Nuclear Safety
2026

Preservation of Key Experimental Datasets for Nuclear Safety

OECD
BETTER POLICIES FOR BETTER LIVES

NEA
NUCLEAR ENERGY AGENCY

Nuclear Safety

Preservation of Key Experimental Datasets for Nuclear Safety

© OECD 2026
NEA No. 7744

NUCLEAR ENERGY AGENCY
ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT

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 the safe, environmentally sound 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 analyses in areas such as energy and the sustainable development of low-carbon economies.

Specific areas of NEA competence include the safety and regulation of nuclear activities, radioactive waste management, 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 publication as:

NEA (2026), *Preservation of Key Experimental Datasets for Nuclear Safety*, OECD Publishing, Paris, https://doi.org/10.82155/vaxj-0x98.

© 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.

Cover photo: Data centre (CC0).

Foreword

The safe operation of nuclear reactors depends on an extensive body of data generated through experimental research programmes and plant operating experience. The efforts and investments required to produce these data are substantial, often involving specialised equipment, large-scale experimental infrastructure, rigorously quality-assured operational databases and the expertise of highly skilled professionals.

Recognising the importance of preserving these valuable resources, the Nuclear Energy Agency (NEA) Committee on the Safety of Nuclear Installations (CSNI) initiated action through its Senior Expert Group on Safety Research/Support Facilities for Existing and Advanced Reactors 2 (SESAR/SFEAR2). This group focused on sustaining the availability of essential experimental facilities and the unique datasets they have produced in support of nuclear safety research.

As part of its conclusions, SESAR/SFEAR2 recommended the establishment of a cross-functional NEA expert group – drawing from the CSNI, the NEA Nuclear Science Committee (NSC), and the Management Board for the Development, Application and Validation of Nuclear Data and Codes (MBDAV) – to identify key experimental datasets and evaluate how they might be effectively preserved for future use. In response, the CSNI Bureau approved the formation and mandate of the Senior Expert Group on Preservation of Key Experimental Datasets (SEGPD) at the end of 2020, as recorded in the summary of the 68th CSNI meeting.

This report presents the status of data preservation efforts for experimental datasets relevant to nuclear reactor safety, along with recommendations developed by the SEGPD to strengthen and sustain these efforts. The report addresses several key technical areas, including neutronics, thermal-hydraulics, severe accidents, fuel performance, long-term operation and fire safety. While the primary focus is on large operating water-cooled reactors and reactor safety, the report also touches on topics relevant to both the front and back ends of the nuclear fuel cycle.

Although the emphasis is on selected subject areas, the recommendations are broadly applicable across all domains of nuclear safety research, encompassing both water- and non-water-cooled reactor technologies and covering existing reactors as well as advanced designs, including small modular reactors (SMRs).

This report was approved by the CSNI on 5 June 2024.

Acknowledgements

The members of the Nuclear Energy Agency (NEA) Committee on the Safety of Nuclear Installations (CSNI) acknowledge the significant contributions of all those who played a key role in the preparation of this report, most notably the members of the Senior Expert Group on Preservation of Key Experimental Datasets (SEGPD).

Special thanks are extended to the Chairs and members of the CSNI working groups who supported the development of this document, including the Working Group on Analysis and Management of Accidents (WGAMA), the Working Group on Fuel Safety (WGFS), and the metal and concrete subgroups of the Working Group on Integrity and Ageing of Components and Structures (WGIAGE). Their valuable feedback on the SEGPD questionnaire and their review of relevant technical content greatly enriched the report.

The NEA also extends its appreciation to the participants of the November 2021 SEGPD workshop, which served as an important forum for gathering insights on data preservation activities.

Further contributions and perspectives were provided through exchanges with other NEA committees, in particular the Nuclear Science Committee (NSC) and the Management Board for the Development, Application and Validation of Nuclear Data and Codes (MBDAV), whose input added depth to the findings presented here.

Finally, the NEA gratefully acknowledges Michelle Bales and James Corson, Chairs of the SEGPD, for their leadership and dedication in guiding the group's work.

Leading authors:

- James Corson, Nuclear Regulatory Commission (NRC), United States;
- Volker Hannstein, Gesellschaft für Anlagen- und Reaktorsicherheit gGmbH (GRS), Germany;
- Didier Jacquemain, NEA;
- Glyn Rossiter, National Nuclear Laboratory (NNL), United Kingdom;
- Nils Sandberg, Strålsäkerhetsmyndigheten (SSM), Sweden;
- Sylvain Suard, Autorité de Sureté Nucléaire et de Radioprotection (ASNR), France.

Contributors:

- Martina Adorni, NEA;
- Michelle Bales, NEA;
- François Barre, Autorité de Sureté Nucléaire et de Radioprotection (ASNR), France;

- Alice Dufresne, NEA;
- Evgeny Ivanov, Autorité de Sureté Nucléaire et de Radioprotection (ASNR), France;
- Jeon-Sik Kim, Korea Atomic Energy Research Institute (KAERI), Korea;
- Daniel Mathers, National Nuclear Laboratory (NNL), United Kingdom;
- Hideo Nakamura, Japan Atomic Energy Agency (JAEA), Japan;
- Alessandro Petruzzi, Nuclear and Industrial Engineering (NINE), Italy;
- Upendra Rohatgi, Brookhaven National Laboratory (BNL), United States;
- Gert Van Den Eynde, Centre d'Étude de l'Énergie Nucléaire (SCK CEN), Belgium;
- Yamane Yuichi, Japan Atomic Energy Agency (JAEA), Japan.

Table of contents

Executive summary ... 9
Key recommendations ... 10
Proposed actions ... 11

List of abbreviations and acronyms ... 13

Chapter 1. **Introduction** ... 17
1.1. Background ... 17
1.2. Objectives and scope ... 18
1.3. Procedure ... 19
1.4. Structure of the report ... 19
1.5. References ... 20

Chapter 2. **Insights from data preservation experiences** ... 21
2.1. Neutronics ... 21
2.2. Thermal-hydraulics ... 26
2.3. Severe accidents ... 30
2.4. Fuel performance in normal operation, anticipated operational occurrences and design basis accidents ... 34
2.5. Material ageing, operational experience, and long-term operation ... 40
2.6. Fire hazard research ... 46

Chapter 3. **Conclusion** ... 51
3.1. Conclusions and recommendations ... 51
3.2. Proposed actions to implement recommendations ... 53

Appendix A. **Questionnaire for gathering information on the current status of key datasets in nuclear safety disciplines** ... 55

Appendix B. **Summary of the SEGPD Information-Gathering Workshop, 18-19 November 2021** ... 57

Executive summary

The safe operation of nuclear reactors relies heavily on a vast body of data generated through experimental research programmes and plant operating experience. Producing such data often requires substantial investments – entailing specialised equipment, highly skilled personnel and large-scale experimental facilities. Recognising the strategic importance of preserving this knowledge base, the NEA Committee on the Safety of Nuclear Installations (CSNI) has taken steps to safeguard key experimental infrastructure and the data it produces.

To this end, the CSNI established the Senior Expert Group on Safety Research/ Support Facilities for Existing and Advanced Reactors (SESAR/SFEAR2), which developed recommendations for supporting existing experimental facilities to ensure their continued availability for nuclear safety research. In a 2021 report, SESAR/SFEAR2 further recommended the formation of a cross-cutting expert group – comprising members from the CSNI, the NEA Nuclear Science Committee (NSC), and the Management Board for the Development, Application and Validation of Nuclear Data and Codes (MBDAV) – to identify key experimental datasets and assess their preservation needs. In response, the CSNI established the Senior Expert Group on the Preservation of Key Experimental Datasets (SEGPD).

This report presents the status of data preservation efforts related to nuclear reactor safety, along with recommendations to strengthen these efforts. It provides examples from a range of technical areas, including neutronics, thermal-hydraulics, severe accidents, fuel performance, long-term operation and fire safety. While the focus is primarily on in-reactor safety for large water-cooled reactors, it also addresses data relevant to the front end and back end of the fuel cycle. The recommendations are broadly applicable across the nuclear safety research community, encompassing both water-cooled and non-water-cooled reactors, and covering both existing and advanced designs – including small modular reactors (SMRs).

The findings in this report are based on the SEGPD's discussions held between 2021 and 2024. The group brought together experts from the CSNI, NSC and MBDAV with extensive experience in nuclear safety research and data preservation. The SEGPD also designed a questionnaire to survey external experts about data preservation practices and needs. The results were discussed at a dedicated workshop in the third quarter of 2021, which informed SEGPD's subsequent plenary discussions. Insights and preliminary recommendations were shared at meetings of CSNI working groups (Working Group on the Analysis and Management of Accidents [WGAMA], Working Group on Fuel Safety [WGFS], and Working Group on Integrity and Ageing of Components and Structures [WGIAGE]), relevant NEA committees (CSNI, NSC, MBDAV), and the NEA Joint Nuclear Safety Research Projects Week held in January 2023.

Key recommendations

1. *Systematic integration of data into the NEA Data Bank*

Given its role as a secure, long-term repository for nuclear safety-related data, the NEA Data Bank should systematically receive and manage datasets produced through NEA joint research projects and used in data assessment activities (e.g. International Standard Problems, benchmarks, quality assurance reviews). Appropriate access controls can be applied to proprietary data.

2. *Emphasis on completeness and quality assurance*

Priority should be placed on the completeness and quality of both the data and accompanying information. Data packages should include raw and processed data, facility descriptions, test boundary conditions and measurement uncertainties. Durable and searchable formats (e.g. ASCII files, PDFs) should be used, with open-source search tools preferred to reduce long-term maintenance costs.

3. *Review and quality assessment of legacy datasets*

Older datasets may lack critical metadata. Therefore, legacy data should be assessed for completeness and quality. Feedback from Data Bank users should be collected and shared to help guide future users.

4. *Guidance on standard data package content*

Although needs vary by technical area, the SEGPD recommends developing general guidance on minimum information requirements (e.g. uncertainty estimates, facility boundary conditions). Current and future projects should be encouraged to comply with these expectations.

5. *Technical-specific guidance initiatives*

Ongoing initiatives – such as THEMPo under WGAMA and the RIA database specifications activity under WGFS – are developing detailed guidance for specific technical domains. These efforts can serve as models for other groups.

6. *The role of benchmark activities*

Benchmarking is essential for data preservation. It allows validation, helps identify gaps, and prompts recovery of missing information, especially when experimentalists and facility operators are still available. Benchmarking activities should be integrated into all current and future joint projects.

This report also provides detailed, area-specific recommendations tailored to each of the nuclear safety technical fields addressed.

Proposed actions

Related to NEA Joint Projects

- Legacy projects: NEA joint project administrators, with support from the Data Bank, should co-ordinate retrieval of missing datasets and negotiate distribution terms with data owners.
- Ongoing/future projects: Early agreement on data preservation and distribution protocols should be established at project inception.
- Post-assessment deliverables: Project teams and data owners should anticipate the delivery of comprehensive data packages following data assessment activities.

Related to the NEA Data Bank

- Awareness and outreach: The NEA should regularly present data preservation strategies at working group and joint project meetings. Awareness among NEA members of available distribution frameworks should be enhanced.
- Identification of at-risk datasets: Working groups should inform the NEA of potentially at-risk datasets and external initiatives where the Data Bank might serve as a final repository. Complementarities with efforts by other international organisations should also be explored.

Related to legacy databases

- Dataset review: The NEA should screen existing databases using clear criteria (e.g. presence of raw/elaborated data, facility descriptions, data accuracy). Determine feasibility and benefit of enhancing dataset completeness and quality assurance.
- Structured reviews: With member support, conduct expert reviews of high-interest datasets distributed by the Data Bank and document quality assurance (QA) levels.
- User feedback: Collect and synthesise user feedback systematically via the Data Bank to inform future preservation activities.

Related to future experimental projects

- High-level guidance: Establish shared expectations for data packages, focusing on elements of broad relevance across safety disciplines.
- Data policy development: Create a clear data policy covering all NEA data-driven activities, defining management practices, access rules and preservation expectations.
- Early engagement: Encourage early discussions about data package content in new projects and data assessment exercises.

Related to code benchmark activities

- Resource allocation: Project proposals should include adequate resources for organising and managing benchmarking activities.
- Data package integration: Results and insights from benchmarking should be incorporated into final data packages to ensure utility and completeness.

