanying reports but it could be encoded logically using a standard such a BPMN (Business Process Model and Notation). It was expressed that it would be desirable to capture issues like the competence needed for certain decisions as well as the iterative nature of these processes, possibly using a version-controlled system which would provide traceability.

A potentially desirable use for process models would be the ability to trace the use of data and knowledge within the safety case. This may then lead to the ability to track where updates are needed and propagate updated data through the system.

A real challenge to modelling safety case processes is that these processes rely heavily on expertise and human judgement, particularly with regard to the safety of repositories over long time frames. A strong view was articulated that these cannot be expressed in an automated process. Furthermore, automation would only provide a limited advantage due to safety cases being updated only infrequently. Related to such views was the belief that the adoption of process models may encounter cultural resistance as such tools may not be accepted by users. This risk may be mitigated by providing appropriate training within organisations.

Change and configuration control

Most discussion around change and configuration control focussed more specifically on version control and version control systems that could be used within a digital safety case system. Version control in existing safety case production is a manual process, if any version management exists at all.

Version and configuration control in a digital safety case should be applied to all files related to the safety case, with practices such as data freezes formally incorporated. Such systems should be able to record aspects like model runs, so that calculations could be reproduced in the future, if needed.

Wider use cases were expressed for systems which could help manage the impact of changes to aspects of the safety case content. The example given focussed on updated knowledge about specific topics and how these would propagate through the safety case. The ideal case would be that when updates to specific areas of the knowledge base are made, a version controlled digital safety case system would identify the relevant areas of the safety case that require review or updating as a consequence, with safety assessors appropriately alerted to such changes. This could extend down to specific parameters in assessment models, which may be used across multiple models and assessment calculations.

Many of the challenges of version and configuration control within a digital safety case system relate to cultural or organisational aspects. A key issue, for example, is the need for such systems to work across an implementing organisation and its commercial supply chain. Commercial suppliers would likely use different systems which are tailored specifically to their individual specialised services. Version and configuration control may face difficulties when updating detailed aspects provided by suppliers in the same way as updates undertaken entirely within an implementing organisation. Therefore, some aspects of the version controlled digital safety case may end up outside the traceable versioned record and limit the ability to propagate changes throughout the safety case appropriately. Other challenges may be raised from user acceptance of such systems, as well as security constraints placed upon them by organisations, which may limit their functionality.

Electronic signing

Proper signing procedures are important for traceability and the enforcement of quality assurance processes. It is meant to ensure “you get what you ordered”. The existing capabilities for electronic signatures go from a basic implementation, in which images of paper signatures are used, to fully electronic signatures which may be associated with cryptographic systems. Electronic signatures can be implemented in existing digital management systems such as SharePoint.

The desired capabilities for electronic signing in a digital safety case system were thought to include connections with a digital tracking system (for example in a workflow system) and connections to competence management systems (to show a SQEP[21] individual has signed).

It was thought that it may be difficult to adapt such systems to organisation changes, including those in supply chain organisations, and that the longevity and durability of electronic signature protocols may be limited. Limited adoption inside of organisations could be a big hurdle could also include some legal barriers.

[21] SQEP: Suitably Qualified and Experienced

Workflow engines and status tracking

Workflows with existing safety case processes are largely manual, with safety cases largely being sets of documents and reports which are updated manually. Formal workflow management systems are, however, commercially available but there has been no reported adaption of such systems for repository safety cases. There was a clear use case, from a regulatory perspective, for the adoption of a workflow within a digital safety case system. This would enable the tracking of adherence to appropriate quality assurance processes and make review cycles traceable and potentially searchable. It would also give a clear audit trail on decisions made in safety case production processes. It was noted that there is a need to ensure that staff are able to make basic challenges and "ask stupid questions" even in the tracked process, in order to avoid either discouraging more information discussion and review of aspects of the safety case, or pushing aspects of this process outside what is recorded within such systems.

Accessing archives

Archiving is largely about maintaining documents over time, which is a key problem for long-term repository programmes. Multiple formats exist for archiving including digital, physical and logical. These may exist in the form of traditional libraries or data servers storing items. For digital material, a difference can be seen between "bitstream conservation" and "content preservation", where the former seeks to conserve the whole gamut of digital information and the latter merely a record of the content that the original digital file contained. For digital archiving it is desirable that such systems are searchable, but that they also contain a degree of user access control, to provide appropriate access and ensure confidential information is only accessed by authorised individuals. It was felt that archiving systems should be relatively simple, although training may also be required for users.

Multiple challenges exist in archiving of repository programme material, particularly in the safety case. This is due both to the longevity of repository programmes, as well as the use of niche digital file formats that are the result of very specialised software that is used. This presents clear costs to the programme which would need to keep maintaining old file formats and software, in order to access archived files and information. The need for such longer-term data preservation may be challenged within programmes, particularly when there are large short-term pressures to produce the next stage of a licence application.

