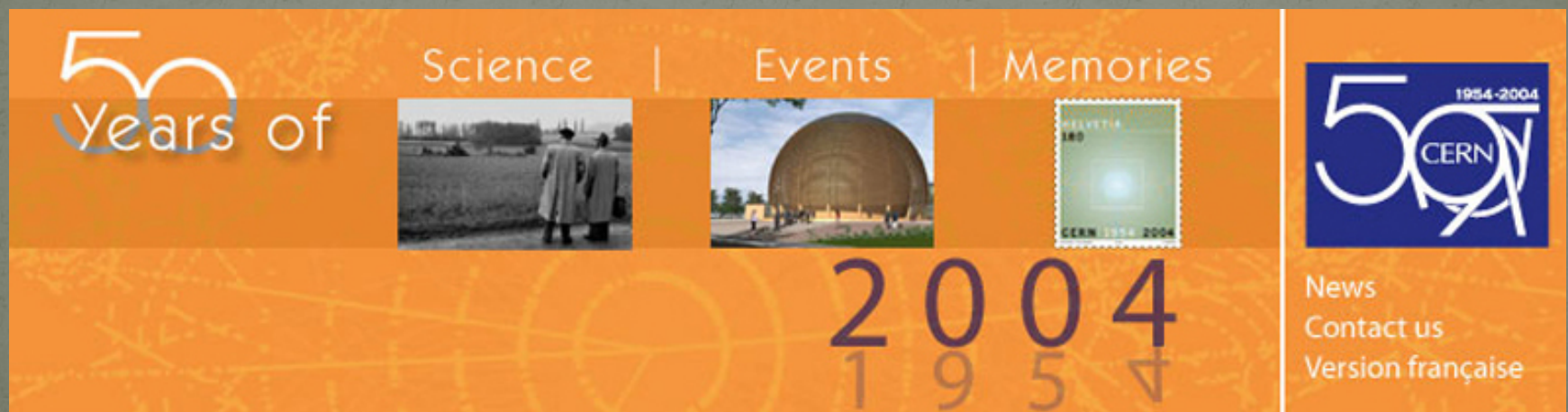


Widening access to historical documents

Digitisation of CERN's archival records

Some issues for the CERN Archive

- How to give the best service to our users?
- Increasing expectations, insufficient resources
- Digitisation – what, why and how



50 Years of Science | Events | Memories

2004
1954

News
Contact us
Version française

The banner features a large '50 Years of' logo on the left, followed by three columns: 'Science' with a photo of two people in a field, 'Events' with a photo of a large dome structure, and 'Memories' with a postage stamp. The year '2004' is displayed in large numbers, with '1954' below it. On the right, there is a CERN 50th anniversary logo and contact information.

Overview



- CERN
- The CERN Archive
- Digitisation and other priorities
- Some of our digitisation projects
 - Theoretical Physics Department reports
 - CERN Council documents
 - The Pauli Archive

CERN - The European Organization for Nuclear Research



- Created in 1954
- The world's largest particle physics laboratory
- Run by 20 European Member States
- Facilities include the 27km circular accelerator

CERN

- Employs around 2,250 people
- 10,000 visiting scientists use CERN's facilities each year
- Education and outreach
- Official languages French and English



CERN



“The Organization shall have no concern with work for military requirements and the results of its experimental and theoretical work shall be published or otherwise made generally available.”

(Convention for the establishment of a European Organization for Nuclear Research, 1953)

The CERN Archive

- 1 linear km of documents
- Recording all aspects of CERN's history
- Archives section of the Scientific Information Service (with Library and OA sections)
- 2 staff members (1.6 fte)



The CERN Archive = The Memory of CERN



- Main functions
 - To select and obtain records of long-term historical interest
 - To keep the records safe (protect their long-term evidential integrity)
 - To make the records available for research



A. Hollier 14/9/2012
(anita.hollier@cern.ch)

Digitisation and other priorities



- Digitisation contributes to the protection of documents and to making them available
- But is it more important to digitise records than to treat the backlog of uncatalogued material?
- Is it more important to digitise records than to be more proactive in obtaining historical material?