List of abbreviations and acronyms

AI	Artificial intelligence
AOO	Anticipated operational occurrences
AGR	Advanced gas-cooled reactor
ASNR	Autorité de Sûreté Nucléaire et de Radioprotection (France)
ATF	Accident tolerant fuel
BE	Best estimate
BWR	Boiling water reactor
CAD	Computer aided design
CEA	Commissariat à l'Energie Atomique et aux Energies Alternatives (France)
CSNI	Committee on the Safety of Nuclear Installations (NEA)
DATIF	Database for IFPE
DBA	Design basis accident
DOE	United States Department of Energy
EDF	Électricité de France
EGRFP	Expert Group on Reactor Fuel Performance (NEA/NSC)
EPRI	Electric Power Research Institute (United States)
EVMR	Ex-vessel melt retention
FAIR	Fire Risk Assessment through Innovative Research Project (NEA)
FR	Fast reactor
HRP	Halden Reactor Project (NEA)
HTR	High-temperature reactor
IAEA	International Atomic Energy Agency
ICSBEP	International Criticality Safety Benchmark Evaluation Project Handbook (NEA)
IE	Integral experiment
IFDA	International fatigue database

IFPE	International Fuel Performance Experiments database (NEA)
IRPhE	International Handbook of Evaluated Reactor Physics Benchmark Experiments (NEA)
ISP	International standard problem
IVMR	In-vessel melt retention
JEEPS	Joint ExpErimental Programme (NEA FIDES-II)
JRC	Joint Research Centre (European Commission)
LOCA	Loss-of-coolant accident
LTO	Long-term operation
LWR	Light water reactor
LWR-UAM	Benchmark for uncertainty analysis in best-estimate modelling for design, operation and safety analysis of light water reactors
MCCI	Molten corium concrete interaction
MBDAV	Management Board for the Development, Application and Validation of Nuclear Data and Codes (NEA)
MTR	Material Test Reactor
NEA	Nuclear Energy Agency
NSC	Nuclear Science Committee (NEA)
ORNL	Oak Ridge National Laboratory (United States)
PIE	Post-irradiation examination
PMO	Plant measurements and observations
PRISME	Fire Propagation in Elementary, Multi-room Scenarios Project (NEA)
PWR	Pressurised water reactor
QA	Quality assurance
RBMK	Reaktor Bolshoy Moshchnosti Kanalnyi (Russia)
RIA	Reactivity-initiated accident
RPV	Reactor pressure vessel
SEGPD	Senior Expert Group on Preservation of Key Experimental Datasets (NEA/CSNI)
SEM/TEM	Scanning electron microscopy/transmission electron microscopy
SESAR	Senior Expert Group on Safety Research (NEA/CSNI)
SFEAR2	Support Facilities for Existing and Advanced Reactors 2 (NEA)

SINBAD	Shielding Integral Benchmark Archive and Database (NEA)
SFCOMPO	Spent Fuel Isotopic Composition database (NEA)
SMR	Small modular reactor
SQL	Structured query language
THEMPO	Thermal-hydraulics Experimental Meta-data Preservation, Collection and Qualification
TRG	Technical review group
USNRC	United States Nuclear Regulatory Commission
VVER	Water-water energetic reactor (Russia)
WCR	Water-cooled reactor
WGAMA	Working Group on Analysis and Management of Accidents (NEA/CSNI)
WGFS	Working Group on Fuel Safety (NEA/CSNI)
WGIAGE	Working Group on Integrity and Ageing of Components and Structures (NEA/CSNI)
WPNCS	Working Party on Nuclear Criticality Safety (NEA/NSC)
WPRS	Working party on scientific issues and uncertainty analysis of reactor systems (NEA/NSC)

Chapter 1. **Introduction**

1.1. **Background**

The safe operation of nuclear reactors depends on an extensive and diverse knowledge base built upon both theoretical and experimental data. These data span a wide spectrum – from calculated and measured fundamental physical parameters (e.g. neutronic and material properties) to targeted experiments designed to isolate specific phenomena and integral tests aimed at understanding the interactions among multiple phenomena under realistic conditions. Such datasets serve either direct purposes, such as establishing safe operating limits for reactors, or indirect ones, such as supporting the development, calibration and validation of models and computer codes used to simulate reactor behaviour under normal and accident conditions.

In some technical areas, formal validation matrices have been developed to identify key phenomena, and associated experimental datasets are used to support the development and qualification of predictive models and codes. To ensure appropriate application of these datasets, it is essential to understand the uncertainties associated with measured parameters. Additionally, for certain complex phenomena, the physical scale of the experimental setup plays a critical role in addressing potential scaling effects and ensuring data applicability.

Generating high-quality experimental data often requires significant investment, including the use of specialised equipment, large-scale test facilities, and highly skilled personnel. Acknowledging this, the NEA Committee on the Safety of Nuclear Installations (CSNI) has undertaken a number of initiatives to preserve essential experimental infrastructure and expertise. These include systematic reviews of critical facilities needed for nuclear safety and the co-ordination of joint research projects to help maintain their long-term availability.

The most recent review, conducted by the SESAR/SFEAR2 group (NEA, 2021), emphasised that preserving access to key experimental facilities must be complemented by the preservation of the datasets themselves. Building on this conclusion, SESAR/SFEAR2 recommended the establishment of a cross-cutting NEA expert group – drawing from the CSNI, the Nuclear Science Committee (NSC), and the Management Board for the Development, Application and Validation of Nuclear Data and Codes (MBDAV) – to identify key experimental datasets and assess the actions needed to preserve them. This group would leverage the collective experience of the NSC, MBDAV and CSNI in data preservation to explore practical solutions for long-term data accessibility. Its tasks would include evaluating options for centralised data libraries, establishing criteria for dataset screening and selection, and defining the contextual information that must accompany datasets (e.g. domain of application, level of qualification, uncertainty estimates, scaling considerations). The group would also be expected to recommend approaches to manage proprietary constraints on data sharing.

The CSNI's mandate is to support and advance the scientific and technical foundations underpinning the safety of nuclear installations. Through its projects and activities, the CSNI has played a pivotal role in the development of several key experimental datasets in support of nuclear safety. Based on a proposal from the CSNI Bureau and following the SESAR/SFEAR2 recommendations, the CSNI formally established the Senior Expert Group on the Preservation of Key Experimental Datasets (SEGPD) at its 68th meeting in December 2020.

1.2. Objectives and scope

The objectives of the Senior Expert Group on the Preservation of Key Experimental Datasets (SEGPD) are outlined as follows (NEA, 2022):

- Assess the range of datasets needed to support the safety of nuclear reactors and identify those considered essential by users. Key considerations may include the cost, time and complexity involved in reproducing the data; the availability of necessary facilities and expertise; and the uniqueness of the experimental conditions.
- Identify existing databases and datasets, with particular attention to those developed under the auspices of the CSNI.
- Identify data gaps and establish priorities for addressing them based on safety relevance and preservation feasibility.

The SEGPD's current focus is on the preservation of datasets directly related to the safety of the existing fleet of large water-cooled reactors. The technical disciplines addressed in this report include neutronics, thermal-hydraulics, severe accidents, fuel performance and safety, ageing and long-term operation, and fire safety. While the primary emphasis is on in-reactor safety datasets and databases, the report also acknowledges that many of the best practices and recommendations are broadly applicable to data relevant to the safety of fuel cycle facilities as well.

Although efforts have been made in the past to capture key nuclear safety research data – including mandates to transfer data to the NEA Data Bank upon the conclusion of NEA Joint Projects – experience has shown that significant gaps remain. Despite some notable successes, there is growing consensus that existing approaches to data preservation can and should be improved. As a result, important datasets – or key elements within them – remain at risk of being lost.

This report identifies some of the most critical datasets across the technical domains reviewed and provides targeted recommendations to ensure their preservation. Given that the resources required for effective data preservation are considerable, the report also proposes a framework for prioritising actions across the different nuclear safety technical areas.

To avoid similar challenges in the future, the report highlights a set of generic best practices and lessons learnt from successful preservation efforts. These should be captured, documented and shared broadly to guide future initiatives and ensure a more systematic and sustainable approach to data preservation.

1.3. Procedure

The mode of operation of the SEGPD was structured around regular plenary meetings, bringing together nuclear safety experts from the CSNI as well as specialists from the NSC and MBDAV with experience in data preservation. Collectively, these experts represent a broad and in-depth understanding of nuclear safety research needs and priorities.

To support its work, the SEGPD developed a questionnaire focused on data preservation practices and challenges, which was distributed widely to experts outside the group. The questionnaire – reproduced in Appendix A – received responses from 20 individuals or organisations. The input gathered played a key role in shaping the content and structure of a virtual workshop held on 18-19 November 2021, which addressed a range of topics including database design, population, usage and the practical challenges associated with both existing databases and the absence of structured repositories. A summary of the workshop discussions is provided in Appendix B. These exchanges significantly influenced the SEGPD's deliberations during subsequent plenary meetings and are reflected in the current status assessment and recommendations presented in this report.

To broaden engagement and gather additional perspectives, the SEGPD's analyses and preliminary recommendations were also presented and discussed at various NEA meetings – including those of the CSNI working groups (WGAMA, WGFS, WGIAGE), and the CSNI, NSC and MBDAV committees. Furthermore, the SEGPD contributed to the January 2023 NEA Joint Nuclear Safety Research Projects Week (NEA, 2024), which provided an important forum for feedback and insight from a wide range of stakeholders.

Notably, participants at the 2023 Joint Projects Week event emphasised the importance of establishing structured frameworks for R&D and data management, and stressed the need to preserve and disseminate high-quality datasets produced through safety research projects. This need was particularly underscored in the areas of fuel safety, thermal-hydraulics and long-term operation (LTO) (NEA, 2024).

1.4. Structure of the report

This report provides an overview of data preservation efforts across several key disciplines relevant to nuclear safety, as detailed in Chapter 2. Specifically, it addresses:

- neutronics (Section 2.1);
- thermal-hydraulics (Section 2.2);
- severe accidents (Section 2.3);
- fuel performance (Section 2.4);
- ageing and long-term operation (Section 2.5);
- fire hazards (Section 2.6).

Each section highlights key datasets, presents examples of existing databases, assesses the current status of preservation efforts, and offers discipline-specific recommendations to strengthen future activities.

In Section 3.1, the report identifies common themes and challenges that emerge across the various technical areas and presents a set of overarching, high-level recommendations. The report concludes in Section 3.2 with a series of proposed near-term actions aimed at enhancing the preservation of legacy data and improving ongoing and future data preservation efforts carried out by NEA working groups and joint research projects.

1.5. References

NEA (2024), "Summary and Main Outcomes of the NEA Nuclear Safety Research Joint Projects Week", OECD Publishing, Paris, www.oecd-nea.org/jcms/pl_91291.

NEA (2022), "Mandate of the Senior Expert Group on Preservation of Key Experimental Datasets (SEGPD)", www.oecd-nea.org/tools/mandates/index/id/10874/lang/en_gb.

NEA (2021), *Nuclear Safety Research Support Facilities for Existing and Advanced Reactors: 2021 Update*, OECD Publishing, Paris, www.oecd-nea.org/7565.

Chapter 2. **Insights from data preservation experiences**

2.1. **Neutronics**

Within the framework of the Nuclear Science Committee (NSC), four key experimental databases related to various aspects of neutron physics are maintained by two of its expert working parties: the Working Party on Scientific Issues and Uncertainty Analysis of Reactor Systems (WPRS) and the Working Party on Nuclear Criticality Safety (WPNCS).

Under the auspices of the WPNCS, the following databases are maintained:

- The International Handbook of Evaluated Criticality Safety Benchmark Experiments (ICSBEP).
- The International Assay Data of Spent Nuclear Fuel Database (SFCOMPO).

Within the WPRS, the corresponding databases include:

- The International Handbook of Evaluated Reactor Physics Benchmark Experiments (IRPhE).
- The Shielding Integral Benchmark Archive and Database (SINBAD).

Each of the ICSBEP, IRPhE, and SFCOMPO databases is overseen by a dedicated Technical Review Group (TRG), which meets regularly to ensure the quality, consistency and relevance of the content. For SINBAD, a Task Force has been established to co-ordinate its ongoing development and enhancement.

It is important to note that most modern applications increasingly rely on **multi-physics modelling**, which offers a more realistic representation of reactor behaviour than traditional single-physics approaches. This shift has brought new emphasis to integrated benchmark data that reflect the complex interactions among various physical phenomena.

In this context, the WPRS has developed several benchmark cases derived from integrated experiments (IEs) and plant measurements and observations (PMOs). While these benchmarks are not yet compiled into a dedicated handbook, they are made available with detailed descriptions and commentary on the assumptions and simplifications required for modelling.

These include:

- Selected transients conducted in the Studsvik R2 fuel testing reactor.
- A benchmark based on PMO data from TVA Watts Bar Unit 1 (TVA-WB1), which is particularly useful for validating multi-cycle depletion modelling.
- Several transient benchmarks included in the LWR Uncertainty Analysis in Modelling (LWR-UAM) benchmark suite.

The following subsections provide a more detailed examination of the four major databases. Each will be discussed in terms of its content, the processes used for validation and maintenance, ongoing developments, key challenges, and opportunities for future improvements. Where relevant, interrelations and overlaps among the databases will also be highlighted.

2.1.1. *International Criticality Safety Benchmark Evaluation Project (ICSBEP)*

The ICSBEP (NEA, 2022; Bess, 2020; Bess, 2023) contains a comprehensive collection of criticality safety benchmark specifications derived from experiments conducted at critical facilities worldwide. These experiments are grouped into three main categories:

- critical, near-critical, or subcritical configurations (587 evaluations comprising 5 121 configurations);
- criticality alarm system configurations (45 configurations, each with multiple detector points);
- fundamental physics experiments relevant to criticality safety (237 configurations).

Additionally, the database includes over 800 configurations categorised as not of benchmark quality. Notably, the criticality alarm system configurations are essentially shielding experiments and may also serve as validation cases for radiation transport and shielding codes.

The ICSBEP is published as a compilation of PDF documents – each describing a benchmark or series of benchmarks – organised into nine volumes totalling approximately 80 000 pages. A companion Java-based application, DICE (Database for the International Criticality Safety Benchmark Evaluation Project), provides a searchable interface with options for basic statistical analysis.

Initiated by the US Department of Energy (DOE), the ICSBEP became an official NEA activity in 1995. It has since benefitted from contributions by experts and institutions in 22 countries.

A rigorous three-stage peer review process is applied to the inclusion of new benchmarks:

1. Initial evaluation, ideally conducted by personnel from the original experimental facility.
2. Internal review by an independent expert from the same institution.
3. External review by a separate institution.

The evaluation must follow strict guidelines, including uncertainty quantification. Once approved by the ICSBEP Technical Review Group (TRG), the evaluation is published in a standardised format comprising:

- description of the experiment and facility;
- evaluation of experimental data;
- benchmark specifications (geometry, materials, keff, uncertainties);
- sample calculation results;
- references and optional appendices (e.g. input files, logbooks).