It should be noted that the workshop discussion primarily addressed internal archiving approaches within organisations. External archiving, such as archives beyond the responsibility of the organisation, introduces its own set of unique challenges, including issues related to a mechanism of transferring responsibility of archiving and compatibility of standards (NEA, 2019[28]).

5.1.3. Q3: How can digital technologies and systems help to improve understanding, presentation and communication of the safety case with both technical and non-technical stakeholders?

To support the discussion in this session, an initial brainstorming exercise was conducted in which participants could contribute topics relevant to Question 3. In a second step, participants selected topics for further discussion. Discussion groups where then divided into two sub-groups, based on the specific interests of each participant, for deeper discussion of one of the selected topics. Discussion of each of these topics is presented below.

Topic 1: How can we reach out to various stakeholders to talk about the safety case with digital tools?

We have to consider a diverse public with different levels of expertise, to which the safety case needs to be explained. This also comprises colleagues and specialists in the waste management organisations and regulatory bodies themselves. For this reason, there is a wish not to complicate the respective reports, but instead to use tools for visualisation and traceability. Uncertain information should also be included, but what is known and what is not known should be stated honestly. Since the safety case is a complex issue, easy access to different levels of complexity needs to be established. However, arguments should be structured in the same way in each level as learning paths for individual expert levels. In addition, text with hyperlinks, adjustable structured information, and reader-specific vocabulary and ontologies should be used. As a link between the public frequently asked questions and the technical domain, tools such as chat bots, system-generated answers and apps for facility walkthroughs were suggested. Furthermore, an encyclopaedia, which is linked to a discussion forum, was proposed.

Topic 2: Interaction of non-technicians with digital solutions

Is a machine answer better than an answer from an expert? What does "better" mean in this context? As we know, a machine answer can be more "correct" than a human answer, but it could also be out of context or wrong. Therefore, humans are still needed to check machine answers and, consequently, to develop trust in the machine answers. A query tool for an expert interacting with the public was suggested. This could introduce problems with resources; however, it also has the potential to provide direct feedback to the safety case.

Are digital systems needed so that non-technical audiences can understand the scientific method? An advantage would be that a digital twin could allow opponents to make their own experiences. Maybe this would speed up the scientific process, to develop arguments and respective counterarguments. This could force opponents to be more honest and to provide more evidence for their claims.

Topic 3: 3D modelling and visualisation of the environment or a future repository to support decision making and to communicate decisions

The use of 3D models in 2D documents was suggested as a way to present data and modelling results in a 1:1 scale. This would facilitate the discussion among different specialists, would allow mistakes to be found at an early stage, and could reveal contradicting requirements. With this tool, the steps of decision making could be followed.

Topic 4: Interactive graphics and media

It may be helpful for data to be plotted based on the specific requests by the reader. The presented data could also be adjusted to the user, depending on the purpose of communication and who is being communicated with. It could enable the user to see the big picture instead of just one aspect. Different layers with different levels of information could be imagined, depending on if the discussion is with experts and regulators or the public. Particularly in the case of 3D data presentation, with orders of magnitude of uncertainty, e.g. probabilistic results, interactive graphics could offer the possibility of showing the median, the minimum and the maximum of the data.

How should the digital data, including unprocessed data, be preserved for future use? It is important to document how the digital system developed with time, and to regularly update the safety case software to avoid the loss of data. Data consistency should be guaranteed, meaning that broken links should be avoided. Otherwise, this could lead to a loss of trust.

5.2. Implementation of the digitisation approaches

5.2.1. Q1: How can a business case support digitisation of the safety case and what are the arguments and components of such a case?

Participants brainstormed on topics related to business case development (strategic, economic, financial case) and strategic approaches (in-house development versus customising commercial software, role of supply chain, delivery partners, etc.) in support of the digitalisation of the safety case. Common points of discussion emerged in the conversations held by the different groups.

The effort and motivation underpinning the development of the business case itself was discussed. Although the meaning of 'business case' in commercial activities may not be applicable in the context of not-for-profit activities like this one, the relevance of developing a *business case* for the digitalisation of the safety case was recognised. The business case should aim to convince stakeholders (managers, administrators, financial officers, and technical and scientific experts) of the quantitative and qualitative value that the digital transition would bring to the organisation and help overcome the reluctance to reconsider long-established, business-as-usual practices in the development of the safety case. Building a business case for digitalisation of the safety case entails the use of resources, but it would offer organisations clarity on the risks, costs and expected benefits that the digitalisation process would bring and, thus, inform the decision-making processes, potentially leading to the investment required to launch the developments.

An estimation of the associated costs was pinpointed as a critical component of the business case. Providing such estimates could prove challenging as these are undefined and variable. Other highly regulated industries (aerospace, pharmaceuticals, oil and gas, etc.) pursuing digitalisation activities could offer lessons learnt and useful insights to apply to the nuclear industry.

