

# Digital Preservation at International Nuclear Information System (INIS)

Dobrica Savić & Germain St-Pierre  
Nuclear Information Section, IAEA  
Vienna Austria



**IAEA**

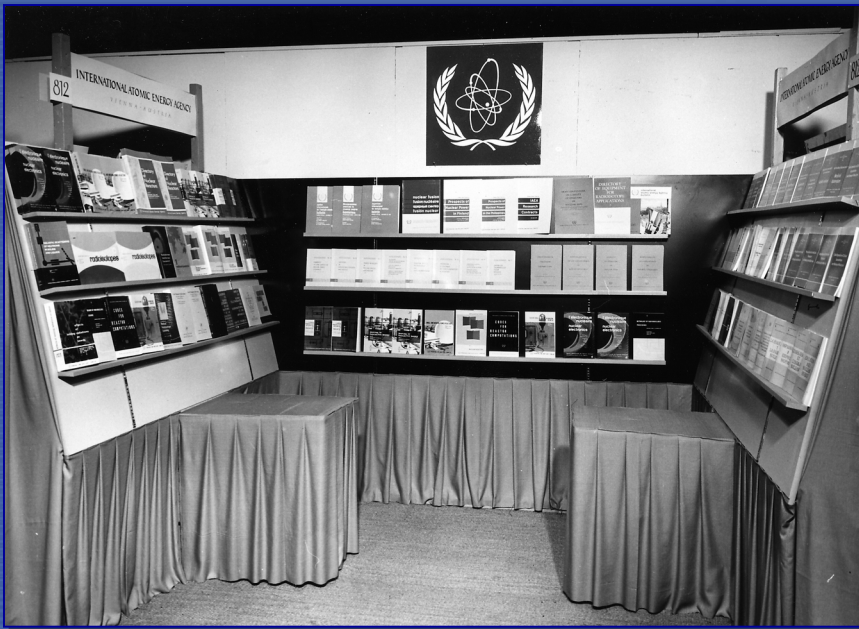
International Atomic Energy Agency

# 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

# 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



# International Nuclear Information System (INIS)

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

The screenshot displays the IAEA INIS Collection Search interface. At the top, there is a navigation bar with links for FAQ, Help, Contact Us, Register, Sign In, and a language dropdown set to English. Below this, the IAEA and INIS logos are visible. The main search area includes tabs for Standard Search, My Selection, and Search History. A search box contains the text 'rn:...45000000', with buttons for Search, Save Query, Subscribe, and Email. A checkbox for 'Limit to results with full text' is checked. The results section shows 'Results 1 - 10 of about 319473. Search took: 6.06 seconds.' A sidebar on the left allows refining the search by Country/Org., listing various entities like the United States, IAEA, France, Russian Federation, JINR, and Poland. The main results area displays a document titled 'Excursions at the places of extraction and processing of inorganic non-metallic materials in Slovakia. The second part', with a file size of 95 MB and a URL: [http://www.iaea.org/inis/collection/NCLCollectionStore/\\_Public/43/095/43095](http://www.iaea.org/inis/collection/NCLCollectionStore/_Public/43/095/43095). The author is listed as 'Jesenak, K. (Department of Inorganic Chemistry, Faculty of Natural Sciences, Comenius Univ., 84215 Bratislava (Slovakia))'.

# 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



The screenshot shows the IAEA INIS Collection Search interface. At the top, there is a blue header with the IAEA logo and the text "IAEA | INIS". To the right of the header, there are links for "ST-PIERRE, Germain", "FAQ", and "Help". Below the header, there is a navigation bar with links for "INIS Home", "Thesaurus", "My INIS", and "Browse". The main content area is titled "INIS Collection Search" and features an "Advanced Search" tab. Below this, there is a search form with the text "Search INIS Collection for documents that...". The form includes two dropdown menus: "Include:" and "But do not include:". The "Include:" dropdown is currently open, showing a list of document types: "Record Type", "Audiovisual", "Book", "Journal article", "Legislative Material", "Miscellaneous", "Multimedia", "Patent", "Report", "Software", "Standard", and "Thesis/Dissertation". The "But do not include:" dropdown is set to "Anywhere". There are also "Add Another" links next to each dropdown menu. A "Clear All" button is located at the top right of the search form. A "Search" button is located at the bottom right of the search form. Below the search form, there is a link for "on Search Highlights".

## 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 (Livelihood)	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 ( $\approx$  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



## 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



# 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

NCL Collection System

Home Queries Statistics About

NCL Documents

Page 17 of 25279, showing 19 records out of 480288 total, starting on record 305, ending on 323 Filter: Off

RN	FLAG	FILE NAME	FILE SIZE	ENTRY DATE	UPDATE DATE
44096784	P	44096784.pdf	407244	2013-10-24 01:01:00	2013-10-02 09:51:00
44096785	P	44096785.pdf	1134409	2013-10-24 01:01:00	2013-10-02 09:51:00
44096786	P	44096786.pdf	132703	2013-10-24 01:01:00	2013-10-02 09:51:00
44096787	P	44096787.pdf	299769	2013-10-24 01:01:00	2013-10-02 09:51:00
44096788	P	44096788.pdf	222334	2013-10-24 01:01:00	2013-10-22 10:50:00
44096789	P	44096789.pdf	399526	2013-10-24 01:01:00	2013-10-22 10:52:00
44096790	P	44096790.pdf	628996	2013-10-24 01:01:00	2013-10-02 09:51:00
44096791	P	44096791.pdf	143643	2013-10-24 01:01:00	2013-10-02 09:51:00
44096792	P	44096792.pdf	1328353	2013-10-24 01:01:00	2013-10-02 09:51:00
44096793	P	44096793.pdf	228360	2013-10-24 01:01:00	2013-10-02 09:51:00
44096794	P	44096794.pdf	376280	2013-10-24 01:01:00	2013-10-22 11:23:00
44096795	P	44096795.pdf	369737	2013-10-24 01:01:00	2013-10-22 11:24:00
44096796	P	44096796.pdf	163027	2013-10-24 01:01:00	2013-10-22 11:26:00
44096797	P	44096797.pdf	267140	2013-10-24 01:01:00	2013-10-22 11:02:00
44096798	P	44096798.pdf	30401211	2013-10-24 01:01:00	2013-08-07 11:25:00
44096799	P	44096799.pdf	12216805	2013-10-24 01:01:00	2013-08-07 11:48:00
44096800	P	44096800.pdf	20999679	2013-10-24 01:01:00	2013-08-12 11:25:00
44096801	P	44096801.pdf	14066140	2013-10-24 01:01:00	2013-08-14 09:09:00
44096802	P	44096802.pdf	6835790	2013-10-24 01:01:00	2013-08-14 09:45:00

INIS Collection Search  
How To...

inscl.iaea.org © 2010-2012 created by NIS Systems Development and Support Group



## 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





## 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 !