Each benchmark is uniquely identified using a structured code: *(Fissile Material)-(Physical Form)-(Spectrum)-(Three-Digit Identifier)*

The ICSBEP is continuously updated with new benchmark contributions.

2.1.2. *International Reactor Physics Experiment Evaluation Project (IRPhE)*

The IRPhE (Bess, 2020; NEA, 2019) contains benchmark specifications based on reactor physics experiments from various international facilities. Many of these experiments are relevant to both reactor physics and criticality safety, and co-ordination with ICSBEP ensures consistency across both databases.

As of 2023, the IRPhE Handbook includes 169 experimental series from 57 facilities, of which 164 are accepted as benchmark quality. The database supports the validation of computational methods and nuclear data for reactor design, safety analysis, and physics evaluations.

Like ICSBEP, IRPhE is presented as a series of PDF documents and is supported by a Java-based application, the IRPhE Database and Analysis Tool (IDAT), which allows users to search benchmarks and perform statistical analysis.

Launched as a pilot project in 1999 and officially adopted by the NSC in 2003, IRPhE shares the same methodologies for evaluation, formatting and quality assurance as the ICSBEP. Typically, IRPhE benchmarks are based on a critical configuration with supplemental data such as:

- reactivity coefficients;
- neutron flux distributions;
- spectral indices;
- reaction rates;
- kinetic parameters.

Benchmark identifiers follow this format: *(Reactor Name)-(Reactor Type)-(Facility Type)-(Three-Digit Identifier)-(Measurement Types)*

The added complexity of reactor physics measurements requires tailored evaluations for each experiment. IRPhE remains under active development with regular additions of new data.

2.1.3. *SFCOMPO – Spent Fuel Isotopic Composition Database*

SFCOMPO (Michel-Sendis, 2017) is a structured database of measured isotopic compositions from spent nuclear fuel samples. It currently includes data from 44 reactors across 8 reactor types, predominantly light water reactors (LWRs), but also including AGR, Magnox, CANDU, and RBMK samples. In total, the database holds 750 fuel samples with over 24 000 measurement entries.

Each dataset is accompanied by fuel design and power history data, which are essential for validating burnup and depletion codes in criticality safety analyses involving burnup credit.

Originating from compilations made by JAERI in the 1990s, SFCOMPO became an NEA initiative in 1997. The current SFCOMPO 2.0 version, developed with ORNL, was released in 2017. It consists of an SQL database with a Java-based user interface that presents the data in a standardised format. Each entry typically includes:

- fuel assembly design (geometry, materials);
- sample and axial location within the rod;
- power history and irradiation conditions;
- measured nuclide concentrations, sample burnup, and measurement date;
- references to original reports and publications.

Not all samples contain complete information – some lack detailed power histories or extensive nuclide measurements. The number of measured nuclides varies widely, from uranium and plutonium only to complete inventories including minor actinides and fission products.

While the data have been verified for consistency and converted into calculation-ready units, most entries have not yet undergone the rigorous evaluation process applied in the ICSBEP. A standard review process is being developed by the TRG. This task is complex and requires multidisciplinary expertise, including reactor physics, radiochemistry and fuel performance analysis.

A recent development is the inclusion of decay heat measurement data for SNF assemblies. Although not directly related to criticality safety, decay heat is intrinsically tied to isotopic composition, making these data a logical addition to SFCOMPO.

Potential future expansions face challenges, including the proprietary nature of some fuel design data and limited public availability of experimental results.

2.1.4. *SINBAD – Shielding Integral Benchmark Archive and Database*

The SINBAD database (Kodeli, 2006; 2021) contains integral benchmark data for shielding applications, grouped into three primary areas:

- reactor shielding and pressure vessel dosimetry (47 datasets);
- fusion neutronics (31 datasets);
- accelerator shielding (23 datasets).

Its primary purpose is to support the validation of radiation transport and shielding calculation tools. SINBAD consists of interlinked HTML pages providing general experiment information and associated PDFs detailing the experiments. Supplementary files – such as code input decks – are included where available.

Established in 1996 as a joint initiative of ORNL/RSICC and the NEA, SINBAD is now managed by the Expert Group on Radiation Transport and Shielding (EGRTS) under the WPRS.

Each entry includes essential experimental details:

- geometry and material compositions;
- radiation source characteristics;
- detector types and placement.

The quality and completeness of entries vary. Some experiments have undergone detailed evaluations similar to ICSBEP standards, while others remain under-evaluated. The diverse user base – including experts in reactor shielding, fusion, and accelerator applications – makes comprehensive coverage a challenge.

Overlap exists between SINBAD and the ICSBEP, particularly regarding criticality alarm system configurations, which serve both criticality and shielding validation roles.

To strengthen SINBAD, a Task Force was established under WPRS. Current priorities at the time of publication include:

- adding new benchmark data;
- re-evaluating legacy experiments;
- supporting modern modelling capabilities (e.g. CAD geometries, variance reduction techniques).

Work is co-ordinated via the NEA GitLab server, which facilitates collaboration and task tracking. Benchmarks are now classified by maturity levels, enabling users to assess evaluation quality and reliability.

2.1.5. *General remarks*

Thanks to their rigorous and systematic evaluation methodologies, the ICSBEP and IRPhE projects represent best practices among the databases managed by the WPNCS and WPRS. The consistent structure of the benchmark data, coupled with a standardised review process, ensures high reliability and usability of the information for validation of computational tools.

However, applying this same approach to other types of nuclear data – such as spent fuel assays or shielding experiments – can be challenging. Criticality safety experiments are somewhat unique in that they typically involve well-defined configurations and standardised quantities (e.g. geometry, composition, keff), making it feasible to develop a uniform modelling framework. These characteristics also support systematic categorisation and benchmarking.

For other domains, such as fuel depletion or shielding, the diversity and complexity of measurements, varying levels of available metadata, and heterogeneous modelling assumptions make such standardisation more difficult. In these cases, it may be necessary to relax certain boundary conditions or adapt the evaluation procedures to accommodate the broader variability in experimental design and data availability.

An alternative or complementary strategy is the use of data assimilation techniques, which can help link experimental data with simulation results across different application domains. This approach could support cross-disciplinary validation and improve the overall coherence of modelling efforts in areas where direct benchmarking is less feasible.

From an organisational standpoint, the successful development and maintenance of these benchmark databases have relied heavily on the long-term commitment of participating institutions and experts. The NEA plays a central role, providing the technical infrastructure, ensuring co-ordination across countries and working groups, and fostering collaboration through dedicated expert bodies and technical review groups.

2.1.6. *References*

Bess, J.D. et al. (2023), "Intrinsic value of the international benchmark projects, ICSBEP and IRPhEP, for advanced reactor development", *Frontiers in Energy Research*, Vol. 11. https://doi.org/10.3389/fenrg.2023.1085788.

Bess, J.D. and T. Ivanova (2020), "Current overview of ICSBEP and IRPhEP benchmark evaluation practices", *EPJ Web of Conferences*, Vol. 239, 18002. https://doi.org/10.1051/epjconf/202023918002.

Kodeli, I.A. and E. Sartori (2021), "SINBAD – Radiation shielding benchmark experiments", *Annals of Nuclear Energy*, Vol. 159, 108254, https://doi.org/10.1016/j.anucene.2021.108254.

Kodeli, I. et al. (2006), "SINBAD Shielding Benchmark Experiments Status and Planned Activities", for the ANS 14th Biennial Topical Meeting of the Radiation Protection and Shielding Division, Carlsbad, NM, United States, 3-6 April 2006, www.oecd-nea.org/science//wprs/shielding/sinbad/RPSD2006-sinbad.pdf.

Michel-Sendis, F. et al. (2017), "SFCOMPO-2.0: An OECD NEA database of spent nuclear fuel isotopic assays, reactor design specifications, and operating data", *Annals of Nuclear Energy*, Vol. 110, pp. 779-788, https://doi.org/10.1016/j.anucene.2017.07.022.

NEA (2022), *International Handbook of Evaluated Criticality Safety Benchmark Experiments (ICSBEP)*, OECD Publishing, Paris, www.oecd-nea.org/jcms/pl_98160.

NEA (2019), *International Handbook of Evaluated Reactor Physics Benchmark Experiments (IRPhE)*, OECD Publishing, Paris, www.oecd-nea.org/download/science/wprs/irphep-handbook/2019.

2.2. Thermal-hydraulics

Thermal-hydraulics encompasses the study of fluid mechanics and heat transfer processes within nuclear reactors – from fuel to coolant and ultimately to the heat sink – under normal operation, anticipated operational occurrences and design basis accidents. Phenomena during beyond design basis accidents are generally addressed within the field of severe accident analysis, covered in the next section. The primary objective of thermal-hydraulic experimental facilities and computational codes is to demonstrate that emergency safety systems meet their design performance criteria (e.g. avoiding significant fuel damage and preventing containment failure).

Decades of international research have built a robust foundation of knowledge in this area, including both separate effects tests (SETs) and integral effects tests (IETs) that cover reactor coolant system (RCS) and containment thermal-hydraulic behaviour. These data have supported the development and validation of various thermal-hydraulic code systems – from high-resolution computational fluid dynamics (CFD) tools to subchannel codes for detailed core analysis, and system codes that simulate the global reactor behaviour, including interactions with auxiliary and secondary systems (e.g. in pressurised water reactors [PWRs]).

2.2.1. *Key datasets and existing databases*

Thermal-hydraulic experiments have been carried out in numerous facilities worldwide since the beginning of the nuclear power era. These experiments typically measure essential parameters such as fluid flow rates, velocities, wall and fluid temperatures, pressure, void fractions, and film thickness. Advanced diagnostic techniques like tomography, radiography, planar laser-induced fluorescence, and high-speed imaging have enabled detailed flow visualisations. Additional experiments address wall heat transfer, mass transfer (e.g. condensation and evaporation), water hammer, cavitation, and flow-induced vibration of structures.

Key thermal-hydraulic datasets have been systematically compiled through successive efforts led by the NEA. Major references include reports published from 1987 onward (NEA, 1987; 1993; 1996a; 1996b; 2001; 2016), and more recently, the online International Experimental Thermal Hydraulics Systems Database (TIETHYS) (NEA, 2024a). Supplementary datasets can be accessed through the NEA Data Bank (NEA, 2024b). Together, these sources provide a comprehensive overview of experimental facilities and test results relevant to PWRs, BWRs and VVERs under transient and accident conditions.

2.2.2. *Data preservation efforts*

The NEA and its working groups have led major initiatives to collect and preserve thermal-hydraulic data through benchmark matrices. The initial CSNI Code Validation Matrix (CCVM) report (NEA, 1987) focused on integral effects tests. This was followed by the 1993 Separate Effects Test Matrix (NEA, 1993), and subsequent updates to the integral validation matrices in 1996 (NEA, 1996b) and for VVERs in 2001 (NEA, 2001).

The Separate Effects Test Matrix provides descriptions of facilities, test parameters, relevant thermal-hydraulic phenomena, and available data for code validation. A cross-reference matrix links facilities to the phenomena they address. These phenomena are applicable not only to PWRs and BWRs but also to advanced water-cooled reactor designs.

The Integral Effects Test Matrix applies a similar methodology, organising data according to facility type, test type, and phenomena covered. Test selection criteria include relevance to expected reactor conditions, data completeness and overall quality. High priority is placed on International Standard Problems (ISPs) and counterpart tests, given their typically rigorous data validation and international collaboration.

While TIETHYS and the Data Bank store much of this information, limitations remain. Some facilities and tests are not included, documentation may be incomplete, and standardisation across datasets is lacking. Many tests date back several decades, and some original documentation has been lost or remains proprietary.

2.2.3. *Analysis of the current situation*

The SET and IET validation matrices serve as a vital foundation for best-estimate (BE) code validation. They enable a quantitative assessment of predictive capability and uncertainty analysis, provided the experimental uncertainties are well characterised. TIETHYS allows users to navigate these matrices and link tests to relevant phenomena. The NEA Data Bank currently holds validated data from around 69 integral experiments and 30 major SETs, available to organisations from NEA member countries.

However, several challenges persist:

- Not all key experiments are documented or included in TIETHYS or the CCVM reports.
- Legacy datasets are at risk due to facility decommissioning and expert retirements.
- The lack of a standardised format for test data limits reusability for modern code validation.

To address these gaps, the NEA Working Group on the Analysis and Management of Accidents (WGAMA) launched the THEMPo (Thermal-Hydraulics Experimental Meta-Data Preservation, Collection and Qualification) initiative. This effort aims to develop a harmonised methodology for the collection, qualification, organisation and use of experimental data, including:

- facility hardware geometry;
- investigated phenomena;
- initial and boundary conditions;
- measurement data and uncertainties.

This methodology will guide the creation of a relational database of experimental metadata, building on TIETHYS and aligned with modern modelling needs. The resulting standard will facilitate faster, more reliable development of validation models for thermal-hydraulic codes.

2.2.4. *Recommendations*

A co-ordinated, long-term framework is essential to secure the future of thermal-hydraulic data preservation. The following actions are recommended:

- Preserve legacy datasets:
  - Identify and collect at-risk datasets, particularly those supporting accident management, safety system validation (including passive systems), and next-generation reactor applications.
  - Prioritise datasets previously used in benchmarks or ISPs, as they typically meet completeness and quality standards.
- Establish benchmark activities:
  - Launch new benchmarks and ISPs using legacy data to exercise and qualify the datasets.
  - Encourage joint projects to include dedicated benchmark activities for newly generated experimental data.
- Address data gaps for advanced reactors:
  - Work with expert groups (e.g. CSNI/WGAMA) to identify missing datasets relevant to SMRs and Generation IV technologies.