In particular, information is required to form an idea of the costs associated to the adoption of digital technologies in the production, maintenance and provision of digital safety cases, in terms of data protection requirements, security updates, software development and technical support. The question of developing in-house tools vs. resorting to commercially available software was debated. On the one hand, developing in-house solutions for the digitalisation of the safety case could prove to be a substantially capital-intensive activity. On the other hand, although resorting to commercial software could seem to be *a priori* a cost-effective alternative, available general-purpose tools may not address the specific needs of the digitalisation of the safety case, which, in turn, could force organisations to

invest in further developments in order to tailor, complement and/or couple these available tools and make them suitable for their intended use. In-house software prototyping for further development by external companies was indicated as an option. The need to survey stakeholders on the tools/formats/standards to develop/adopt depending on their level of involvement (regulatory, implementation, technical support, public communication, ministerial review, etc.) adds a layer of complexity to the definition of the technical requirements of such tools.

Despite cost estimation and technical challenges, the groups clearly identified quantitative and qualitative benefits that the business case needs to factor in, including:

- **Long-term economic return**: The digital transition could help streamline safety case related activities by turning them into optimised, repeatable processes. These efficiencies, along with automation where applicable, in safety case production could lead a long-term economic return, as well as to shorter licensing processes.
- **Multi-stakeholder collaboration**: Digital safety cases could enhance collaboration and communication between internal and external stakeholders through instant, remote accessibility to all information and data used in their production.
- **Trust-building**: Digital safety cases could allow different stakeholders to have access to transparent, traceable and quality-assured information, which could strengthen accountability and build trust among the stakeholders involved. Specifically, the digital transition could represent an opportunity in the public communication domain, by enabling accessibility to civil society.

The discussion groups recognised that many organisations and stakeholders are going through a digitalisation of their products and processes. The digitalisation of the safety case would align with this trend and help organisations make the most of new opportunities and findings that it could bring to the industry, and avoid the potential hidden costs of not pursuing digitalisation in this way.

5.2.2. Q2: How do organisations migrate to a digital safety case?

The discussion on this question focussed on pre-proposed topics prepared by the workshop organisation committee. Discussion groups were divided into two sub-groups, based on the specific interests of each participant, for deeper discussion of one of the selected topics.

Topic 1: Roadmaps for delivery and prioritisation of activities (e.g. staged roll out, staff training on approaches and systems, development of management system procedures)

The roadmaps for delivery of digital safety case systems should be linked to skill development of the waste management organisation. Together with repository development, the pivot to digital should be planned. Already at an early stage, it should be considered that systems have to talk to each other in the future. When a digital system/tool is created, there should be time allowed for trial and error. Furthermore, sufficient time should be left for stakeholder/user involvement. If there is a choice, tools should be prioritised which are reuseable in the future by most people.

Digitalisation should be led by the safety case producer and supported by digital experts at each stage. Co-operation between the respective waste management organisations is a key factor. For instance, Posiva Oy, the Finnish waste management organisation, was not participating in this NEA workshop, but could have some contributions to the digital safety case.

Topic 2: New skills and education needed

There will be a need for a new generation of experts with certain skill sets, particularly experience in advanced data science. The term "super-user" was used to described these experts, though it was also discussed that this specific expertise could be outsourced. The relevant skills comprise the ability to think in a structured way (and in a "digital way"), to develop the digital safety case systematically within the waste management organisation. It is important to hand over knowledge when people leave the company. In Korea, this is implemented by a Wikipedia-type prototype platform.

Topic 3: Managing digital information over the very long time frames associated with a disposal facility (system lifetime and data migration)

What has to be done now to manage lifetime storage of safety case information? Who pays for it and what are the actual costs for storage of information? What time frame has to be covered? It should be considered that there is still some information not in digital formats. That information would need to be converted. Criteria have to be defined as to which information will be accepted for storage. Different standards would be needed for different aspects. For instance, different formats should be provided to access information. The different levels of complexity, particularly regarding interaction of software and hardware, need to be considered. One key aspect is that the stored information should be regularly tested, if it is still accessible.

Topic 4: Dealing with legacy information (e.g. determining priority of what to digitise and the value of image scanning versus production of full machine-readable formats)

The content and presentation of legacy information can take a number of forms, depending on how it was produced or stored. Broadly, legacy information can be divided into three classes:

1. Information which still exists and is known but is not machine-readable.
2. Information which has been lost entirely.
3. Information which is available but not traceable.

For information class (1), there already exist methods and tools such as optical character recognition (OCR) and NLP, which convert the information into a machine-readable format. In the case of information class (2), if possible, attempts should be made to reconstruct the information, while capturing and storing it using current best practices. For information class (3), a hybrid approach (both analog and digital) would be necessary. Inverse modelling can be tried in the case of experimental data. Expert knowledge should be used to decide if the information is trustworthy.

The question is how much time and resources should be invested to deal with legacy information?