Changing expectations

- Increasing demand for online access
- Advocacy issues
 - Securing continued support for the Archive
 - Profile raising
 - Tangible products



Tim Berners-Lee (CERN) - inventor of the WWW

Three of our digitisation projects – mixed motivation



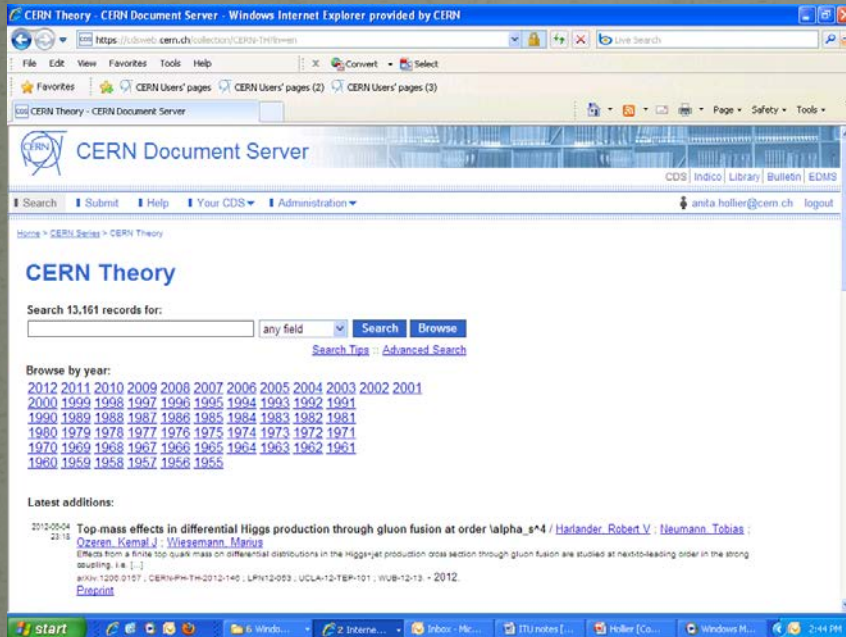
- Theoretical Physics Department reports
 - Widening access, support for OA
- CERN Council documents
 - Widening access, support for greater openness by CERN's governing body
- The Pauli Archive
 - Widening access, conservation, collaboration

1. Theoretical Physics Department reports

- Objectives and scope
 - Working with the Library to complete their online full-text collections (the 'green road' to Open Access)
 - Access to one of the oldest departmental series
 - Around 12,000 records
- Workflow
 - Identify missing documents (inconsistent numbering systems)
 - Check for electronic versions held elsewhere
 - Find and scan paper copies (over 2,500 documents; work carried out in-house)
 - Update database (new cataloguing and standardisation of old cataloguing) with open access to full texts
 - QA (scans and originals)

Results

- Results
 - All full texts scanned & OCR'd
 - Available in Library database
 - Browsable by year



2. CERN Council documents

- In September 2008 CERN's governing body, the CERN Council: "approved the President's proposal to have all paper copies of past Council documents scanned with a view to making them available electronically".

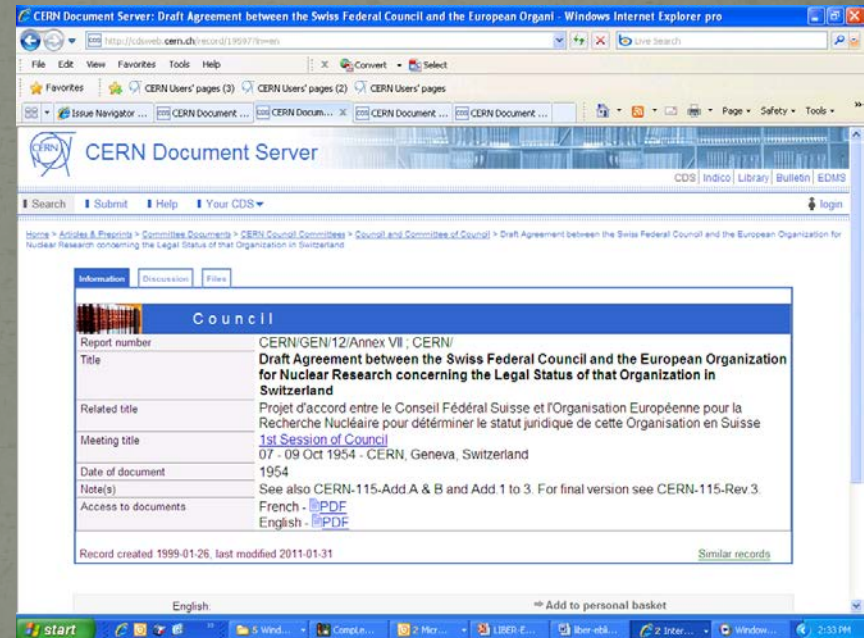


Objectives and scope

- Approximately 21,000 records held in the Archive
- Improve access to the documents by providing online full texts in the Library database
- Opportunity to streamline and reorganise the database collections and standardise catalogue records
- Need to respect closure periods (30 or 5 years)
- Extra resources: one person working 50% for one year plus scanning/OCR expenses (c. 82 KCHF; 0.43 CHF per page)
- Project took around 30 months to complete (approx. one person full time)