- Leverage AI tools:
  - Explore the use of artificial intelligence and machine learning to support metadata integration, pattern recognition and data mining for incomplete datasets.
- Implement THEMPo:
  - Apply and validate the THEMPo methodology on key historical datasets as a priority.
  - Use the resulting framework to guide current and future joint projects in data structuring, verification and qualification.
- Ensure early agreement on data dissemination:
  - During joint projects, reach consensus on long-term data preservation and sharing strategies.
  - Put distribution agreements in place before project completion.

Preservation efforts will require substantial resources, including expert time and digital infrastructure. Stakeholders across the nuclear sector should recognise the long-term cost benefit of supporting a shared, high-quality thermal-hydraulic data foundation to ensure safe deployment of existing and future reactor technologies.

2.2.5. *References*

NEA (2024a), TIETHYS – The International Experimental Thermal Hydraulics Systems Database, www.oecd-nea.org/tiethysweb.

NEA (2024b), NEA Data Bank catalogue of experiments, databases, and projects, www.oecd-nea.org/jcms/rni_6525/data-bank.

NEA (2016), "A State-of-the-Art Report on Scaling in System Thermal-hydraulics Applications to Nuclear Reactor Safety and Design", OECD Publishing, Paris, www.oecd-nea.org/jcms/pl_19744.

NEA (2001), "Validation Matrix for the Assessment of Thermal-Hydraulic Codes for VVER LOCA and Transients", OECD Publishing, Paris, www.oecd-nea.org/jcms/pl_17492.

NEA (1996a), "Evaluation of the Separate Effects Tests (SET) Validation Matrix", NEA/CSNI/R(1996)16, OECD Publishing, Paris, www.oecd-nea.org/jcms/pl_16132.

NEA (1996b), "CSNI Integral Test Facility Validation Matrix for the Assessment of Thermal-Hydraulic codes for LWR LOCA and Transients", OECD Publishing, Paris, www.oecd-nea.org/jcms/pl_16134.

NEA (1993), "Separate effects test matrix for thermal-hydraulic code validation, Vol. 1 and 2", OECD Publishing, Paris, www.oecd-nea.org/jcms/pl_15968.

NEA (1987), "CSNI Code Validation Matrix of Thermo-Hydraulic Codes for LWR LOCA and Transients", OECD Publishing, Paris, www.oecd-nea.org/jcms/pl_15784.

2.3. Severe accidents

The research in the area of severe accidents has primarily focused on understanding key phenomena and assessing the effectiveness of severe accident management (SAM) measures in large-scale light water reactors (LWRs), including PWRs, BWRs, and VVERs, and to a lesser extent, in pressurised heavy water reactors (PHWRs) such as CANDU reactors. These efforts have significantly supported the development of SAM strategies for Generation III and III+ reactors. Most experimental campaigns and datasets pertain to core-originated severe accidents, with only limited studies addressing scenarios that initiate in spent fuel pools (SFPs).

The current state of knowledge reflects decades of experimental research, code development, and insights gained from real events – most notably the Three Mile Island Unit 2 (TMI-2) and Fukushima Daiichi accidents. Although the Chernobyl accident is less relevant to Western water-cooled reactor designs, it still provides valuable information, especially for long-term aspects such as radioactive release from ageing fuel-containing materials like lava-like fuel-concrete matrices.

2.3.1. *Key datasets and existing databases*

Severe accidents involve a highly complex sequence of coupled physical and chemical processes, starting with core heat-up and progressing through material degradation, melting, relocation within the pressure vessel, potential vessel failure, and ex-vessel phenomena. The primary safety concerns include radiological source term behaviour, integrity of the reactor pressure vessel and containment, and the ability to cool degraded fuel both in-vessel (IVMR) and ex-vessel (EVMR).

In spent fuel pool accidents, important phenomena include fuel uncovery, heat-up and degradation of fuel assemblies and racks, potential criticality, and fission product release. While many of the relevant phenomena are similar to in-core accidents, SFP events often involve increased potential for air ingress scenarios.

Given the wide range of disciplines involved – including thermal-hydraulics, material science, chemistry, aerosol physics, and combustion science – there is no single comprehensive database for severe accident research akin to those developed for neutronics or thermal-hydraulics. Efforts such as the NEA CSNI Code Validation Matrix (CCVM) and the European Commission's STRESA database have faced challenges in consolidating data due to proprietary restrictions, incomplete documentation, and the fragmented nature of data collection.

Severe accident datasets span multiple categories:

- Fundamental data from small-scale analytical experiments on basic phenomena such as heat transfer, oxidation kinetics, chemical reactions (eutectics, iodine chemistry), aerosol behaviour, etc.
- System-level experimental data from reduced-scale but coupled-phenomena facilities specific to severe accident progression, e.g.:
  - Containment thermal-hydraulics and safety system efficiency: BUBCON, HYMERES, PANDA, SETH, THAI, THEMIS.
  - Hydrogen/CO combustion risk in containment: SETH, THAI, THEMIS, HYMERES.

  - Core melt progression and retention: CORA, QUENCH, MASCA, MCCI, OLHF, RASPLAV, ROSAU, SERENA.
  - Fission product release from fuel: VEGA, VERCORS, VERDON.
  - Fission product transport and deposition: BIP, ESTER, LOFT-FP, THAI, THEMIS.
- Integral experiments simulating complex accident sequences across multiple systems, most notably the PHEBUS-FP programme, which has been foundational for validation of severe accident codes and ISP activities.

Integral experiments involving radioactive materials are costly and technically demanding. Many related facilities have been dismantled or are under threat of closure, making preservation of existing datasets an urgent priority.

Past accident data (TMI-2, Fukushima Daiichi) also remain invaluable for code benchmarking and validation (e.g. NEA, 2009; 2021), and integrated severe accident system codes (MELCOR, MAAP, ASTEC) serve as repositories of collective knowledge, developed and validated against a wide array of data sources. These codes often incorporate stand-alone modules for specific phenomena (e.g. MCCI).

Severe accident data used in code validation matrices are typically preserved by the organisations responsible for code development, but broader and durable preservation mechanisms are still needed.

2.3.2. *Data preservation efforts*

The NEA WGAMA made significant contributions towards consolidating severe accident data through the CCVM effort, culminating in the 2014 validation matrix report (NEA, 2014). The European Commission developed the STRESA database as a complementary initiative (Pla Freixa, 2015).

Additional NEA efforts include the collection of data packages from several joint projects, available on request from the NEA Data Bank (NEA, 2024), such as:

- Containment TH and combustion: BUBCON, HYMERES, THAI, SETH.
- Core melt progression and coolability: MASCA, MCCI, OLHF, RASPLAV, SERENA.
- Fission product release and transport: BIP, LOFT-FP, STEM, THAI.
- Accident analyses: TMI-VIP, BSAF.

Ongoing joint projects (e.g. ESTER, PANDA, QUENCH-ATF, ROSAU, THEMIS) will similarly have their datasets archived at the NEA Data Bank once confidentiality periods lapse.

The NEA also preserves datasets from integral experiments used in ISPs (e.g. CORA, FALCON, FARO, LOFT-FP, MARVIKEN), which are available upon request.

Efforts to collect and manage accident data have also focused on historical events. For TMI-2, key resources include the TMI-2 Data Summary Report (McCormick, 1987), the NUREG Knowledge Management Digest (2022), and the INL Knowledge Management Library (2016). For Fukushima Daiichi, the ARC-F and FACE projects continue to support data collection and knowledge management through to at least 2026.

In the domain of fundamental chemistry, NEA is co-ordinating international efforts to build thermodynamic databases critical for severe accident modelling, including:

- TAF-ID – focused on equilibrium calculations of in-/ex-vessel corium phases and fission product chemistry.
- TCOFF – supporting the development and benchmarking of thermodynamic data through experimental studies and database comparisons (e.g. TAF-ID and NUCLEA).

These databases are essential for simulating advanced materials (e.g. ATFs) and fuels for Generation IV designs. The COPS project (started May 2024) will investigate corium properties, further supporting code development and validation.

At the EU level, the 2023 launch of the SEAKNOT project is another major milestone. It aims to define a severe accident research roadmap, including a comprehensive review of validation databases for current and near-term reactor technologies.

2.3.3. *Analysis of the current situation*

Before CCVM and STRESA, there was no co-ordinated effort to preserve severe accident datasets. These initiatives, while important, were hampered by access restrictions, lack of systematic documentation, and limited stakeholder engagement. Many datasets remain in the custody of originating institutions, often without long-term preservation strategies. In some cases, internal databases have become obsolete or inaccessible, and important experimental facilities have been dismantled.

Even among the datasets preserved at the NEA Data Bank, not all have undergone thorough review for completeness or quality assurance. Feedback from users of the Data Bank has not been consistently collected or used to update or supplement existing documentation.

While the NEA and European Commission have acknowledged the need for co-operation (e.g. NEA-SEAKNOT), no formal mechanisms exist to ensure co-ordinated long-term preservation post-project completion.

2.3.4. *Recommendations*

It is recommended to establish a co-ordinated and sustained collaborative framework for the long-term preservation of severe accident datasets. This framework should:

- Address legacy data preservation gaps, building on CCVM, STRESA, and current gap analyses at the NEA and SEAKNOT. This includes:
  - Identifying and preserving, at the NEA Data Bank, datasets critical to SAM analysis for current and future reactors.
  - Prioritising datasets linked to key safety strategies (IVMR, EVMR, hydrogen mitigation), severe accident code validation matrices, and applications to new reactor technologies.
  - Reviewing completeness and quality-assurance levels of datasets, especially those not previously exercised through benchmarks or ISPs.

- Propose and co-ordinate new benchmark activities and ISPs, especially using legacy datasets, to reinforce validation and quality-assurance.
- Identify future data needs, especially for advanced technologies (e.g. SMRs, Generation IV), with expert support (e.g. CSNI/WGAMA).
- Develop a co-operation mechanism with SEAKNOT, including discussion of the NEA Data Bank's potential role as the long-term repository of reference datasets.
- Explore the use of artificial intelligence (AI) to facilitate metadata integration and cross-referencing across fragmented datasets.

For current and future NEA Joint Projects, it is essential to:

- Integrate data preservation planning into project execution phases, with clear dissemination and distribution agreements in place.
- Define the structure and documentation of data packages to be submitted to the NEA Data Bank.
- Include benchmark activities within projects to validate and exercise experimental data.

Regarding thermodynamic data, continued co-operative efforts are recommended to support the development and qualification of datasets relevant for new fuels and reactor designs. Projects like TAF-ID, TCOFF and COPS are vital in identifying modelling gaps and building high-quality data repositories.

For real accident data (e.g. Fukushima Daiichi), the continued engagement of Japanese organisations through projects such as ARC-F and FACE is essential. A long-term dissemination strategy for these datasets should be established and agreed with all partners.

2.3.4. *References*

Clément, B. et al. (2005), "Thematic network for a Phebus FPT1 international standard problem (THENPHEBISP)", *Nuclear Engineering and Design*, Volume 235, Issues 2-4, pp. 347-357, https://doi.org/10.1016/j.nucengdes.2004.08.057.

Guéneau, C. et al. (2021), "TAF-ID: An international thermodynamic database for nuclear fuels applications", *Calphad*, Vol. 75, 102212, https://doi.org/10.1016/j.calphad.2020.102212.

INL (2016), "The Knowledge management library for the Three Mile Island Unit 2 Accident of 1979", https://tmi2kml.inl.gov/HTML/Page1.html.

McCormick, R.D., J.L. Anderson and D.W. Golden (1987), *TMI-2 Data Summary Report*, EGG-TMI-7843, EG&G Idaho, United States, https://doi.org/10.2172/6120550.

NEA (2024), NEA Data Bank catalogue of integral experiments data, databases, benchmarks and safety joint projects, www.oecd-nea.org/tools/ie/list.

NEA (2021), *Benchmark Study of the Accident at the Fukushima Daiichi Nuclear Power Plant: Summary Report*, OECD Publishing, Paris, www.oecd-nea.org/7525.

NEA (2015), "Status Report on Spent Fuel Pools under Loss-of-Coolant Accident Conditions Final Report", OECD Publishing, Paris, www.oecd-nea.org/jcms/pl_19596.

NEA (2014), "Containment Code Validation Matrix Report", OECD Publishing, Paris, www.oecd-nea.org/jcms/pl_19413.

NEA (2009), "Ability of Current Advanced Codes to Predict Core Degradation, Melt Progression and Reflooding: Benchmark Exercise on an Alternative TMI-2 Accident Scenario", OECD Publishing, Paris, www.oecd-nea.org/jcms/pl_18720.

NUREG (2022), "Three Mile Island Accident of 1979 Knowledge Management Digest", NUREG/KM-0001, US Nuclear Regulatory Commission, Washington, www.nrc.gov/reading-rm/doc-collections/nuregs/knowledge/km0001/index

Pla Freixa, P., G. Pascal, J. Tanarro Colodron and A. Annunziato (2015), "The new STRESA tool for preservation of thermalhydraulic experimental data produced in the European Commission", *Annals of Nuclear Energy*, Vol. 85, pp. 893-902, https://doi.org/10.1016/j.anucene.2015.07.007.

2.4. Fuel performance in normal operation, anticipated operational occurrences and design basis accidents

This section discusses datasets and databases for fuel performance in normal operation, anticipated operational occurrence (AOO) and design basis accident (DBA) conditions. The main discussion is on water-cooled reactor (WCR) fuel, but similar considerations will apply to other fuel types, including pressurised heavy water reactor (PHWR) fuel, advanced gas-cooled reactor (AGR) fuel, fast reactor (FR) fuel, and high-temperature reactor (HTR) fuel. However, given the worldwide predominance of WCRs in terms of commercial power reactors, the datasets and databases associated with other fuel types are less common than for WCRs, and, in the case of FR and HTR fuel, tend to be historical. This specific point is returned to later in the recommendations.