5.2.3. Q3: What does the vision of an "ideal" digital safety case look like and what future work would be needed to achieve this?

This discussion session split this question into two aspects. The first aspect covered what an ideal digital safety case would look like and the second aspect focussed on the work needed to achieve one.

Ideal digital safety case

The ideal digital safety case would need to take account of the process for safety case production, namely that a wealth of knowledge produced and curated by domain experts is then organised and used by safety case specialists, who are generalists not domain experts, to produce the safety case. Digital tools may make aspects of this safety case production process easier.

The documentation of the safety case is an area where this ideal digital safety case could be impactful. Such an ideal system could make it significantly easier to bring strands of specialised knowledge together. To enable this, an ideal digital system could enable very fine-grained referencing of documentation, meaning it would be very easy to trace evidence through the safety case documents, directing users directly to sections or paragraphs rather than to potentially large reports where the generalist reader might be unable to trace the thread further. Such a system could also enable this thread tracing in the opposite direction, allowing domain experts to see how their topics map to the safety case. This could also enable searching or querying of how topic-specific information is used in the safety case, or if it is not used, and why this choice has been made. This would then enable both the safety case generalist and the domain expert to understand how and to what extent domain-specific knowledge is contributing to the safety case arguments. Safety case documentation could also be subject to audience specific views; for example, levels of technical complexity displayed to users could be tailored depending on viewer preferences. This could also include more general filters of information displayed to the user.

The methodology and process for compiling and reviewing a safety case could also benefit from this ideal digital safety case. The ideal from this perspective is that of a "live" safety case. In this paradigm, changes or updates in knowledge or understanding are easily fed into the digital safety case system, which would then be used to identify areas of the safety case that require review or update. Any updates made to the safety case would be trackable and traceable, showing the QA processes that has occurred for the updates to be made. The "live" safety case would be used to produce 'frozen' versions which could then be used for licence application submissions. There was, however, scepticism and caution associated with this ideal vision. There will likely always be human expertise involved in identifying areas of the safety case that require updating as the result of new information. Furthermore, such a 'live' model introduces a risk associated with the sum of small changes. On their own, some changes may not be significant but added together, they may be. The 'live' model in a digital safety case may not be able to detect such risks. However, this could be mitigated by introducing periodic full updates to the safety case, which are similar to the process of doing (full) Periodic Safety Reviews of Nuclear Power Plants, in addition to 'live' models.

Future work needed to achieve it

For the work needed to achieve an ideal digital safety case, it was discussed that such an ideal digital safety case is still a relatively immature concept, and that the concept itself would require more development. This was reflected in the limited replies provided by organisations to the survey on digital safety case systems for this workshop. A discussion of future work was therefore focussed on moving this area forward and not on the specifics associated with implementing a digital system.

One of the key elements was that a more mature comprehension of the relevant digital capabilities available was required for safety case specialists, but also that a greater comprehension of safety cases was needed for technology experts in order to appropriately understand what the safety case is, and what it sets out to achieve. The aim would be to

ensure that clear dialogue could be established between the safety case specialists and the technology experts, enabling a clear articulation of digital safety case needs on one hand, and the communication of technological capabilities on the other hand. It was felt that NEA forums would be a good place for such a dialogue to occur and that the current workshop was a good first step in doing so.

It was also felt that a clear articulation of best practice in terms of digital safety cases would be a good output to be agreed on within NEA forums. This would then enable organisations to build appropriate business cases for the development of appropriate systems to meet safety case needs. This would need to consider what is currently implemented within waste management organisations, as well as potential examples from outside the field. Such work may also consider developing proof of concept systems in isolated areas, to inform good practice and explain what the ideal digital safety case would be. In particular, user stories of the system could be developed, such as for the safety case author, the technical specialist, the regulator, etc. It was felt this work could be cross cutting across NEA groups, such as the EGSSC, the IGSC, etc., but that it would be useful to develop.

Further future work that was discussed surrounding common standards that digital safety case systems would use. In particular, it was felt that common ontologies, dictionaries and other linked data related standards could be developed in common, through NEA fora and that such standards would likely be needed in any ideal digital safety case system.

6. Conclusion and recommendations

The workshop on digital safety cases (DSC) highlighted the strong interest and enthusiasm among participants regarding the potential benefits DSCs offer. However, a common understanding of what constitutes a DSC is still lacking. Additionally, digitalisation should not only focus on the presentation format but also on the entire production process. To facilitate the digital transition, it is essential to systematically assess the needs and requirements of all stakeholders, including implementers, regulators, and other parties, and align these needs with the capabilities of digital technologies.

Information management must be integrated throughout all stages of safety case development. Ensuring that regulators and other stakeholders are involved in the digitalisation process from the beginning is crucial for the successful review of deliverables. Early consideration of interoperability and agreements on data preservation and exchange formats are necessary to prevent complications later in the licensing process.