Workflow

- Identify all documents from each meeting
- Scan and OCR (initially in-house, later outsourced; controlled using barcodes)
- Edit existing catalogue records or create new ones, combining French and English versions
- Apply access control
- QA (scans and originals)



Results - looking for 'Austria' in metadata finds only four records...



CERN Document Server: austria - Search Results - Windows Internet Explorer provided by CERN

http://cdsweb.cern.ch/search?ln=en&cc=Council-related+Committees&p=austria&f=&action_search=Search&c

CERN Document Server

Search Submit Help Your CDS login

Home > Articles & Preprints > Committee Documents > CERN Council Committees > Search Results: austria

CERN Council Committees

Search: austria any field Search Browse

Search Tips :: Advanced Search

Search collections: CERN Council Committees *** add another collection ***

Sort by: latest first desc - or rank by - Display results: 10 results split by collection Output format: HTML brief

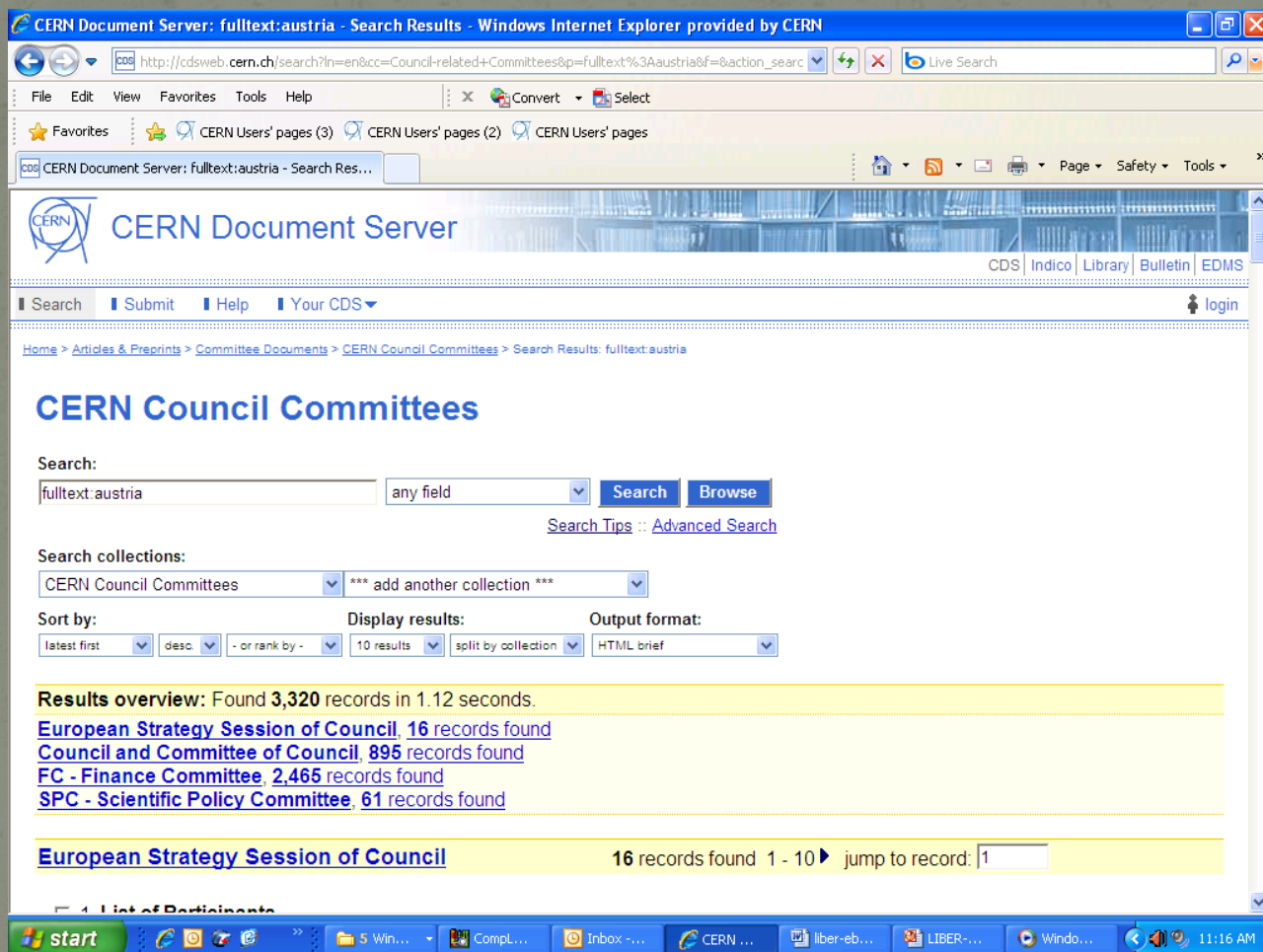
Results overview: Found 4 records in 0.03 seconds.