2.4.1. *Datasets and databases in general terms*

Datasets associated with fuel performance in normal operation and anticipated operational occurrences (AOOs) primarily contain data from experimental irradiations of short test rods ("integral" tests) in material test reactors (MTRs), or from commercial irradiations of full-length rods in power reactors. The commercial irradiations could be of lead test rods (LTRs), rods in lead test assemblies (LTAs), or rods in reload assemblies (all as part of relevant surveillance programmes).

Datasets associated with fuel performance under DBA conditions may take place in-pile in MTRs, out-of-pile in a hot cell facility, or out-of-pile and out-of-cell. More specifically, reactivity-initiated accident tests are performed in special-purpose MTR facilities capable of generating rapid power pulses (NEA, 2022b), while tests simulating loss-of-coolant accident (LOCA) conditions have been performed in both MTRs and in out-of-pile facilities (NEA, 2009; NEA 2024a). RIA and LOCA tests may be performed on irradiated fuel rods or on fresh fuel rod samples. Typically, RIA and LOCA tests are performed on short rod segments, so full-length commercial rods need to be refabricated into shorter segments before performing the DBA tests.

Data from in-pile or out-of-pile separate-effects tests (for example, creep tests) on fuel or cladding materials are also relevant to fuel and cladding performance under normal, AOO, and DBA conditions, but are only described in passing here given these are primarily used to generate material property data.

Data from experimental irradiations include (a) details of the fuel rod design and manufacturing parameters, including any data on pre-characterisation and (where relevant) pre-irradiation; (b) details of the irradiation performed (including, where relevant, the experimental procedure); (c) "online" (in-reactor, through-life) measurements from experiment instrumentation; and (d) post-irradiation examination (PIE) data. Data from commercial irradiations include only (a), (b) and (d) (since fuel irradiated commercially cannot contain the detailed instrumentation used in MTRs). The online data include measurements of fuel and cladding temperatures via thermocouples, of fuel and cladding elongation via extensometers, and of rod internal pressure via pressure transducers. Similar measurements are made during DBA tests, both for those conducted in MTRs and those conducted out-of-pile. These tests may include additional sensors to determine the timing of fuel rod failure or to measure the mass of hydrogen generated by high-temperature steam oxidation of the cladding.

The PIE data include those from both non-destructive and destructive examination, and from both poolside inspections and hot cell examinations. Non-destructive PIE data include rod length measurements via metrology; fuel stack length measurements via gamma scanning; clad diameter measurements via profilometry; and clad oxide thickness measurements via an eddy current technique. Destructive PIE data include rod internal pressure, fission gas release and rod free volume measurements via rod puncture; measurements of sizes and concentrations of microstructural features (like fission gas bubbles or dislocations), or of details of macroscopic features (like grain sizes and crack patterns), via ceramography or metallography; actinide and fission product isotope or element concentration measurements via electron probe microanalysis (EPMA), X-ray fluorescence (XRF), secondary ion mass spectrometry (SIMS) and other techniques; and fuel density measurements via immersion.

Online data (and in some cases PIE data) consist of "raw data" and "processed data". The raw data comprise instrument signals, for example a thermocouple reading in millivolts, while the processed data are the raw data converted to the quantity to be measured, for example temperature. The raw data are immutable ("fixed"), while the processed data may change with time if the conversion process (which in turn is based on instrument calibration) is modified. Conversion can be complicated by irradiation effects, for example thermocouple decalibration (where the signal to temperature proportionality changes with neutron fluence as the elemental composition of the thermocouple wires changes due to transmutation) or pressure transducer drift (where irradiation creep of the pressure transducer bellows changes the compression of the bellows for a given pressure, and hence changes the measured signal). Furthermore, some raw data to processed data conversions involve calculations; of particular importance for fuel performance data is the conversion of neutron detector signals to rod powers via neutronics code calculations (where powers are not otherwise determined, for example by calorimetric means).

Electronic "databases" maintained for fuel performance in normal operation, AOOs, and DBAs include "data collections", "formal databases" and "code validation databases". There are also still paper archives of reports covering old experiments, and data retained in non-electronic media (for example, paper traces of instrument readings) in some instances (in particular where data acquisition pre-dated electronic recording and storage).

Data collections comprise archived datafiles and reports in some prescribed structure, for example a simple electronic folder structure, or nested webpages; formal databases are data collected in computer software for data management (for example, Microsoft Access); and code validation databases are input files and measured data files associated with irradiations used for fuel performance code validation, often combined with aspects of a data collection (for example, archived experimental reports) or a formal database (for example, a "front-end" electronic index of validation cases and their details). Formal databases are typically of the relational variety, where common data fields for irradiations are defined and populated in a structured manner for each irradiation. Other varieties, for example data lakes, where unstructured data are maintained and accessed via advanced search techniques, also exist.

2.4.2. *Specific datasets and databases*

Datasets for fuel performance in normal operation, AOOs and DBAs come from (i) international research programmes, (ii) national research programmes, (iii) fuel vendor test irradiations, and (iv) PIE of commercial fuel. The datasets (i)-(iii) are generally associated with irradiations carried out in material test reactors (MTRs), though (i) and (ii) also include DBA tests performed out-of-pile in hot cell facilities.

Important international research programmes include:

- the NEA Halden Reactor Project (HRP);
- the NEA Studsvik Cladding Integrity Project (SCIP);
- various other Studsvik programmes focusing on the effects of power ramping (INTER-RAMP, OVER-RAMP, DEMO-RAMP, SUPER-RAMP and TRANS-RAMP), rod overpressure (ROPE) and behaviour of defective fuel (DEFEX);
- the NEA Second Framework for Irradiation Experiments (FIDES-II) programme and its associated Joint Experimental Programmes (JEEPs) involving fuel performance during normal operations, AOOs, and design basis accidents;
- the Jules Horowitz Reactor (JHR) Joint International Programmes (JIPs);
- the Risø Fission Gas Release Projects;
- the Nuclear Fuel Industry Research (NFIR) programme led by the Electric Power Research Institute (EPRI);
- the Gadolinia Fuel Investigation to Extended Burnups (GAIN) led by Belgonucléaire and the Belgian Nuclear Research Centre (SCK CEN);
- the programme on Tests Relative to High Burnup Limitations Arising from Transient Incidents Occurring Normally in LWRs (TRIBULATION) led by Belgonucléaire and SCK CEN;
- the High Burnup Effects Program (HBEP) led by Battelle Pacific Northwest Laboratories (BNW);
- the PWR Reference Irradiation of MOX fuels (PRIMO) programme led by Belgonucléaire and SCK CEN, and a follow-up programme on mixed oxide (MOX) fuel behaviour, FIGARO;
- the REP-Na test programme and the CABRI International Program (CIP) involving RIA tests in the CABRI reactor in France and the NSRR reactor in Japan.

HRP, SCIP, FIDES-II, CIP and NFIR are ongoing programmes (though the fuel and materials portion of HRP is currently winding down due to the 2018 closure of the Halden reactor). The JHR JIPs are future planned programmes. The remaining items are historical research programmes that were completed some years ago. The corresponding MTRs include the Halden Boiling Water Reactor (HBWR) in Norway; the Studsvik R2 reactor in Sweden; the Risø DR3 reactor in Denmark; the BR2 and BR3 reactors in Belgium; the Osiris reactor, CABRI, and JHR in France; TREAT and ATR in the United States; NSRR in Japan; and the LVR-15 reactor in Czechia. Many of these reactors have been shut down; BR2, LVR-15, CABRI, ATR, TREAT, and NSRR are operating, and JHR is under construction.

The primary database that is internationally available for fuel performance is the NEA/IAEA International Fuel Performance Experiments (IFPE) database [NEA, 2024b]. There are also databases associated with international research programmes and national research programmes which are accessible only by the member organisations. Finally, fuel vendors and other fuel performance code developers (often national research organisations or regulatory bodies) maintain their own proprietary databases, with a focus on applications for fuel performance code validation.

The IFPE database is a data collection, but it is "coupled" to a formal database DATIF (Database for IFPE), which is of the relational kind (NEA, 2024c). The data are mainly rod irradiation data, but some separate-effects test data are also included. Both IFPE and DATIF are under active maintenance by NEA under NEA Data Bank quality assurance (QA) procedures. However, DATIF is still under development, and so currently contains only a fraction of the data in IFPE, and only those associated with rod irradiations (not separate-effects tests, given the different data field requirements). IFPE contained as of 2023 ~70 data packages covering irradiations of ~1 450 rods, fuel specimens and cladding specimens. The irradiations are from many national and international research programmes, as well as from some commercial irradiations. Copies of selected IFPE data packages, of all IFPE data packages, or of DATIF can be requested by NEA Data Bank signatories and then installed locally at signatory organisations (no remote access to data package contents or DATIF).

The HRP database has historically been a bespoke electronic system for storage of measured data (the Test Fuel Data Bank, or TFDB), with paper and electronic archives of reports and other documentation (for example, technical drawings), but, from 2021, has been replaced with the HRP Legacy Data Lake containing measured data, reports, PIE results and other documentation as well as a search engine tool under paying license. Due mainly to the maintenance costs (including licence fees) and software dependencies associated with the Data Lake, the HRP Legacy Data Lake was discontinued at the end of 2023 and replaced with a simpler database containing all material collected in the Data Lake and an open-source semantic search engine tool. HRP members were provided at the end of 2023 with a hard drive containing this HRP database and the open-source search engine. The data base has also been provided to the NEA at the end of 2023. Copies are installed locally by HRP member organisations. As of early 2024, the NEA will maintain the master copy of the database. The NEA was also granted the right to distribute the HRP database upon request to users in NEA member country organisations.

IFPE/DATIF and the HRP Legacy Database are considered examples of current best practice in the preservation of data on fuel performance in normal operation and AOOs.

Data from some international research programmes (in particular HRP), national research programmes and commercial irradiations have been used in a number of international benchmark exercises for WCR fuel performance modelling. Benchmark

exercises performed under the NEA and IAEA – which have in some cases also involved preparation of data packages for IFPE – include:

- MOX fuel performance benchmarks performed under the remit of NEA's Expert Group on Reactor-based Plutonium Disposition (TFRPD); irradiations included a PRIMO rod irradiation, the HRP irradiation of two MOX rods, and the US Department of Energy's weapons-grade MOX fuel irradiation performed in the Advanced Test Reactor (ATR);
- code-to-code and code-to-data comparisons carried out as part of IAEA's D-COM (1982-1984), FUMEX (1993-1996), FUMEX-II (2002-2007), FUMEX-III (2008-2012), and FUMAC (2014-2018) co-ordinated research projects (CRPs) (Killeen, 2007; Pizzocri, 2018; Zhang, 2018; IAEA, 2019);
- pellet-cladding mechanical interaction (PCMI) benchmarks performed under the remit of the NEA Expert Group on Reactor Fuel Performance (EGRFP); irradiations included HRP irradiations carried out in the IFA-118 and IFA-629.4 experiments;
- a series of code-to-code and code-to-data comparisons of reactivity-initiated accident tests performed under the NEA's Working Group on Fuel Safety (WGFS), conducted between 2009 and 2020 and summarised in (NEA, 2022a).

A burst fission gas release benchmark, using data from the CEA's REGATE and HATAC experiments, was completed in October 2024 under the remit of the EGRFP; a benchmark exercise to evaluate the completeness of the HRP database was recently launched by the WGFS and was completed in June 2025, as was work to use advanced technology/accident-tolerant fuel (ATF) irradiation and separate-effects test data in benchmark exercises as part of the IAEA's CRP on Testing and Simulation for Advanced Technology and Accident Tolerant Fuels (ATF-TS) (Zhang, 2024). In addition, benchmarks have been carried out, are being carried out, or are planned to be carried out as part of some of the international research programmes described above (including SCIP and FIDES-II).

Code benchmarks have also been performed for fast reactor fuel and for TRISO fuel through the NEA Expert Group on Innovative Fuels (EGIF) and as part of various IAEA CRPs, but such efforts are beyond the scope of this report.

2.4.3. *Analysis of the current situation*

The preceding section describes various repositories and databases relevant to the fuel performance area. At present, only the IFPE database is largely accessible in electronic form to organisations in NEA member countries. Since 2024, the NEA also provides electronic access to Halden Reactor Project data. Data from other projects are available only to the project members (e.g. the SCIP and CIP SharePoint sites). Fuel vendor databases are inaccessible to the public, but data are likely well preserved for commercial interests.

In the case of older experimental programmes, data may only be described in project reports and may not be available in a useable format for computer code development and validation. However, there have been efforts to preserve some of these old data; IFPE was created in part for this purpose. Other efforts include NEA Working Group on Fuel Safety state-of-the-art reports on fuel performance under RIA (NEA, 2022b) and LOCA (NEA, 2009; NEA, 2024a) conditions, which compile information about past experiments and provide some of the key experimental results.

One major challenge is that data may not include all the information needed for computer code validation. There is often limited information about the as-fabricated fuel (e.g. grain size, pore distribution, etc.) and about the fuel irradiation history, including the thermal-hydraulic boundary conditions (e.g. fuel rod linear heat rate vs. time, coolant temperature and flow rate). This is not only a problem with older programmes but also with ongoing joint projects, in part due to the proprietary nature of fuel fabrication processes and of commercial reactor operations. Fuel rod segments used in many integral tests are first irradiated in a commercial reactor before testing in an MTR or a hot cell. Assumptions must be made when evaluating the data, which increases the uncertainty in any conclusions drawn from the experiments.