To foster progress, it is recommended to initiate easy kick-start projects within national programmes to identify challenges and opportunities in DSC. Advanced digital technologies from both nuclear and non-nuclear sectors offer valuable insights and applications which could be integrated into DSC development for radioactive waste management.

Collaboration between safety case experts and digital technology specialists presents significant opportunities for innovation and knowledge transfer. Continued exchange and collaboration among these experts are essential for advancing the field.

While individual organisations are developing digital solutions tailored to their specific needs, these efforts are often isolated from a broader vision for digitalisation of safety cases. A comprehensive understanding of the current state of digital approaches is needed, along with co-ordinated development efforts to address gaps in expertise and ensure better integration across the field.

Attention must also be given to the risks of digital obsolescence, such as the discontinuation of products or data formats, which could lead to the loss of information and knowledge. Selecting tools with sustainable data storage and export options, as well as considering the longevity of tools and languages, is vital for preserving information and knowledge over time.

Knowledge management must be explicitly addressed to prevent data loss, especially from an intergenerational perspective. The lack of transparency in the tools used by Waste Management Organisations (WMOs) poses challenges for knowledge transfer, underscoring the need for improved communication and sharing of information across the industry.

With these results of the workshop, the EGSSC is collecting elements and ideas for the development of the vision for the DSC. Some examples are listed as follows:

- Ensuring traceability of data and information throughout the disposal programme and safety case development process is crucial. This includes opportunities for process modelling approaches and the integration of specialised knowledge strands during safety case production.
- Leveraging both general system descriptions based on scientific, state-of-the-art knowledge and structured safety argumentation (e.g. through Claims, Arguments & Evidence (CAE) frameworks). This would also allow for the separation of content from structure or output format, enabling content reuse and minimising duplication errors.
- The DSC should offer interactive features for querying and navigating through the safety case, making it more accessible and user-friendly.
- Fostering interoperability within an organisation, between organisations (e.g. supply chain), and on an international level, is essential. This requires common standards, dictionaries, ontologies and linked data principles, to ensure seamless sharing of information and resources.
- The DSC should facilitate the demonstration of compliance with regulatory standards and processes, along with maintaining an audit trail for updates to the safety case and relevant decision-making processes.
- The DSC should support change control and version control processes, allowing for the management of updates in the knowledge base (e.g. new R&D results) and their impact on safety case content, and the validity of safety arguments. This moves towards the concept of a ‘live’ safety case, where changes are immediately integrated, and the safety case is consistently maintained in its most up-to-date form.
- Implementing proper electronic signing procedures and ensuring comprehensive quality assurance throughout the DSC.
- Supporting formal workflows for scenario development, managing calculation cases and other essential processes within the safety case lifecycle.
- Ensuring proper internal archiving, with robust access control mechanisms to safeguard sensitive information.
- Incorporating 3D modelling, interactive graphics, and media to enhance stakeholder communication and understanding of the safety case.
- Allowing the DSC to be adaptable in terms of the level of detail presented, depending on the user's background, expertise, or preferences, ensuring it meets the needs of diverse audiences.

References

AKSW Research Group (n.d.), *RDFUnit: An RDF Unit-Testing Suite*. [1]

Arlight and K. LLC (2026), *LichtAR*. [4]

Bailey, L. and D. Billington (1998), *Overview of the FEP Analysis Approach to Model Development, NIREX SCIENCE REPORT S/98/009*. [10]

Claris International Inc. (2026), *Claris FileMaker*. [21]

DBpedia Association (n.d.), *DBpedia Databus*. [2]

Fu, X. et al. (2015), "Supporting the annotation of chronic obstructive pulmonary disease (COPD) phenotypes with text mining workflows", *Journal of Biomedical Semantics*, Vol. 6/1, https://doi.org/10.1186/s13326-015-0004-6. [6]

IAEA (2012), *The Safety Case and Safety Assessment for the Disposal of Radioactive Waste for protecting people and the environment, Specific Safety Guide No. SSG-23*, IAEA, Vienna, http://www-ns.iaea.org/standards/. [25]

Jofche, N. et al. (2023), "PharmKE: Knowledge Extraction Platform for Pharmaceutical Texts using Transfer Learning", *Computers*, Vol. 12/1, p. 17. [5]

Kelly, M. and D. Billington (1997), "Scenario Analysis and Conceptual Model Development, Using FANFARE", *MRS Online Proceedings Library*, Vol. 506/1, pp. 701-708, https://doi.org/10.1557/PROC-506-701. [14]

Kelly, T. (2008), "Are Safety Cases Working?", *Safety-Critical Systems Club Newsletter*, Vol. 17/2, pp. 31-33. [24]

Kelly, T. (1998), *Arguing Safety - A Systematic Approach to Managing Safety Cases*, University of York, www.researchgate.net/publication/2565194. [22]

NEA (2022), *IGSC Safety Case Symposium 2018: Current Understanding and Future Direction for the Geological Disposal of Radioactive Waste*, www.oecd-nea.org/jcms/pl_75947 (accessed on 31 March 2025). [8]

