



# News media processing and interactive presentation

IFLA 2017 News Media Satellite Conference | Dresden

**Claus Gravenhorst**

Director Strategic Initiatives

# About CCS | Some facts

- CCS - Content Conversion Specialists is a privately owned company with headquarters in Hamburg, Germany
- Technology company developing market-leading software and hardware for the creation and display of digital collections
- Participating in European research projects:
  - METAe – The Metadata Engine (2000 – 2003)
  - ENP – Europeana Newspapers Project (2012 – 2015)
- Participating in US research project:
  - Library of Congress (2004), NDNP specification

# About CCS | Selected references

- References for software solutions:
  - Brightsolid (UK), Library of Congress and various National Libraries in Europe, Australia and Asia
  - service providers like Contentra, Digital Divide Data, ...
- References for services:
  - The British Library, Dutch National Library, National Library of Norway and ...
  - publishers like Springer, FAZ, ...
- Today CCS is one of the world leaders in the provision of information through digitization and conversion.

# ENP – Europeana Newspapers Project

- CCS, as technical project partner, provided its expertise and docWorks technology to set up and operate a mass digitization workflow for creating high quality structured content from 2 million scanned newspaper pages provided by 5 library partners
- Page volume:  
BNF=1.000 k, NLE=500 k , SUB HH=480 k, NLF=90 k, SBB=10 k
- The distributed OLR workflow enabled the contribution of project partners (content providers) to the integrated quality assurance process
- CCS has also contributed to the specification of the ENMAP\* metadata model

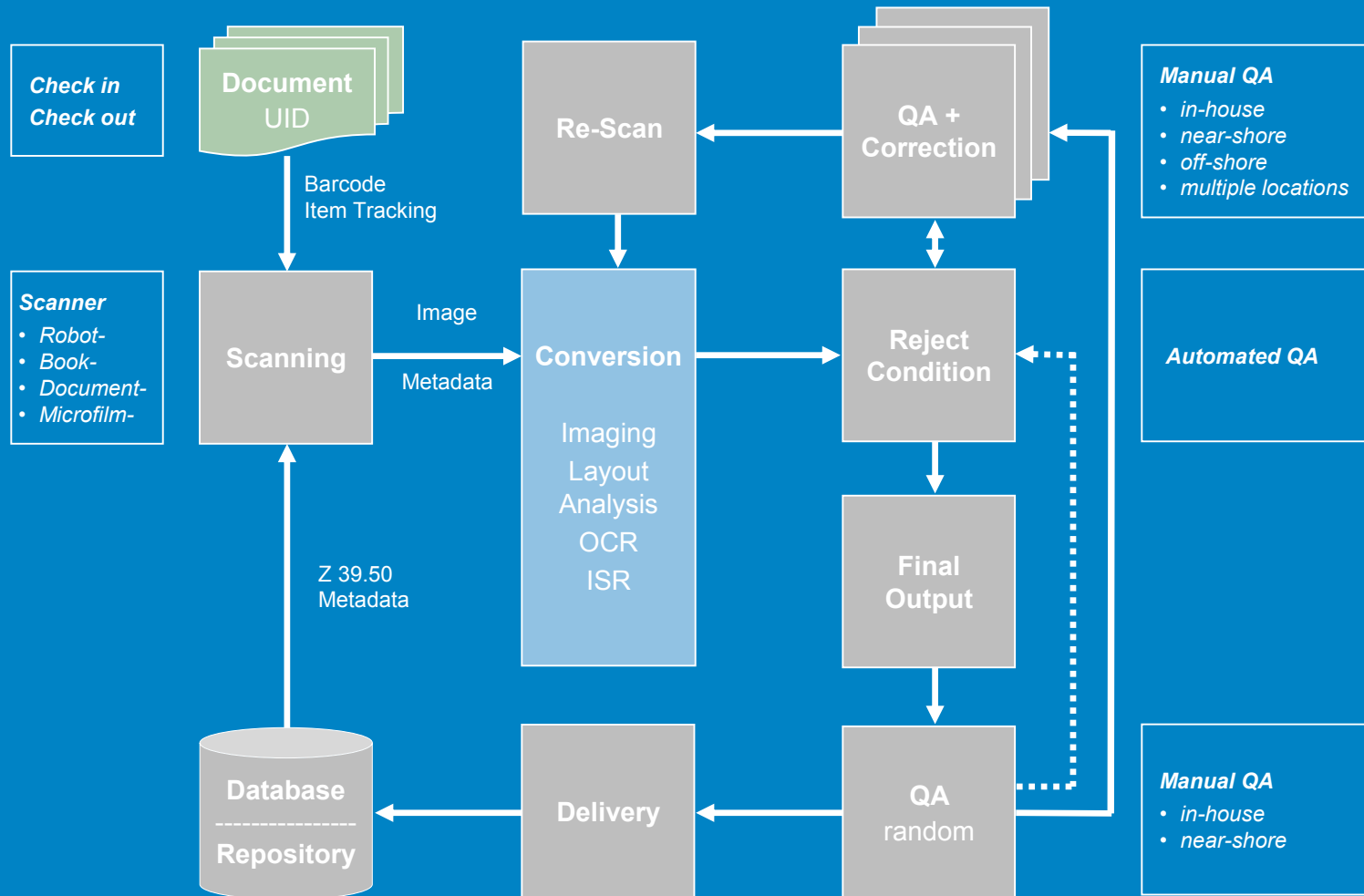
\* ENMAP = Europeana Newspapers Mets Alto Profile