A second challenge is that many legacy datasets do not include information about experimental uncertainties. As mentioned, there are uncertainties due to a lack of fuel fabrication and irradiation history data, but there are also measurement uncertainties during the tests and uncertainties introduced when the raw instrument data are converted into processed data. Uncertainties are not always provided in project reports or with the experimental data.

There have been recent efforts to elaborate the information that should be collected as part of fuel performance experiments, as well as the information needed for code validation (e.g. Section 4 of Van Uffelen, 2019). The NEA Expert Group on Reactor Fuel Performance is currently writing a report on experimental data requirements for fuel performance modelling, while the NEA Working Group on Fuel Safety has an ongoing activity whose objective is to specify the information that should be included in a database of RIA experiments. Such activities are expected to provide guidance on data collection and preservation to ongoing joint projects and to owners of legacy data.

2.4.4. *Recommendations*

In the fuel performance area, the top priority is to collect and preserve data from joint safety research projects, document the quality assurance level and agree on the long-term distribution for the data. This is currently being addressed for the Halden legacy database through a WGFS activity and by the NEA.

Moving forward, strategies for data integration in databases need to be commonly decided. For example, EGRFP and WGFS must consider whether to include legacy datasets (e.g. from the Halden Reactor Project) in IFPE and add them to DATIF or whether to store them in a different database. The groups must consider the specific needs for different applications (e.g. the WGFS RIA database activity, FIDES-II, and other needs).

2.4.5. *References*

IAEA (2019), *Fuel Modeling in Accident Conditions (FUMAC)*, IAEA-TECDOC-1889, International Atomic Energy Agency, Vienna, www.iaea.org/publications/13604.

Killeen, J.C., J.A. Turnbull, E. Sartori (2007), "Fuel Modelling at Extended Burnup: IAEA Coordinated Research Project FUMEX-II", *Proceedings of the 2007 International LWR Fuel Performance Meeting, San Francisco, California, 30 September-3 October 2007*, Paper 1102, www.osti.gov/biblio/21229289.

NEA (2024a), *State-of-the-art Report on Nuclear Fuel Behaviour in Loss-of-Coolant Accident (LOCA) Conditions*, OECD Publishing, Paris, www.oecd-nea.org/7483.

NEA (2024b), "International Fuel Performance database", www.oecd-nea.org/jcms/pl_36358.

NEA (2024c), "DATabase for IFpe (DATIF)", www.oecd-nea.org/jcms/pl_33943/database-for-ifpe-datif.

NEA (2022a), *Reactivity-Initiated Accident Fuel Rod Codes Benchmark Phases I-III: Synthesis Report*, OECD Publishing, Paris, www.oecd-nea.org/7577.

NEA (2022b), *State-of-the-art Report on Nuclear Fuel Behaviour under Reactivity-initiated Accident Conditions*, OECD Publishing, Paris, www.oecd-nea.org/jcms/pl_74371.

NEA (2010), "Benchmarking Calculations on Halden IFA-650 LOCA Test Results", OECD Publishing, Paris, www.oecd-nea.org/jcms/pl_18944.

NEA (2009), "Nuclear Fuel Behaviour in Loss-of-coolant Accident (LOCA) Conditions: State-of-the-art Report", OECD Publishing, Paris, www.oecd-nea.org/jcms/pl_18878.

Pizzocri, D. et al. (2018), "IAEA FUMAC Benchmark on the Halden, Studsvik and Quench-L1 LOCA tests", TOPFUEL 2018 Conference, Prague, Czechia, www.researchgate.net/publication/326557541_IAEA_FUMAC_Benchmark_on_the_Halden_Studisvik_and_QUENCH-L1_LOCA_tests.

Van Uffelen, P. et al. (2019), "A review of fuel performance modelling", *Journal of Nuclear Materials*, Vol. 516, pp. 373-412, https://doi.org/10.1016/j.jnucmat.2018.12.037.

Zhang, J. (2018), "IAEA FUMAC Benchmark on the Uncertainty and Sensitivity Analysis for Fuel Rod Code Simulation of the Halden LOCA Test IFA-650.10", TOPFUEL 2018 Conference, Prague, Czechia, www.researchgate.net/publication/334304303_iaea_fumac_fuel_rod_code_benchmark_uncertainty_and_sensitivity_analysis_on_the_halden_loca_tests.

Zhang, J. et al. (2024), "Contribution of IAEA coordinated research projects to light water reactors advanced technology fuel testing and simulation", *Nuclear Engineering and Design*, Vol. 418, 112910, https://doi.org/10.1016/j.nucengdes.2024.112910.

2.5. Material ageing, operational experience, and long-term operation

This section addresses the preservation of datasets related to material ageing, material property databases, and operational experience relevant to long-term operation (LTO) of nuclear power plants. While the focus is primarily on water-cooled reactors (WCRs), the same principles apply to other reactor technologies.

Material ageing research involves systematic testing of metals, concrete and polymer-based materials to understand ageing mechanisms, degradation progression, and the ability of structures, systems and components (SSCs) to maintain their intended safety functions over extended periods. Such testing assesses critical damage initiation and growth, mechanical resistance, leak-tightness (e.g. of passive components or containment structures), and insulation integrity of electrical cables. In some cases, repair processes are also studied.

Testing may be conducted on:

- Simulant materials, subjected to accelerated ageing under prototypical environmental conditions, often including irradiation.
- Harvested materials taken from decommissioned or operating nuclear plants, offering insights into ageing under actual LTO conditions.

To interpret results from harvested materials, detailed historical data are essential. In both cases, macro- and microstructural characterisation (e.g. using SEM or TEM) is typically performed before and after testing.

Datasets collected through these efforts may include:

- manufacturing and fabrication records;
- irradiation parameters (for irradiated materials);
- testing conditions and time-series results;
- pre- and post-test characterisations.

Material databases contain information on the physical and mechanical properties of materials used in commercial nuclear power plants, often documenting changes due to ageing, and may include trend curves for predictive modelling. Operational experience (OPEX) databases record failure types, material properties, root causes and reliability indicators, supporting ageing management and decision making.

These datasets and databases serve to:

- develop and validate models for crack initiation, crack growth and other key degradation mechanisms;
- validate trend curves for material behaviour;
- support evidence-based ageing management practices.

2.5.1. *Ongoing collective data preservation efforts*

This section reviews collective international efforts for data preservation in the LTO domain. While many utilities maintain national databases – often proprietary and tailored to their reactor types and operational environments – only a limited subset of these data is shared internationally. These national resources typically contain ageing-related data on specific reactor designs (e.g. CANDU, PWR, BWR, VVER), components and material types (e.g. specific alloys, concrete mixes, cable insulation systems).

Utilities and operators also develop databases to support plant ageing management programmes, often aligned with IAEA guidance; see IAEA (2022). However, most are not accessible for collaborative safety research, highlighting the value of international initiatives.

The following subsections summarise major international data preservation efforts.

International metal material databases

- IAEA International RPV Database (Kryukov, 2012): Collected surveillance and irradiation embrittlement data from 14 countries. A standardised data framework was established for secure and interchangeable use. The database is no longer actively maintained.
- AGE60+ Project (European Commission, 2015-2016): Focused on data relevance for extending LWR operation beyond 60 years. The project consolidated databases related to MnMoNi steels, VVER RPV materials, thermal ageing of low-alloy steels, IASCC, and concrete ageing. A new embrittlement trend curve (ETC) for VVER-440 was developed. The project concluded after 18 months.
- ENTENTE Project (2020-2024): Funded under Euratom H2020, this initiative sought to design, populate and apply a comprehensive RPV embrittlement database. It incorporated data from earlier Euratom projects and applied integrated computational materials engineering (ICME), AI, and machine learning to improve predictive capabilities. 27 organisations across 12 countries participated.
- IFDA – International Fatigue Database (launched 2023): Led by the USNRC, IFDA consolidates fatigue data from earlier bilateral and multilateral collaborations (e.g. INCEFA+, INCEFA-SCALE). Data entry is standardised using templates and uploaded to MatDB with appropriate quality assurance. JRC oversees data management and training.

These databases support safety upgrades and LTO decisions for current plants and may extend to advanced designs. ENTENTE and IFDA are being positioned for long-term operation, with the NEA Data Bank envisioned as their eventual custodian.

- NEA RPV Integrity Database Initiative: Developed under WGAMA–WGIAGE to compile temperature-dependent data on RPV steels for in-vessel melt retention (IVMR) analysis. Properties relevant for finite element modelling and damage assessment are covered.
- IAEA Research Reactor Databases:
  - Ageing Database: Collects good practices and ageing resolution case studies.
  - Materials Database (IAEA, 2019): Compiles irradiated core structural material properties, to be developed into a relational database.

International concrete material databases

- IAEA Concrete Ageing Database (Seni, 1996): Developed with AECL (now CNL) to compile ageing data for concrete structures. The project established procedures for data collection, validation, and access.
- AGE60+ Concrete Module: Highlighted the need for shared databases on concrete ageing, including waste disposal contexts. Currently inactive.
- ACES Project (2020-2025): Develops a concrete creep and shrinkage database to improve LTO assessment of civil structures (European Commission, 2024b).

International electrical cable databases

- SCAP and CADAK Projects (NEA, 2007; NEA, 2010; NEA, 2018): Focused on ageing and condition monitoring of safety-critical cables. The resulting databases compiled ageing mechanisms, failure modes, and testing protocols. Although valuable, these initiatives were discontinued due to lack of sustained international support.

Data from international experimental programmes

- Halden Reactor Project: Included materials ageing research (e.g. IASCC and creep testing). The legacy database, now preserved at the NEA Data Bank, includes reports and PIE data. Selected data could be migrated into IFPE-like formats to enhance usability.
- SMILE Project: Investigates harvested metal components from Swedish reactors. Long-term data preservation plans are under discussion with the NEA Data Bank.
- FIDES-II INCREASE: A recent project that will address ageing mechanisms. Data preservation frameworks are being developed, leveraging Halden's experience.

International event databases for systems and components

- CODAP (Component Operational Experience, Degradation and Ageing Programme): As a successor to the OPDE and the Stress Corrosion Cracking and Cable Ageing Project (SCAP), projects, CODAP compiles event data on passive component degradation. The database supports inspection programme evaluation, PSA input, and regulatory review. Now in its fifth phase, CODAP benefits from strong regulatory support and regular analysis publications (NEA, 2019; NEA, 2023).

2.5.2. ***Analysis of the current situation***

Establishing and maintaining material and event databases requires significant resources, including data collection, validation, QA, formatting and access management. The proprietary nature of much of the data adds complexity.

Nevertheless, extensive experience has been gained in developing harmonised procedures for data submission and quality assurance. The ENTENTE and IFDA databases are leading efforts at the time of writing, but it remains necessary to:

- ensure these efforts comprehensively address LTO needs for current and advanced reactors;
- evaluate the feasibility of a dedicated RPV database for IVMR;
- investigate why data efforts for concrete and polymers have not continued and assess future needs.

A recurring issue is the absence of long-term preservation and dissemination plans at project inception, creating risks of data loss or obsolescence post-project.

Databases should be designed with user needs in mind, ensuring discoverability and format compatibility. Access rights and proprietary data handling need to be clarified upfront and maintained post-project. Enhanced sharing of utility data would greatly benefit broader safety analyses and license-renewal decisions.

Among NEA experimental projects, only Halden data are currently accessible to all NEA members. SMILE and FIDES-II data are still under access restrictions. Long-term access and export control policies should be discussed.

For event databases like CODAP, experience shows that sustainable operation requires:

1. international critical mass;
2. national commitment to detailed data coding;
3. legal frameworks for data protection;
4. long-term membership engagement.

On harvesting research, EPRI has begun building a database of harvested components. A broader international initiative to catalogue these resources should be considered.

2.5.3. *Recommendations*

- International co-ordination: Organise a high-level forum (NEA, EC, IAEA) to review the status of LTO-related data preservation, identify gaps in metal, concrete and polymer material datasets, and define policies for long-term access and maintenance.
- Long-term planning: All databases (e.g. ENTENTE, IFDA, CODAP) and experimental projects (e.g. SMILE, FIDES-II) should include preservation, access and data protection plans from the outset. Early co-ordination with the NEA Data Bank is recommended.
- WGIAGE role: The NEA Working Group on Integrity and Ageing (WGIAGE) should co-ordinate the prioritisation of collective data efforts to support life extension and deployment of new reactor designs, including a potential RPV database for IVMR.
- Harvested materials library: A collective effort should be considered to establish a centralised library or catalogue of harvested components and associated operational data.
- Halden data use: Consider launching data review campaigns or benchmarks using selected Halden datasets to extract further insights.
- Sustaining CODAP: National regulators should continue to allocate staff and resources to ensure high-quality submissions and the ongoing utility of CODAP.
- Modern technologies: Encourage integration of machine learning (ML) and artificial intelligence (AI) techniques in future database development and data mining activities.

2.5.4. *References*

EC (2024a), *European Database for Multiscale Modelling of Radiation Damage, European Database for Multiscale Modelling of Radiation Damage*, ENTENTE Project, News & Multimedia, H2020, CORDIS, European Commission, https://doi.org/10.3030/900018.

EC (2024b), *Towards Improved Assessment of Safety Performance for Long-Term Operation of Nuclear Civil Engineering Structures*, ACES Project, News & Multimedia, H2020, CORDIS, European Commission, https://doi.org/10.3030/900012.

IAEA (2022), *Ageing Management and Long Term Operation of Nuclear Power Plants: Data Management, Scope Setting, Plant Programmes and Documentation*, Safety Reports Series No. 106, International Atomic Energy Agency, Vienna, www.iaea.org/publications/14693.