NEA (2022), *NEA RWMC Workshop on Developing Safety Cases for Various Radioactive Waste Disposal Facilities: Needs and Challenges of RWMC Organisations*. [29]

NEA (2019), *Preservation of Records, Knowledge and Memory (RK&M) Across Generations Final Report of the RK&M Initiative*, www.oecd-nea.org/jcms/pl_15088 (accessed on 27 July 2023). [28]

NEA (2018), *Metadata for Radioactive Waste Management*, www.oecd-nea.org/jcms/pl_15062/metadata-for-radioactive-waste-management?details=true (accessed on 26 July 2023). [12]

NEA (2012), *Methods for safety assessment of geological disposal facilities for radioactive waste : outcomes of the NEA MeSA Initiative.*, OECD Publishing. [16]

OASIS (2012), *OASIS Darwin Information Typing Architecture (DITA) TC*, www.oasis-open.org/committees/dita/charter.php (accessed on 26 March 2025). [13]

Open Energy Platform (n.d.), *Open Energy Databus*. [3]

SKB (2019), *Post-closure safety for a proposed repository concept for SFL: Main report for the safety evaluation SE-SFL*. [20]

SKB (2015), *Safety analysis for SFR Long-term safety: Main report for the safety assessment SR-PSU*. [19]

SKB (2011), *Long-term safety for the final repository for spent nuclear fuel at Forsmark: Main report of the SR-Site project*. [18]

SKB (2006), *Long-term safety for KBS-3 repositories at Forsmark and Laxemar – a first evaluation: Main Report of the SR-Can project*, www.skb.se. [11]

SKB (1999), *Deep repository for spent nuclear fuel: SR 97 - Post-closure safety*. [17]

Syncro Soft SRL (2002-2026), *Oxygen XML Editor*, www.oxygenxml.com/. [15]

Thompson, P. et al. (2016), "Text mining the history of medicine", *PLoS ONE*, Vol. 11/1, https://doi.org/10.1371/journal.pone.0144717. [7]

U.S. Nuclear Regulatory Commission (2025), *10 CFR Part 2, Subpart J: Procedures Applicable to Proceedings for the Issuance of Licenses for the Receipt of High-Level Radioactive Waste at a Geologic Repository*, www.ecfr.gov/current/title-10/chapter-I/part-2 (accessed on 27 July 2026). [9]

UK Ministry of Defence (1996), *JSP 430: Ship Safety Management System Handbook*. [23]

WENRA (2014), *Radioactive Waste Disposal Facilities Safety Reference Levels*. [26]

World Wide Web Consortium (n.d.), *PROV-O: The PROV Ontology*, 2013, https://www.w3.org/TR/prov-o/ (accessed on 26 March 2025). [27]

Appendix A. Workshop agenda

NEA IDKM/EGSSC Workshop on Digital Safety Case Methods and Development

Tuesday 25 October – Wednesday 26 October 2022
Venue: NH Hotel Berlin Mitte am Checkpoint Charlie, Leipziger Str. 106-111, 10117
Berlin, Germany

Optional Site Visit (Monday 24 October)
Site Visit to Morsleben repository Transportation and refreshments will be provided. Departure time and place: 6 am from Berlin (meeting point at the entrance of the NH Hotel Berlin Mitte am Checkpoint Charlie - Leipziger Str. 106-111, 10117). Estimated arrival time to Berlin: 5.30 pm.

DAY 1 (Tuesday 25 October)	
08:45	*REGISTRATION*
Opening of the Workshop Chair: Jonathan KINDLEIN, *Head of Post Closure Safety Analyses Group, BGE, Germany*	
09:15	Stefan STUDT, *Chairman of the Management Board, BGE, Germany* Luca PICIACCIA, *Senior Adviser, DSA, Norway / EGSSC Chair;* Ulrich NOSECK, *Head of Safety Case Site Selection Department, GRS, Germany/ IGSC Co-chair*
Session 1 – Digitisation and Digitalisation experiences Chair: Alexander CARTER, *Senior Environmental Safety Case Manager, NWS, United Kingdom*	
09:30	**Linked Data** Khalil AHMED, *Senior Solutions Architect, Epimorphics, United Kingdom*
10:00	**Knowledge Graphs** Sebastian HELLMANN, *Executive Director of DBpedia Association, Institute of Applied Informatics (InfAI), Germany*
10:30	**Virtual and Augmented Reality – Applications in Production Engineering** Philipp KLIMANT, *Chief Executive Engineer and Head of Business Unit, Chemnitz University of Technology and Fraunhofer Institute IWU*
11:00	*BREAK*
11:30	**Unlocking knowledge from Textual Data using Natural Language Processing** Riza BATISTA-NAVARRO, *Senior Lecturer, University of Manchester, United Kingdom*
12:00	*LUNCH*

