

DATA ON DEMAND

Collections as Data as a bridge between
collections as infrastructures

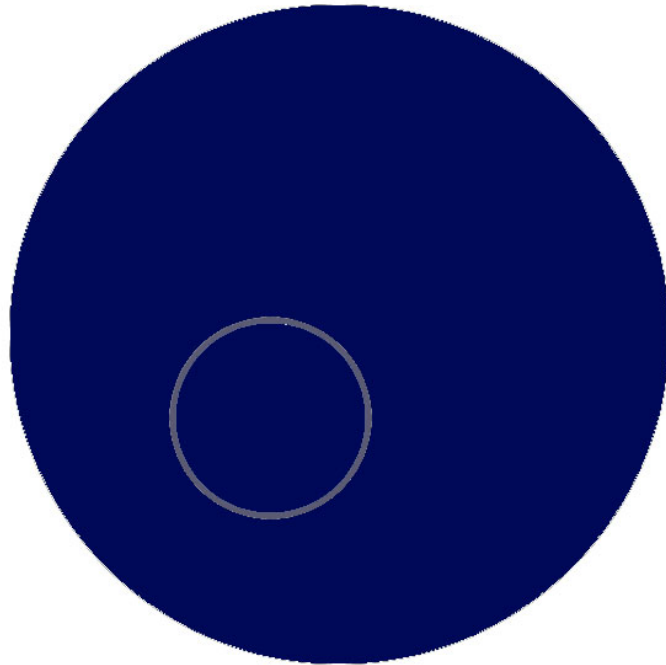
Ditte Laursen, Senior researcher, Ph.D., Royal Danish Library
Collections as Data: Collaborating across
Data Spaces for Cultural Heritage and Open Science
KBR | Royal Library of Belgium, 19th - 20th February 2024



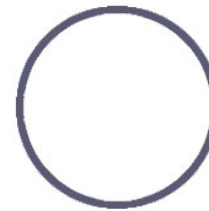
**DET KGL.
BIBLIOTEK**

Royal Danish Library

SERVICE: DATASETS ON DEMAND



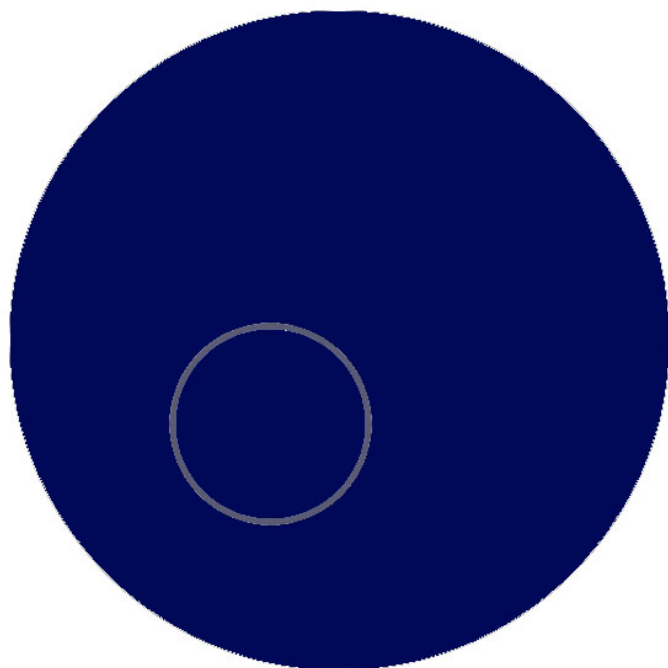
~ 10 petabytes



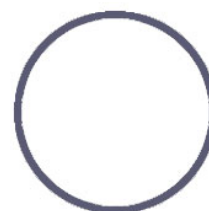
A set of objects that constitutes a meaningful unit.
Can be based on material type, content, author, provenance, period, etc.



RECORDS DANISH PRINTED BOOKS 1600-1900



410 million records



900.000 records
national bibliography



**DET KGL.
BIBLIOTEK**

Royal Danish Library

POTENTIAL

History of the book collection (size, genres, languages, publication places ...)

Understanding the cultural context in Denmark compared with collections in other countries

Specific threads, like genres (ie. cook books), periods (ie. the golden age) ...



**DET KGL.
BIBLIOTEK**

Royal Danish Library

CHALLENGES

Extraction tools
410 millions records

Selection of data
Merged classification systems
Internal administrative data
Licensed content

Cleaning
Records (not OCR)
No authority database
Too trivial for a programmer



Librarian

User

Size
900.000 records

Format
Marc21 xml to .csv

Quality
Not ready-made data

Sense-making of data



BENEFITS

- Open access
- Open science
- Open education



Librarian

User

Free from the constraints of the library's infrastructure facilities

Possible to handle data as a dataset