# Structure Analysis | Newspaper

- General rule system enables recognition of words, text lines, text blocks, columns and classification of text blocks, illustrations, advertisements, tables and the following page types:
  - title page (the title page of an issue)
  - content page (a page that consists of content/text only)
  - illustration page (a page that has at least one illustration)
  - advertisement page (a page that contains adverts only)
- Structure analysis through classification of headlines and grouping of zones into articles (incl. article continuation)



# Distributed Digitization Workflow



# docWorks & Kitodo

- 1-year project with Saxon State Library Dresden (2016)
- Goal: enhance quality for Kitodo users by integrating docWorks features “Image pre-processing”, “Layout analysis”, “OCR (Tesseract)” and “Structure Analysis” into the Kitodo workflow
- Beside newspapers also other document types like magazines and books are supported
- Automated server process:
  - Kitodo provides images, Kitodo XML and XML setting file
  - automated background processing through docWorks
  - docWorks provides layout, text and structural information
  - automated data import by Kitodo
- re-use scenario will be defined and fixed in 2<sup>nd</sup> half of 2017
- Next steps: further improvement of integration, continuous development of new docWorks features

# Content Display - The challenge

- How to meet the needs of today's users?
- Benefit from your knowledge about the content!
- Foster new ways to access and experience content!
- How about offering selected content in virtual exhibitions?
- How about combining the presentation of origin items, their digital copies and related digital content?



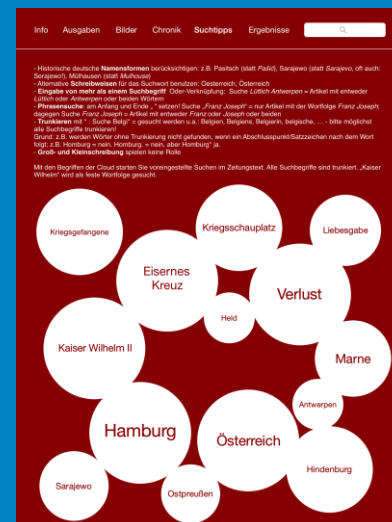
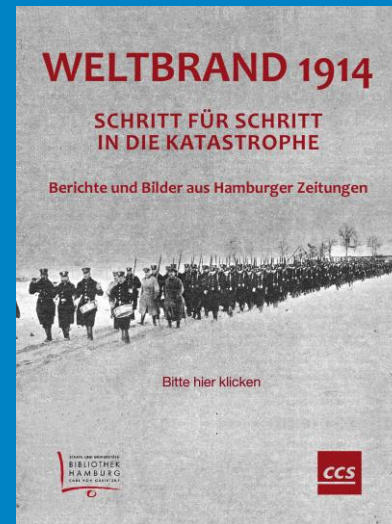
# CES – Content Experience Solutions – Touchpad

## 1914 world war one

- Created for State and University Library Hamburg
- Runs on iOS and Android
- Collection of approx. 1000 pages
- Incremental download
- Fulltext index for searching
- Cloud with predefined search terms
- Chronic with links to Wikipedia
- Picture collection

<https://itunes.apple.com/de/app/weltbrand-1914-bilder-und/id863960289?l=en&mt=8>

<https://play.google.com/store/apps/details?id=de.stabihh.weltbrand.android>



# The MagicBox (I)



- Presentation of and access to curated content, both origin and digital items
- Users can interact with the digital content
- Video and 3D animation can be combined with the showcased items
- It's all about contextualizing objects and related story telling

See videos on Vimeo:

<https://vimeo.com/108877847>

<https://vimeo.com/189295246>

# The MagicBox (II)

- Product history:

  - 2014 – prototype introduced in Europe

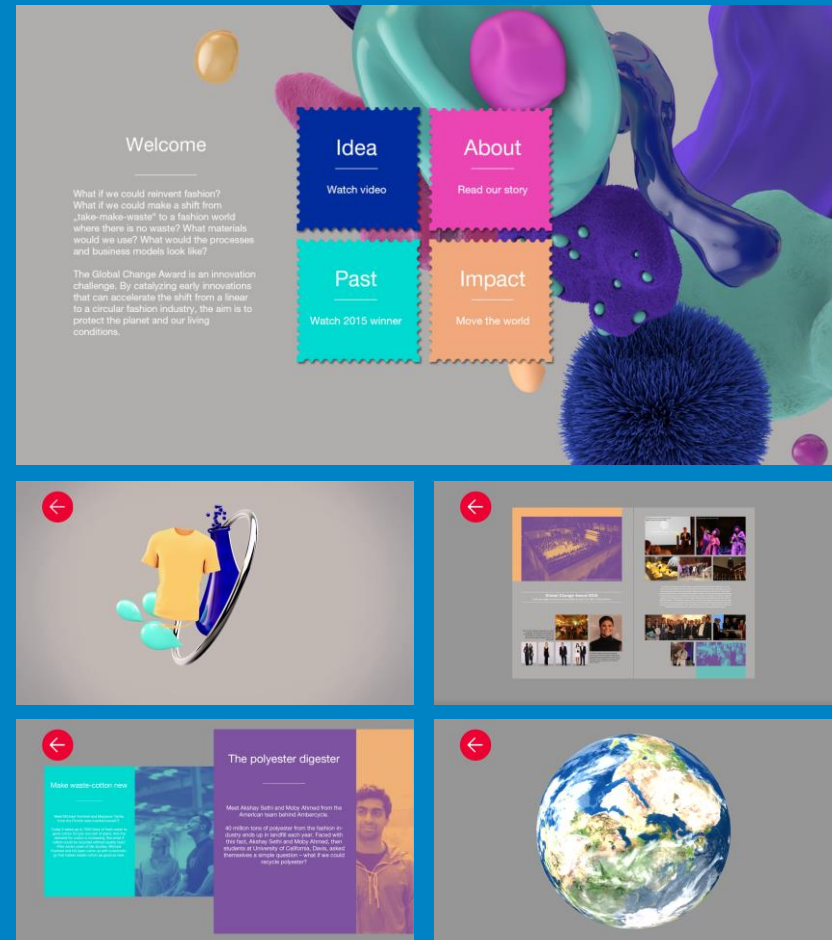
  - 2015 – global product launch

- References:

  - National Library Vietnam (2 boxes)
  - Datacom IT, Australia
  - Saudi Arabia
  - University Library Belgrade, Serbia
  - Leica Camera AG, Germany
  - Irish Distillers, Ireland
  - Cleveland Public Library, USA
  - Gold Coast City Council Library, Australia
  - Lancaster University Library, UK
  - University Library of Florida, USA
  - Guinness, Ireland
  - National Library of Slovakia
  - University Library Leuven, Belgium
  - Sharjah Library, UAE

# MagicTouch

- Software application for arranging, editing and presenting multimedia content on PC-operated touch devices
- Go live with your exhibition stories by selecting the corresponding digital content, easily arrange it with the integrated DataManager and present it on a touch screen
- It's all about contextualizing digital objects like books, magazines, pictures, videos, 3D models and the related storytelling
- Support for various standard formats: MP4, AVI, MOV, WMV, JPG, TIF, PNG, JP2, BMP, XML, 3DS, OBJ
- MagicTouch as turnkey solution:
  - Kiosk with 32"/42" Full HD touch screen
  - Integrated industrial PC with MS Windows OS
  - Integrated soundbar



# Conclusion

- CCS provides solutions and relevant project experience to support current and future digitization projects/programs of the News Media Community
- We are part of this community since more than 20 years and looking forward to entering into next generation partnerships

# Thank you!

**Claus Gravenhorst**

Director Strategic Initiatives

**CCS Content Conversion Specialists GmbH**

Weidestr. 134

22083 Hamburg

Germany

T +49 40 227130-16

F +49 40 227130-11

M +49 176 12713016

[c.gravenhorst@content-conversion.com](mailto:c.gravenhorst@content-conversion.com)

[www.content-conversion.com](http://www.content-conversion.com)