DAY 1 (Tuesday 25 October)	
Session 2 – Stakeholder perspectives and national use cases Chair: Shogo NISHIKAWA, *Knowledge Management Expert, NUMO, Japan*	
13:00	**Main takeaways from the Digital Safety Case Questionnaire** Olivier DE CLERCQ, *Inspector – Expert in Waste Management and Disposal, FANC, Belgium*
13:05	**Input from outside the Safety Case community: Improving Safety Case content and communication using electronic tools** Klaus-Jürgen RÖHLIG, *Executive Director - Professor of Repository Systems, Technical University of Clausthal, Germany*
13:35	**VISI Safety Case Management System** Alexander CARTER, *Senior Environmental Safety Case Manager, NWS, United Kingdom*
14:05	**Use of databases for FEP, activities and input data in SKB's post-closure safety assessments** Niko MARSIC, *Expert Radionuclide Transport, SKB, Sweden*
14:35	**Regulatory perspectives on the use of Digital Safety Cases – requirements, expectations and independent Safety Case review** Olivier DE CLERCQ, *Inspector – Expert in Waste Management and Disposal, FANC, Belgium*
15:05	*BREAK*
15:30	**Panel Discussion on Session 1 and 2** Moderator: Jesus S. MARTINEZ, *Nuclear Energy Agency - Secretariat* Key questions: • **Q1**: How can digital technologies and systems benefit knowledge and information management and help demonstrate a 'Golden Thread' between safety claims and evidence? • **Q2**: How can digital workflows increase auditability and efficiency in the safety case production process? • **Q3**: How can digital technologies and systems help to improve understanding, presentation and communication of the safety case with both technical and non-technical interested parties?
16:45	*BREAK*
17:00	**Summary of Panel Discussion** Oliver HALL, *Post-closure Safety Specialist, NWS, United Kingdom*
17:15	*END OF DAY ONE*

DAY 1 – Workshop dinner
Workshop dinner offered by BGE Buffet dinner at 6:30 – 9:30 pm. Restaurant: Erdinger am Gendarmenmarkt Berlin, Jägerstraße 56, 10117 Berlin

DAY 2 (Wednesday 26 October)

Session 3 – Digitalisation of the Safety Case

Chair: Ann-Kathrin LEUZ, *Head of Disposal and Safety Analyses Section, ENSI, Switzerland*

Participants split into three breakout groups according to a predefined list produced by NEA. Each group has a mix of skills, experience and interests. Each breakout group will work with a facilitator and rapporteur to collect ideas and thoughts on three key questions relating to the digitalisation of a radioactive waste safety case. After the allotted time, the group will change breakout room to consider the next question, with the next facilitator and rapporteur.

Session 3a – Opportunities and Challenges

- **Q1**: **How can digital technologies and systems benefit knowledge and information management and help demonstrate a 'Golden Thread' between safety claims and evidence?**
  Facilitator: Khalil AHMED; Rapporteur: Olivier DE CLERCQ
  Possible discussion topics:
  - Linked data and semantic authoring
  - Use of methodologies like 'argumentation' / 'Claims, Argument, Evidence' to structure safety case information
  - Use of metadata

- **Q2**: **How can digital workflows increase auditability and efficiency in the safety case production process?**
  Facilitator: Riza BATISTA-NAVARRO; Rapporteur: Oliver HALL
  Possible discussion topics:
  - Process modelling
  - Change and configuration control
  - Electronic signing
  - Workflow engines and status tracking
  - Accession to archives

- **Q3: How can digital technologies and systems help to improve understanding, presentation and communication of the safety case with both technical and non-technical interested parties?**
  Facilitator: Julia DOSE; Rapporteur: Claudia JOSEPH
  Possible discussion topics:
  - Interactive graphics and media
  - Use of virtual or augmented reality
  - Review and addressing of comments, within an RWMO and with Regulators

9:00	**Welcome and Opening Remarks**		
9:05	Breakout session 3a.1		
	Group 1: Q1	Group 2: Q2	Group 3: Q3
10:00	Breakout session 3a.2		
	Group 1: Q2	Group 2: Q3	Group 3: Q1
11:00	*BREAK*		
11:15	Breakout session 3a.3		
	Group 1: Q3	Group 2: Q1	Group 3: Q2
12:15	*LUNCH*		
13:15 – 14:00	Summary of Session 3a and Discussion • Summaries by Q1 / Q2 / Q3 rapporteurs • Discussion		

