MONK
An AI-based application in the Digital Humanities

Ove Kähler
Vice President Operations
kahler@brill.com

#HumanitiesMatter
Brill

- Scholarly publisher founded in 1683
- Offices in Leiden (NL), Boston (US), Singapore (SG), Beijing (CN), Paderborn (DE)
- +160 staff
- +300 journals, +1450 books per year
- Major reference works (encyclopedia, primary sources etc)
- Leading in selected areas of the humanities, social sciences & international law

#HumanitiesMatter
Primary Sources
Facsimile, microfiches

#HumanitiesMatter
Primary Sources Online

Example: Newspapers
Primary Sources Online
Example: Western Quran translations
Primary Sources Online
Example: “Prize Papers”
Primary Sources Online
What’s the challenge

• OCR on HISTORIC documents does not work well
• OCR on handwritten manuscripts doesn’t work at all
• Transcribing handwritten manuscripts is very time-intensive and costly

#HumanitiesMatter
Primary Sources Online
What’s the solution

Image credits: Lambert Schomaker
Primary Sources Online
What’s the solution ... MONK

Making Sense of Illustrated Handwritten Archives

Providing access to the hidden treasures of the ‘Natuurkundige Commissie’ (1820-1850)

#HumanitiesMatter
MONK
What is it

• Makes handwritten documents full-text searchable
• Based on image recognition & machine-learning

• Takes into account sloppy letter shapes & individual writer styles
• Applicable also to non-Latin scripts
How MONK works
Individual word labels

Monk - Hit list word shape (Carolina) PP Part1 Set01 [*]  bk#PP_Part01_0.

No codf {};
Items in hit list of Carolina numbered: 0-3999. (now showing: 0-27)

#HumanitiesMatter
How MONK works
Labelling words per line
How MONK works
Labelling full lines
How MONK works
Labelling full lines with autosuggestions
How MONK works

Makes use of human trainers
MONK
How it benefits scholars (and Brill)

• Publish test collection OA, including annotations and being fully searchable
• Make existing primary source collections fully searchable
• Offer MONK as Software-as-a-service to digitize other collections
Thank you!

Ove Kähler
Vice President Operations
kahler@brill.com

#HumanitiesMatter