- [Council and Committee of Council](#), 1 records found
- [FC - Finance Committee](#), 3 records found

Council and Committee of Council 1 records found

1. Reintegration of the Members of the CERN Personnel into the Social Security Systems of the Member States - Draft Agreement with Austria
Réintégration des membres du personnel du CERN dans les systèmes sociaux des Etats membres priet d'accord avec l'Autriche

Searching the full texts finds 3,320 records



CERN Document Server: fulltext:austria - Search Results - Windows Internet Explorer provided by CERN

http://cdsweb.cern.ch/search?ln=en&cc=Council-related+Committees&p=fulltext:%3Aaustria&f=&action_search

CERN Document Server

Search Submit Help Your CDS login

Home > Articles & Preprints > Committee Documents > CERN Council Committees > Search Results: fulltext:austria

CERN Council Committees

Search: fulltext:austria any field Search Browse

Search Tips :: Advanced Search

Search collections: CERN Council Committees *** add another collection ***

Sort by: latest first desc - or rank by - Display results: 10 results split by collection Output format: HTML brief

Results overview: Found 3,320 records in 1.12 seconds.

- [European Strategy Session of Council, 16 records found](#)
- [Council and Committee of Council, 895 records found](#)
- [FC - Finance Committee, 2,465 records found](#)
- [SPC - Scientific Policy Committee, 61 records found](#)

European Strategy Session of Council 16 records found 1 - 10 ► jump to record: 1

1 List of Participants

start 5 Win... Compl... Inbox... CERN... liber-eb... LIBER... Windo... 11:16 AM

Combine search terms to find the information you want

A screenshot of a Windows Internet Explorer browser window. The address bar shows the URL "http://cdsweb.cern.ch/record/17882/files/CM-P00076309-e.pdf". The browser interface includes a menu bar (File, Edit, Go To, Favorites, Help), a toolbar with navigation buttons, and a search bar. The main content area displays a PDF document with the following text:

well justified the additional work it entailed for the Finance Office.

After fairly lengthy negotiations, CERN had signed the Site Agreement with the authorities of the Canton of Geneva. In that connection, he wished to stress how much the Organization appreciated the helpful attitude of Mr. Dutoit, Head of the Public Works Department, and how pleasant negotiations with his department had always been. Finally, the contract with Dr. Steiger, CERN's architect, had recently been renewed and CERN would thus benefit from his services in the drawing up of plans for the new buildings envisaged. The remarkable qualities of the building in which the Council was now meeting emphasized the value of his help.

6. ACCESSION OF AUSTRIA (Item 10 of the agenda)

The PRESIDENT invited the Council to settle the question of the accession of Austria. He recalled that Austria had requested to join CERN, in accordance with the provisions of Article III, 2 of the Convention. That country's formal request had not been communicated to the Member States three months before the Council Session, as stipulated in the Convention; nevertheless, Austria had long ago informed the Governments of the Member States through diplomatic channels of her wish to join the Organization. The consideration of Austria's request at the present Session might thus not be regarded as contrary to the spirit of the Convention. If all the delegations were in favour of the admission of Austria, there would be a definite advantage in her becoming a member of the Organization as from 1 July, 1959, since she would thus be in a position to share the advantages

3. The Wolfgang Pauli Archive



- 1900 – birth, Vienna
- 1924 – postulates the exclusion principle
- 1928 – professor at the ETH Zürich, CH
- 1930 – postulates the existence of the neutrino
- 1940-45 – professor at the Institute for Advanced Study, Princeton, USA
- 1945 – Nobel Prize for his exclusion principle
- 1958 – death, Zürich