Session 3b – Implementation of Digitisation

- **Q1: How can a Business Case support digitisation of the safety case and what are the arguments and components of such a case?**
  Facilitator: Axel LIEBSCHER; Rapporteur: Jesus S. MARTINEZ
  Possible discussion topics:
  - Strategic case
  - Economic case (e.g. cost-benefit analysis and success factors)
  - Financial case (e.g. estimating costs)
  - Commercial approaches (e.g. in-house development versus customising commercial software, role of supply chain, delivery partners)
- **Q2: How do organisations migrate to a digital safety case?**
  Facilitator: Julia DOSE; Rapporteur: Claudia JOSEPH
  Possible discussion topics:
  - Roadmaps for delivery and prioritisation of activities (e.g. staged role out, staff training on approaches and systems, development of management system procedures)
  - New skills and education needed
  - Managing digital information over the very long time periods associated with a disposal facility (system lifetime and data migration)
  - Dealing with legacy information (e.g. determining priority of what to digitise and the value of image scanning versus production of full machine-readable formats)
- **Q3: What does a vision for an 'ideal' digital safety case look like and what future work would be needed to achieve this?**
  Facilitator: Klaus-Jürgen RÖHLIG; Rapporteur: Oliver HALL
  Possible discussion topics:
  - Role of NEA in developing a shared vision
  - International vs. national guidance
  - Future work needed in EGSSC roadmap

14.00	Breakout session 3b.1		
	Group 1: Q1	Group 2: Q2	Group 3: Q3
14:30	Breakout session 3b.2		
	Group 1: Q2	Group 2: Q3	Group 3: Q1
15:00	Breakout session 3b.3		
	Group 1: Q3	Group 2: Q1	Group 3: Q2
15.30	Summary of Session 3b and Discussion • Summaries by Q1 / Q2 / Q3 rapporteurs • Discussion		
16:15	*BREAK*		

Session 4 – Conclusions and Next Steps
Chair: Jonathan KINDLEIN, *Head of Post Closure Safety Analyses Group, BGE, Germany*

16:30	**Summary of Workshop, takeaways and potential next steps** Oliver HALL, *Post-closure Safety Specialist, NWS, United Kingdom*
16:45	**Closing remarks** Jonathan KINDLEIN, *Head of Post Closure Safety Analyses Group, BGE, Germany*
17:00	*END OF THE WORKSHOP*

Appendix B. Questionnaire

EGSSC Workshop on Digital Safety Case Methods and Development

Questionnaire

Please complete the following set of questions using as much narrative information as you feel appropriate to inform the Workshop Programme Committee.

Section A - Background

A1. To provide context, please identify (name) the organisation for which this questionnaire is being completed, including its purpose and, if a disposal organisation, the stage of the programme (e.g. generic, down-selection of sites, construction, operation etc.).

Section B – Information Management

B1. To what degree have you formally modelled, the flow of information into and within your organisation, ending in the safety case (e.g. Assessment Model Flowcharts[22])? How is this connected formally to your business processes?
B2. To what degree do you use metadata within your organisation, especially for management of information relevant to safety case? Do staff appreciate/understand and realise the benefits of this? Can you provide examples on how the introduction of metadata has improved the efficiency of your organisations' activities?
B3. Do you have a single vocabulary/glossary which is consistently used across the safety case and its inputs? How do you ensure this consistency across multiple teams/organisations? How are changes to the vocabulary managed and cascaded if term definitions change with time?
B4. How many dedicated staff do you have employed directly or indirectly with managing information for the safety case (including IT/development staff, knowledge management staff)? Do you have plans/an understanding for how this may progress as the programme and company evolves (e.g. through a siting process)?
B5. How is your safety case structured? Do you use a logical hierarchy across or within documents? Are techniques such as CAE (Claim Argument Evidence)/Argumentation used to structure or visualise the safety case and ensure its completeness?

Section C - Systems and Tools

C1. Do you currently have any form of content management system for your safety case? If so, what capabilities does it have (e.g. interactivity, search functionality, workflow/approval, versioning/configuration management etc.) and who can access/access and update it? Can the system be used to check the consistency of safety case claims and quickly "drill-down" into more detailed information, models and data, to ensure their correctness (traceability)? How have you taken into account the needs of the regulator and other additional stakeholders (e.g. general public, local communities…)?

[22] e.g. https://skb.se/upload/publications/pdf/TR-14-11.pdf

C2. Does the safety case make use of interactive charts or infographics to better communicate the safety case to different stakeholders (Regulators, internal team members, wider society etc.)?
C3. Do you currently have any form of system to manage input data and models, scenarios, and output PA results for use by the safety case? If so, how does the system handle model versioning and verification as well as run management? How does the system interact with the safety case to ensure traceability and consistency guaranteeing that up-to-date results are referenced which support safety case claims?
C4. Do you have any system(s) to track/audit communications with stakeholders and to allow feedback to be received from such (including Regulators)? Are these automated or do you rely on manual steps such as the e-mails of documents etc. to progress?
C5. Do you have regulatory/policy needs to submit information to an external archive (e.g. a National Archive) and if so, do you understand how such a transfer would occur technically? How is the selection of safety case information needing archiving carried out?

Section D – Workshop Interests

D1. Are there any particular topics or speakers you would like to see discussed/presenting at the Workshop?
D2. Please, indicate whether your organisation would be interested in participating to the Workshop with a presentation on the topic