Digital Preservation at
International Nuclear Information System (INIS)

Dobrica Savić & Germain St-Pierre

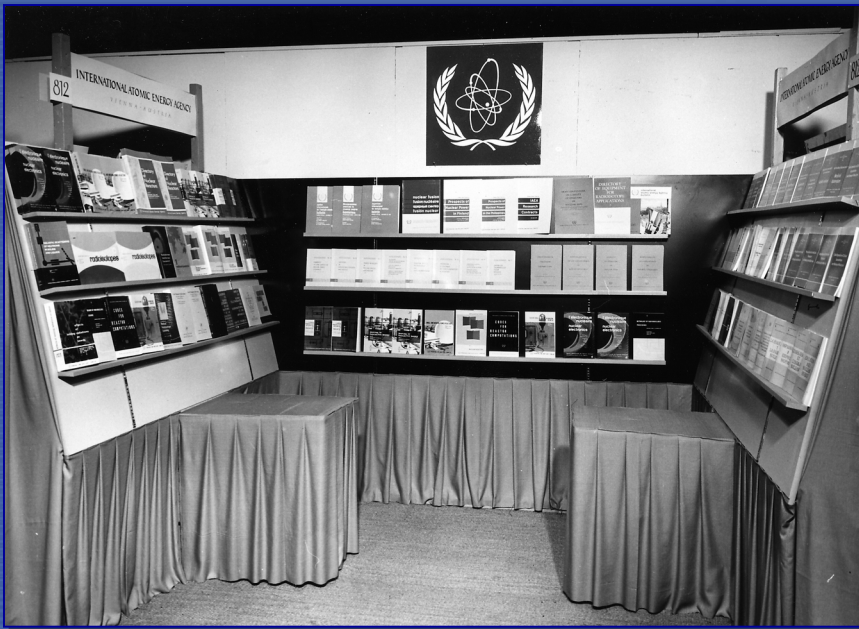Nuclear Information Section, IAEA

Vienna Austria

IAEA

International Atomic Energy Agency

Contents

- About INIS
- INIS Collection
- NCL Processing and Distribution
- INIS NCL Collection on Microfiche (1970-1996)
- INIS Imaging System (INISIS) (1997-2003)
- First Steps in Digitization: Imaging Pilot Project (Summer 1995)
- INIS Imaging System 2000 (INISIS2K) (2003-2009)
- Current Technical Infrastructure (2010 to present)
- Microfiche Digitization Project
- Other Digitization Activities
- Conclusions

IAEA

International Nuclear Information System (INIS)

- One of the world's largest collections of published information on the peaceful uses of nuclear science and technology
- 43 years of international cooperation
- Operated by the IAEA in collaboration with 128 countries and 24 international organizations

IAEA

International Nuclear Information System (INIS)

- *Free online access* to a unique collection of non-conventional (grey) literature *www.iaea.org/inis*
- Sophisticated Google-based search engine
- Unique multilingual nuclear thesaurus

INIS Collection

- Over 3,5 million bibliographic records
- More than one million NCL records
- Around 320,000 full-texts in PDF
- Reports, journal articles, books, legislative materials, multimedia, patents, software, standards, thesis/dissertations

INIS Collection

- 70% of all NCL (over 700,000 items) available in electronic or microfiche format

- Around 300,000 NCL items available from other sources

- Digitization of the microfiche collection is ongoing

- The INIS Collection counts over **480,000** full-texts in PDF (status 2013-10-24)
  - Around **320,000** full-texts are *publicly* available
  - 160,000 full-texts require *clearance procedure*
-

NCL Processing and Distribution

	Production System	NCL Processing	NCL Distribution
1970-1996	Microfiche-based (B&H, TDC Documate, Anacomp)	Conversion from paper to microfiche (Note: NCL from U.S.A. and Japan received on microfiche)	Diazo duplicates
1997-2003	INISIS Image-based (JouveScan)	Conversion from paper to TIFF	NCL CD-ROM with INISIR viewer
2003-2009	INISIS2K (InputAccel) IDPS (Livelink)	Conversion from paper to PDF + OCR Ingestion of PDF files	NCL CD-ROM Online access (subscribers only) E-mail / FTP
2010 to present	NCL Collection System (PixEdit + ABBYY FineReader)	Conversion from paper to PDF + OCR Ingestion of PDF files	NCL DVD Free online access E-mail / FTP

INIS NCL Collection on Microfiche (1970-1996)

- 312,000 NCL reports on microfiche
- 500,000 bibliographic records
- Over 1 million physical microfiches (≈ 17 million pages)
- Diazo duplicates sent to INIS customers

First Steps in Digitization: Imaging Pilot Project (Summer 1995)

- There is always a beginning to a success story
- B/W scanning, document feeder, but no duplex scanner available
- Output: the first CD burned on the IAEA premises

INIS Imaging System (INISIS) (1997-2003)

- Image-based production system by Jouve Systems
- B/W scanning, image enhancement, link to bibliographic metadata using bar codes, validation and CD image creation
- Format supported: single-page TIFF Group 4
- 2 b/w Fujitsu production scanners, 5 workstations, 2 servers

IAEA

INIS Imaging System 2000 (INISIS2K) (2003-2009)

- 32-bit modular Capture System (EMC InputAccel)
- B/W, Greyscale and Color Scanning, OCR, output to PDF
- Integration with Livelink-based INIS Data Processing System (IDPS)
- 2 B/W Fujitsu production scanners, 1 Kodak i260 color scanner, dedicated workstations for image enhancement and OCR

IAEA

Current Technical Infrastructure (2010 to present)

- Robust NCL Collection System to support its continuous growth
- Flexible tools to support different projects:
    - PixEdit 7 (scanning, image enhancement)
    - ABBYY FineReader Corporate Edition 11 (multilingual OCR)
- 2 scanners (Fujitsu fi-5750c with VRS, Kodak i1440) that support color, greyscale and b/w scanning
- 2 Sunrise microfiche scanners
- 4 workstations with quad-core processors

IAEA

Microfiche Digitization Project

- The microfiche digitization started in 2002 as support to document delivery service

- Scanning part had to be outsourced to 3 different companies

- Image enhancement, OCR and output to PDF done in-house

- 2 Sunrise microfiche scanners (300 true dpi)

- Special scanning utility (Scanflow)

- Image enhancement with PixEdit 7

- OCR with ABBYY FineReader 11

IAEA

Other Digitization Activities

- Support digitization initiatives from Member States

    France, Mexico, Sweden, USA, Serbia, etc.

- Digitization of old IAEA publications

    Proceedings Series, Safety Series, IAEA Bulletin, IAEA Board of Governors, International Nuclear Data Committee (INDC), etc.

Conclusions

- Digitization projects require:
  - serious planning
  - substantial funds
  - qualified staff (internal or external)
  - awareness of standards
  - well defined purpose
  - Maintenance of high quality digitization process is a must
  - Document preparation, selection of proper scanning techniques, type of equipment, and adherence to current standards, determine digitization success or failure
- Digitization should not be a goal in itself - its ultimate use and usefulness must always be taken into account
- Meaningful and searchable metadata should accompany any digitized collection, making online search and retrieval efficient
- Long term preservation needs to be considered in order to ensure future sustainability and availability of the digitized collection

Nothing endures but change!

Heraclitus (2500 years ago)

Thank you !