Pauli's scientific legacy

- Archive includes correspondence, manuscripts and notes, photographs, books, certificates, prizes and other objects
- Donated to CERN by his widow, Franca Pauli



Objectives and scope

- Protect paper originals
- Improve access
- Collaboration with ETH Archive
- Around 3,800 items
- Diverse formats, some fragile



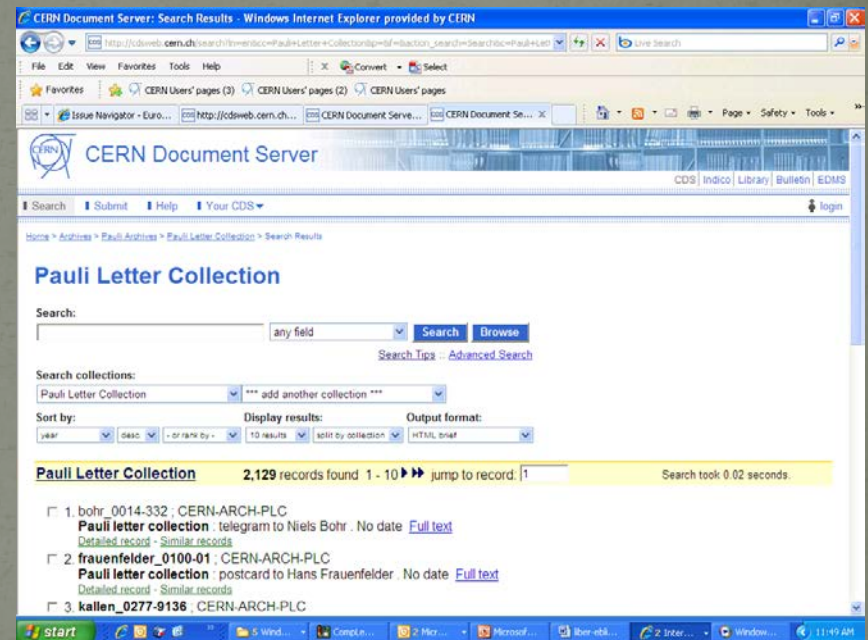
Workflow

- Cataloguing
- Scanning (in-house, or by local specialist firm); no OCR
- Cross-reference to published volumes of correspondence
- Copyright issues (letters, photos)



Results

- Over 2,000 letters, 200 photos, and 1,400 other documents scanned and made available online



The screenshot shows a Windows Internet Explorer browser window displaying the CERN Document Server search results for the Pauli Letter Collection. The browser's address bar shows the URL: http://cdsweb.cern.ch/search/?method=cw+Pauli+Letter+Collection&op=wf+&action_search=Search&loc=Pauli+Let. The page title is "CERN Document Server: Search Results - Windows Internet Explorer provided by CERN". The search results are displayed under the heading "Pauli Letter Collection". The search criteria are: "any field" and "Search". The search results are sorted by "year" and display 10 results. The first three results are:

1. bohr_0014-332 ; CERN-ARCH-PLC
Pauli letter collection : telegram to Niels Bohr. No date [Full text](#)
[Detailed records](#) - [Similar records](#)
2. frauenfelder_0100-01 ; CERN-ARCH-PLC
Pauli letter collection : postcard to Hans Frauenfelder. No date [Full text](#)
[Detailed records](#) - [Similar records](#)
3. kallen_0277-9136 ; CERN-ARCH-PLC

The search results indicate that 2,129 records were found, and the search took 0.02 seconds. The browser's taskbar at the bottom shows the start button and several open applications, including Windows Explorer, Microsoft Word, and Internet Explorer.

Comparison of projects

	Scope	Special issues	Approach	Main motivation
Theoretical Physics Dept. reports	2,500 documents scanned (from total of 12,000)	Historic series Hunting for e-versions and missing items	In-house OCR Library DB	Wider access (support of OA)
CERN Council documents	21,000 documents scanned (EN & FR)	Key records Access restrictions Extra funding	Mostly outsourced OCR Library DB	Wider access (part of Council's push for openness)
Wolfgang Pauli archives	3,600 documents and 200 photos scanned	Diverse formats Fragile Copyright Wide interest	In-house and outsourced No OCR Archive DB	Conservation Wider access (popular) Collaboration with ETH