IAEA (2019), *Material Properties Database for Irradiated Core Structural Components for Lifetime Management for Long Term Operation of Research Reactors*, IAEA-TECDOC-1871, IAEA, Vienna, www.iaea.org/publications/13529.

Kolluri, M. et al. (2023), "Structural Materials research on parameters influencing the material properties of RPV steels for safe long-term operation of PWR NPPs", *Nuclear Engineering and Design*, Volume 406, 112236, https://doi.org/10.1016/j.nucengdes.2023.112236.

Kryukov, A. et al. (2012), "International database on reactor pressure vessel materials for long-term operation of NPPs", *Nuclear Technology*, Vol. 180, Issue 3, pp. 443-449, https://doi.org/10.13182/NT12-A15355.

Lydell, B., D. Escrig Forano and J. Riznic (2021), "OECD Nuclear Energy Agency CODAP database project on passive component operating experience: An international collaboration in materials research", *Nuclear Engineering and Design*, Volume 380, 111280, https://doi.org/10.1016/j.nucengdes.2021.111280.

Lydell, B. and J. Riznic (2008), "OPDE – The international pipe failure data exchange project", *Nuclear Engineering and Design*, Volume 238, Issue 8, pp. 2115-2123, ISSN 0029-5493, https://doi.org/10.1016/j.nucengdes.2008.01.002.

NEA (2023), "Use of NEA Database Projects Operating Experience Data for Probabilistic Safety Assessment", OECD Publishing, Paris, www.oecd-nea.org/jcms/pl_88697.

NEA (2019), "NEA CODAP Project Topical Report: A Review of the 'Post-1998' Experience with Thermal Fatigue in Heavy Water and Light Water Reactor Piping Components", OECD Publishing, Paris, www.oecd-nea.org/jcms/pl_73832.

NEA (2018), "Cable Ageing in Nuclear Power Plants: Report on the First and Second Terms (2012-2017) of the NEA Cable Ageing Data and Knowledge (CADAK) Project", OECD Publishing, Paris, www.oecd-nea.org/jcms/pl_19866.

NEA (2012), "OECD-NEA Piping Failure Data Exchange Project (OECD-NEA OPDE) Final Report", OECD Publishing, Paris, www.oecd-nea.org/jcms/pl_19232.

NEA (2010), "Technical Basis for Commendable Practices on Ageing Management SCC and Cable Ageing Project (SCAP) Final Report", OECD Publishing, Paris, www.oecd-nea.org/jcms/pl_19010.

NEA (2007), "SCAP: the NEA project on stress corrosion cracking and cable ageing", *NEA News 2007 – No. 25.1*, OECD Publishing, Paris, www.oecd-nea.org/jcms/pl_14294.

Ortner, S. and M. Brumovský (2017), "Age 60+ – Applicability of ageing related data bases and methodologies for ensuring safe operation of LWR beyond 60 years", *Proceedings of the ASME 2017 Pressure Vessels and Piping Conference*, Volume 6A: Materials and Fabrication, Waikoloa, Hawaii, 16-20 July 2017, V06AT06A075, ASME, https://doi.org/10.1115/PVP2017-65283.

Seni, C. and L. Ianko (1996), "An International Database of Nuclear Concrete Structures Ageing", in Dubois, J.E., Gershon, N. (eds), *Industrial Information and Design Issues, Data and Knowledge in a Changing World*, Springer, Berlin, Heidelberg, https://doi.org/10.1007/978-3-642-80286-7_32.

USNRC (2023), "2nd Workshop on International Harvesting Cooperation Summary Report: Workshop held on November 17, 2022 in Stockholm, Sweden", www.nrc.gov/docs/ML2334/ML23347A144.pdf.

2.6. Fire hazard research

Since the early 2000s, the Institut de Radioprotection et de Sûreté Nucléaire (IRSN), now the Autorité de Sûreté Nucléaire et de Radioprotection (ASNR), has led several experimental programmes focused on fire risk in nuclear installations. These initiatives – conducted either independently or under the auspices of the NEA – are carried out in collaboration with various national and international partners. Their overarching objectives are to enhance scientific understanding of fire phenomena in nuclear environments and to support the development and validation of predictive fire simulation tools used in safety assessments.

A prominent example of such efforts is the PRISME series of large-scale experimental projects (acronym for *Propagation d'un incendie pour des scénarios multi-locaux élémentaires*), undertaken within the framework of NEA joint research initiatives. These experiments were conducted at IRSN's GALAXIE experimental platform, notably using the DIVA facility, a confined and mechanically ventilated installation designed for realistic fire tests.

From 2006 to 2022, three PRISME phases were implemented. The projects explored critical aspects of fire dynamics in nuclear settings, including:

- the effects of confinement and mechanical ventilation on fire development;
- smoke and heat propagation from the fire compartment to adjacent rooms;
- electrical cabinet and cable tray fire propagation towards representative targets;
- the impact of spray systems on fire dynamics in confined spaces;
- fire spread along extended cable trays.

Alongside the experimental work, a dedicated analytical group assessed the predictive capabilities of various fire modelling tools, with numerical validation activities performed by participating organisations. The results of these efforts, including benchmark outcomes and analytical insights, are extensively documented (Audouin, 2013).

Since 2023, a follow-up experimental programme – FAIR (Fire Risk Assessment through Innovative Research) – has been launched under NEA auspices. FAIR builds on PRISME by exploring new fire scenarios and research topics to generate complementary datasets (March, 2023). Key findings from the PRISME series were presented at the NEA Joint Nuclear Safety Research Projects Week in January 2023 (Suard, 2023).

2.6.1. *Data generation and preservation*

Experimental projects such as PRISME and FAIR generate a broad spectrum of data, encompassing documentation, raw and processed measurements, and visual records. Data are typically organised into structured directories on dedicated platforms containing:

- Meetings materials: Agendas, minutes and presentation materials in Word, PowerPoint, or PDF formats.
- Reports: Official contractual reports (PDF).
- Quick look reports: Preliminary test summaries (PDF) for timely partner updates.
- Experimental campaigns: Subfolders for each test campaign, containing:
  - test specifications and analysis reports (PDF);
  - photographs (PNG, JPG);
  - video footage (MP4, AVI);
  - raw and validated data files (TXT, XLS/XLSX), often accompanied by graphical representations (PNG) for readability.

Each data file includes a header summarising test information, data types (raw or validated), variable names and units (e.g. time in seconds, temperature in °C, mass in kg).

Project data and documentation are accessible on request through the NEA Data Bank Computer Program Services (www.oecd-nea.org/tools/ie/list). A dedicated project website managed by the NEA provides access for participants. To streamline collaboration, a Gforge system was established during PRISME-1, serving as a shared repository for all project data and communications. PRISME 1-3 datasets are archived on this platform, while FAIR project data are hosted on a dedicated SharePoint site. These services are maintained by ASNR's IT department, which also provides user support.

After each test campaign, data are saved on ASNR's internal servers and then uploaded to the collaborative platforms used in ASNR and NEA projects. For NEA joint projects, a copy is also stored on NEA servers. Collectively, these datasets form the comprehensive experimental database of the project.

2.6.2. *Use of data*

The use of experimental fire data varies by stakeholder:

- Regulatory authorities primarily rely on reports and presentations to support expert review, build institutional knowledge, and stay informed about the latest research developments in fire safety.
- Academic and research institutions, often through confidentiality agreements, utilise the data for model validation and development under collaborative R&D frameworks.
- Operators and TSOs use the data to improve simulation tools and derive practical insights, especially when experimental scenarios are representative of operational conditions.

Benchmark exercises are periodically organised to assess the predictive accuracy of fire simulation tools against the experimental data (Plumecocq, 2023; Audouin, 2011). These exercises are critical for validating numerical models and serve as a practical check on the completeness and quality assurance of the datasets used.

2.6.3. *Principles of database implementation*

Beyond providing project data to international partners, ASNR has undertaken significant internal efforts to preserve and capitalise on its fire research data. In line with regulatory and ethical requirements – including the French national charter for research ethics (26 January 2015) – ASNR ensures that all experimental protocols are documented and all raw data preserved to guarantee full traceability and reproducibility.

To this end, ASNR has developed a dedicated internal experimental database, designed with a uniform structure to facilitate long-term access, comparability across test campaigns, and support for advanced analysis. The database contains:

- raw experimental results;
- validated datasets processed using standardised and documented signal treatment procedures;
- derived quantities, such as heat release rates.

The database is operated using tools developed in the IDL (Interface Description Language) environment. Validated datasets are used internally for physical analysis and model validation and are uploaded to project websites (Gforge or SharePoint) as part of NEA collaborations.

2.6.4. *Recommendations*

As in other areas of nuclear safety research, it is strongly recommended that joint fire safety projects:

- Formalise data preservation and dissemination plans early in the project lifecycle, with clear agreements on long-term access and distribution.
- Define the content of data packages to be transferred to the NEA Data Bank by the end of the project.
- Conduct benchmarking exercises to validate data quality and completeness.

These recommendations are particularly relevant for the FAIR project.

ASNR has made notable progress in fire data preservation, including the development of a centralised database incorporating PRISME and FAIR data. It would be valuable to explore – possibly within the FAIR project – the potential for broader international access to this database. Such discussions should also address the sustainability of access, maintenance and potential co-funding mechanisms to ensure the long-term availability and utility of this important research resource.

2.6.5. *References*

Audouin, L. et al. (2011), "Quantifying differences between computational results and measurements in the case of a large-scale well-confined fire scenario", *Nuclear Engineering and Design*, Vol. 241, Issue 1, pp. 18-31, https://doi.org/10.1016/j.nucengdes.2010.10.027.

Audouin, L. and J.L. Torero (2013), "Special Issue on PRISME – Fire Safety in Nuclear Facilities", *Fire Safety Journal*, Vol. 62, Part B, pp. 79-220, https://doi.org/10.1016/j.firesaf.2013.11.002.

March, P. et al. (2023), "*FAIR: A New OECD Fire Risk Research Project Under Development as a Follow-Up to PRISME 3*", 17th International Post-Conference Seminar on Fire Safety in Nuclear Power Plants and Installations, www.researchgate.net/publication/372406827.

NEA (2024), NEA Nuclear Safety Research Joint Projects, www.oecd-nea.org/jcms/pl_72332.

Plumecocq, W. et al. (2023), "Common OECD/NEA FIRE and PRISME Cable Benchmark Exercise", *Fire Safety Journal*, Vol. 139, https://doi.org/10.1016/j.firesaf.2023.103815.

Suard, S. et al. (2024), "Fire propagation in a mechanically ventilated multi-compartment facility: OECD/NEA project PRISME 3", *Fire Safety Journal*, Vol. 143, https://doi.org/10.1016/j.firesaf.2023.104062.

Suard, S. (2023), "Session 3: Joint Projects for Safety in Operation, Learnings and Perspectives", presentation at the NEA Nuclear Safety Research Joint Projects Week, A Success Story and Opportunities for Future Developments, 9-13 January 2023, pp. 26-44, www.oecd-nea.org/upload/docs/application/pdf/2023-01/2023_nea_nuclear_safety_research_joint_projects_week_day3.pdf

Suard, S. et al. (2019), "Fire development in multi-compartment facilities: PRISME 2 project", *Fire and Materials*, Vol. 43, Issue 5, pp. 433-435, https://doi.org/10.1002/fam.2754.

Chapter 3. **Conclusion**

3.1. **Conclusions and recommendations**

Following a comprehensive review of data preservation efforts across key nuclear safety research areas, SEGPD members offer a series of recommendations aimed at strengthening data preservation and accessibility practices throughout the nuclear safety community.

Central role of the NEA Data Bank: SEGPD members reaffirm the critical role of the NEA Data Bank as a central, secure and durable repository for high-value nuclear safety data. They strongly recommend systematic collection at the NEA Data Bank of datasets generated under NEA-led safety research activities, particularly those from joint safety research projects and datasets used in data assessment exercises such as International Standard Problems (ISPs), benchmarks and quality assurance reviews.

While acknowledging the existence of proprietary or export control constraints, members emphasise that data collection should remain the default objective. Distribution agreements and appropriate access control mechanisms should be established to address ownership or confidentiality concerns. The NEA should also identify and prioritise missing datasets – those not yet collected – through collaboration with CSNI working groups, and work to encourage their submission for preservation and controlled dissemination.

Availability of data assessment results: Reports and results from data assessment activities – such as benchmarks and ISPs – should be made available to data users as part of the NEA Data Bank's data packages. These materials provide valuable insights into data quality and enhance transparency.

Raising awareness and expanding scope: The NEA should continue efforts to raise awareness among its member countries regarding the capabilities of the NEA Data Bank, particularly its utility in hosting key datasets developed either under NEA auspices or in external projects with relevance to nuclear safety. Notably, the NEA has successfully archived data from experiments conducted outside NEA projects and is increasingly seen by members as a long-term repository for externally developed databases, such as those from the European ENTENTE project and the International Fatigue Database (IFDA). With support from the CSNI and NSC working groups, the NEA should identify and prioritise such external datasets for integration, evaluating associated resource needs and considering the establishment of dedicated joint data preservation projects where appropriate.

Data distribution from joint projects: In the context of joint projects, the SEGPD recommends increased awareness-building among project members regarding the options for controlled data distribution after the end of the agreed non-disclosure period. This could support more systematic collection and long-term archiving of project data.

Data quality, completeness, and accessibility: In the collection, formatting and archiving of data, priority should be given to:

- Ensuring completeness and quality assurance.
- Using durable and standardised formats (numerical, text, audio, image, video).
- Implementing an intuitive and open-source search structure, especially for extended datasets (e.g. Halden legacy data).

Advanced analytics tools may be useful, but completeness and clarity of core data and metadata must come first, as application-specific needs can vary widely.

To ensure high data quality, the SEGPD highlights the need for datasets to include:

- pre- and post-test material characterisation;
- raw and processed data;
- analysis scripts and supporting documentation;
- information on accuracy and uncertainty;
- full descriptions of the test facilities and procedures.

Legacy data challenges: The SEGPD recognises that some legacy datasets may not meet current standards of completeness or quality assurance. Nonetheless, they retain value – especially if accompanied by clear indications of their limitations. Even basic screening criteria (e.g. existence of raw and elaborated data, availability of metadata) would help users assess usability for modern applications, including for new reactor designs like SMRs.

User feedback collection: To strengthen future data use, the NEA Data Bank should collect user feedback more systematically. This could include distributing a brief questionnaire with data packages, focusing on the completeness and quality of the data. Such feedback should be synthesised and used to inform future users and guide improvements.

Guidance for future projects: Looking ahead, the SEGPD recommends the development of a general guidance document defining the expected content of data packages. While each technical domain has specific needs, all members stress the importance of systematically including uncertainty information and boundary condition data. Though difficult to recover retrospectively, these elements should be prioritised in current and future projects.

Data preservation from project outset: Data preservation strategies should be addressed from the beginning of joint experimental projects. This includes agreeing on the content and format of final data packages to be delivered to the NEA Data Bank. Project Management Boards should be made aware of their responsibility to formally approve the final data package, with proposals prepared by the Programme Review Group, including relevant benchmarking outcomes.

Balance of functionality and preservation needs: Regarding the design and features of data preservation tools, the SEGPD reiterates that completeness, quality assurance, durability and searchability must take precedence. In certain technical areas (e.g. thermal-hydraulics, fuel performance), ongoing efforts to develop standardised approaches to data formatting and archiving – such as the THEMPO activity (CSNI/WGAMA) and a parallel

initiative for RIA data (CSNI/WGFS) – should be continued and supported. These initiatives are expected to provide valuable guidance for future database development, subject to availability of appropriate frameworks and resources (e.g. joint projects).

Importance of benchmark activities: The SEGPD strongly recommends integrating benchmark exercises into all NEA joint projects. These exercises not only validate simulation tools but also help verify the completeness and quality of the supporting datasets. As benchmarking requires dedicated co-ordination and documentation, the SEGPD recommends that joint project proposals include appropriate resources for this purpose and that operating agents factor in data quality outcomes from benchmarking activities when finalising data packages.

3.2. Proposed actions to implement recommendations

3.2.1. *NEA Joint Projects*

- Legacy projects: NEA Joint Project administrators, with support from the NEA Data Bank, should continue efforts to collect missing data packages from past projects, define appropriate distribution rules, and engage with data owners (e.g. current efforts for the Halden legacy database).
- Ongoing and future projects: NEA project administrators should define data preservation and distribution schemes early in the project lifecycle, ensuring consensus among participants and data owners.
- Data assessment activities: The NEA Secretariat should proactively plan for data delivery and define distribution agreements with data owners during and after data assessment activities.

3.2.2. *NEA Data Bank*

- Promotion and awareness: The NEA Secretariat should systematically present the NEA Data Bank's capabilities and data access schemes at working group meetings and joint project events. This includes promoting the NEA Data Bank's role in collecting benchmark results and preserving key datasets.
- Technical group co-ordination: CSNI working groups should regularly inform the NEA Secretariat of datasets at risk, emerging data preservation initiatives, or databases under development externally that may benefit from long-term archiving at the NEA.

3.2.3. *Legacy databases*

- Screening for completeness and QA: The NEA should conduct a review of existing datasets and databases using simple screening criteria (e.g. availability of raw and elaborated data, supporting documentation, facility information). Based on this, the NEA should assess the value and feasibility of further enhancement.

- Quality documentation: For high-interest datasets, the NEA should co-ordinate detailed reviews of completeness and quality assurance, and document findings in the distributed data packages.
- User feedback: The NEA should collect and analyse user feedback on legacy datasets to inform users about data strengths and gaps. Feedback can be collected through distribution channels or organised review initiatives.

3.2.4. *Future experimental programmes*

- Guidance on data package content: The NEA Secretariat should develop high-level guidance on essential elements of data packages in collaboration with data producers and users.
- Data policy development: A clear, unified data policy should be established across NEA activities, covering data collection, management, access rights, and preservation.
- Early planning: Anticipated discussions on data packages should be included at project inception, including during ISP and benchmark planning phases.

3.2.5. *Code benchmark activities*

- Integration in project proposals: The NEA Secretariat should ensure that operating agents allocate appropriate resources for co-ordinating benchmarking activities as part of project proposals.
- Using benchmark outcomes: Benchmark results should directly inform the design and content of the final data packages, ensuring they reflect data usability, quality and completeness.

Appendix A. **Questionnaire for gathering information on the current status of key datasets in nuclear safety disciplines**

Introduction

At its 68th meeting in December 2020, the CSNI approved the establishment of the Senior Expert Group on the Preservation of Key Experimental Datasets (SEGPD). The SEGPD's objective is to provide the CSNI – with input from the NSC and MBDAV – with strategic recommendations on how to preserve key datasets that form a critical part of the knowledge base underpinning nuclear safety.

The first SEGPD meeting took place on 17-18 May 2021. During this meeting, senior experts shared their experiences with various experimental databases and began defining the desired outcomes of the group. A key point raised during discussions was the need to gather structured information from technical experts across nuclear safety disciplines to identify:

- vulnerable datasets at risk of being lost;
- datasets critical to safety code validation or licensing bases that are inadequately preserved; and
- examples of well-managed and effectively preserved datasets.

Questionnaire

To that end, we would greatly appreciate your insights and perspectives on the following questions:

1. Existing Datasets and Experimental Projects

   In your technical area, what types of experimental datasets or key projects currently exist?

2. Preservation Status

   How are these key datasets currently being preserved?

   Which datasets do you consider vulnerable to loss?

   Which datasets do you consider well preserved, and what makes them so?

3. Benchmarking Activities

   What major benchmark exercises have been conducted in your area?

4. Validation Bases for Safety Codes

   What reference computer codes are used to support safety cases in your field?

   What major datasets form the basis for validating these codes?

5. Database Use and Experience

   What has been your experience using relevant databases?

   What limitations have you encountered?

   In your view, what defines a “good” database, and what are the key challenges in database use and maintenance?

Appendix B. Summary of the SEGPD Information-Gathering Workshop, 18-19 November 2021

On 18-19 November 2021, experts from across a range of nuclear safety disciplines convened virtually to discuss the current state of data preservation in their respective fields. The workshop was organised by the SEGPD as part of its foundational activities to shape the group's mandate and forward work. It followed the issuance of a pre-workshop questionnaire that gathered feedback on challenges and good practices in data preservation.

The workshop was structured into discussion blocks focusing on:

- database design;
- database population and use; and
- challenges and lessons learnt from both existing and missing databases.

Each block included introductory remarks, invited presentations and discussions informed by responses to the questionnaire (Appendix A). Key cross-cutting themes emerged, demonstrating the interconnected nature of data preservation challenges and solutions.

1. Database design

Insights from questionnaire

- Effective databases must not only store comprehensive data, but also include powerful search capabilities and quality assessment tools.
- Advanced and intuitive search functionality is essential, as users often rely on basic keywords.
- Early definition of data formatting requirements improves database usability and coherence.

Invited presentation

Mr Scott Holcombe (IFE) shared insights from Halden's experience digitising 60+ years of experimental data using the Data Lake platform. He emphasised the importance of ongoing user-developer engagement and advisory input throughout the process.

Key points

- Clear data structures and early identification of end-users improve efficiency.
- Continuous interaction between developers and end-users is critical.
- Database architecture must be adaptable over time to evolving needs.
- Expandability should be user-driven to reduce IT overhead.

Discussion themes

- File formats (e.g. ASCII, .txt), raw vs. processed data, and traceability of processing steps.
- Discipline-specific needs must be addressed with tailored solutions.

2. Database population

Insights from questionnaire

Data preservation must go beyond test results to include:

- (1) Input/initial specimen characterisation.
- (2) Experiment rationale/context.
- (3) Post-test analysis and interpretation.
- Metadata on test conditions is crucial for assessing dataset applicability and quality.
- Traceability to original documentation is essential.
- Capturing parameter interdependencies is critical (e.g. as in thermodynamic databases).
- Ideally, everything should be preserved – from blueprints and material specs to boundary conditions and raw test outputs.

Invited presentation

Mr Ian Hill (OECD/NEA) described the IRPhEP Handbook, identified in the questionnaire as a model resource. He highlighted the value of ongoing updates and comprehensive metadata to ensure long-term usability.

Key points

- Characterising uncertainties during database population is essential.
- Experimental setup and methodology are as important as the results themselves.

Discussion themes

- Balancing completeness with usability.
- Single-purpose vs. multi-purpose databases.
- Reusability of legacy data and related limitations.

3. Database use

Insights from questionnaire

Examples of well-regarded databases include:

- ICSBEP/DICE and IRPhEP.
- SFCOMPO.
- PHEBUS-FP and THAI-based projects.
- STEX and MCCI (ANL).
- IFPE/DATIF and Halden fuel performance data.

Keyword-based searches are often insufficient; machine learning could improve user access and relevance of results.

Invited presentation

Ms Alice Dufresne (OECD/NEA) shared insights from the NEA Data Bank. She discussed content screening procedures, challenges in managing access rights – especially for legacy data – and emphasised the importance of clear usage guidelines for end-users.

Key points

- Tutorials and user guides support accessibility.
- Datasets with active management are in higher demand.

4. Challenges in data preservation

Insights from questionnaire

- Lack of accuracy statements for instruments.
- Missing or incomplete facility descriptions, often due to staff retirement.
- Experimental uncertainties and error estimates are frequently omitted.
- Long-term durability of storage media and formats must be ensured.
- Many datasets are incomplete – especially lacking raw data or test specifications.

Discussion themes

- Missing boundary/initial conditions, measurement accuracy, and uncertainty data hinder validation.
- Important clarifications are often shared only orally during project meetings.
- Quality labels or indicators could help identify reliable datasets.

- Examples like the IRSN fire database demonstrate the value of capturing full test inputs and metadata.
- The need for standardised guidelines on data formats and preservation approaches.

Invited presentation

Dr Alessandro Petruzzi presented the CAPS proposal under CSNI/WGAMA: *Harmonization of Methodologies for Thermal-Hydraulics Experimental Meta-Data Preservation, Collection and Qualification (THEMPo)*. He proposed transitioning from document-based to relational, data-centric databases with standardised metadata collection practices.

Key points

- Expert collaboration frameworks can support data population without excessive dedicated resources.
- A well-structured database increases efficiency and prevents repeated data retrieval efforts.
- Lack of preserved data compromises the validity of safety analysis.

5. Cross-cutting points

- Data preservation demands significant resources and long-term commitment; high-level support and prioritisation are essential.
- The cost of inaction – loss of data and reduced benefit to future generations – should be clearly articulated.
- Preservation requires not just design and population, but also continuous digitisation, curation, and maintenance.
- Capturing data iteratively during the project life cycle ("as-you-go") avoids the risks of a final "data dump."
- Ownership, access rights, and data use policies must be defined early in a project's lifecycle.
- Parallel activities like modelling and code validation help define required data granularity and promote timely documentation.
- Preservation supports knowledge transfer and enables modern approaches such as data mining and AI-driven analysis.
- Future needs (e.g. multi-physics simulations) will require integrated, multidisciplinary datasets from common experiments.
- Despite past commitments, current efforts are insufficient; The SEGPD should work to reinforce, integrate, and build upon existing tools and initiatives.

NEA PUBLICATIONS AND INFORMATION

The full catalogue of publications is available online at www.oecd-nea.org/pub.

In addition to basic information on the Agency and its work programme, the NEA website offers free downloads of hundreds of technical and policy-oriented reports. The professional journal of the Agency, *NEA News* – featuring articles on the latest nuclear energy issues – is available online at www.oecd-nea.org/nea-news.

An NEA monthly electronic bulletin is also distributed free of charge to subscribers, providing updates of new results, events and publications. Sign up at www.oecd-nea.org/bulletin.

Visit us on LinkedIn at www.linkedin.com/company/oecd-nuclear-energy-agency or follow us on X (formerly known as Twitter) @OECD_NEA.

OECD/NEA PUBLISHING, 2 rue André-Pascal, 75775 PARIS CEDEX 1

Preservation of Key Experimental Datasets for Nuclear Safety

High-quality experimental data are essential for the efficient management of knowledge and maintaining the scientific basis for the next generation. This report provides an overview of data preservation in key areas of nuclear safety, such as neutronics, thermal hydraulics, severe accidents, fuel performance, ageing management, and fire hazards. It also identifies the most critical data for future nuclear safety assessments. Based on the work of the NEA Senior Expert Group on the Preservation of Key Experimental Datasets (SEGPD), the report highlights vulnerabilities, best practices, and opportunities for ensuring long-term access to experimental knowledge. It provides practical recommendations for improving the completeness, quality and usability of data, and emphasises the role of the NEA Data Bank as a secure repository for preserving the data. This publication is designed to be a key reference for technical experts, research organisations, regulators, and anyone interested in maintaining a high-quality database for nuclear safety.

Nuclear Energy Agency (NEA)
46, quai Alphonse Le Gallo
92100 Boulogne-Billancourt, France
nea@oecd-nea.org www.oecd-nea.org

NEA No. 7744