



Content Conversion Specialists

CCS – sponsor presentation

2015 IFLA International News Media Conference, Stockholm

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), NDNF 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

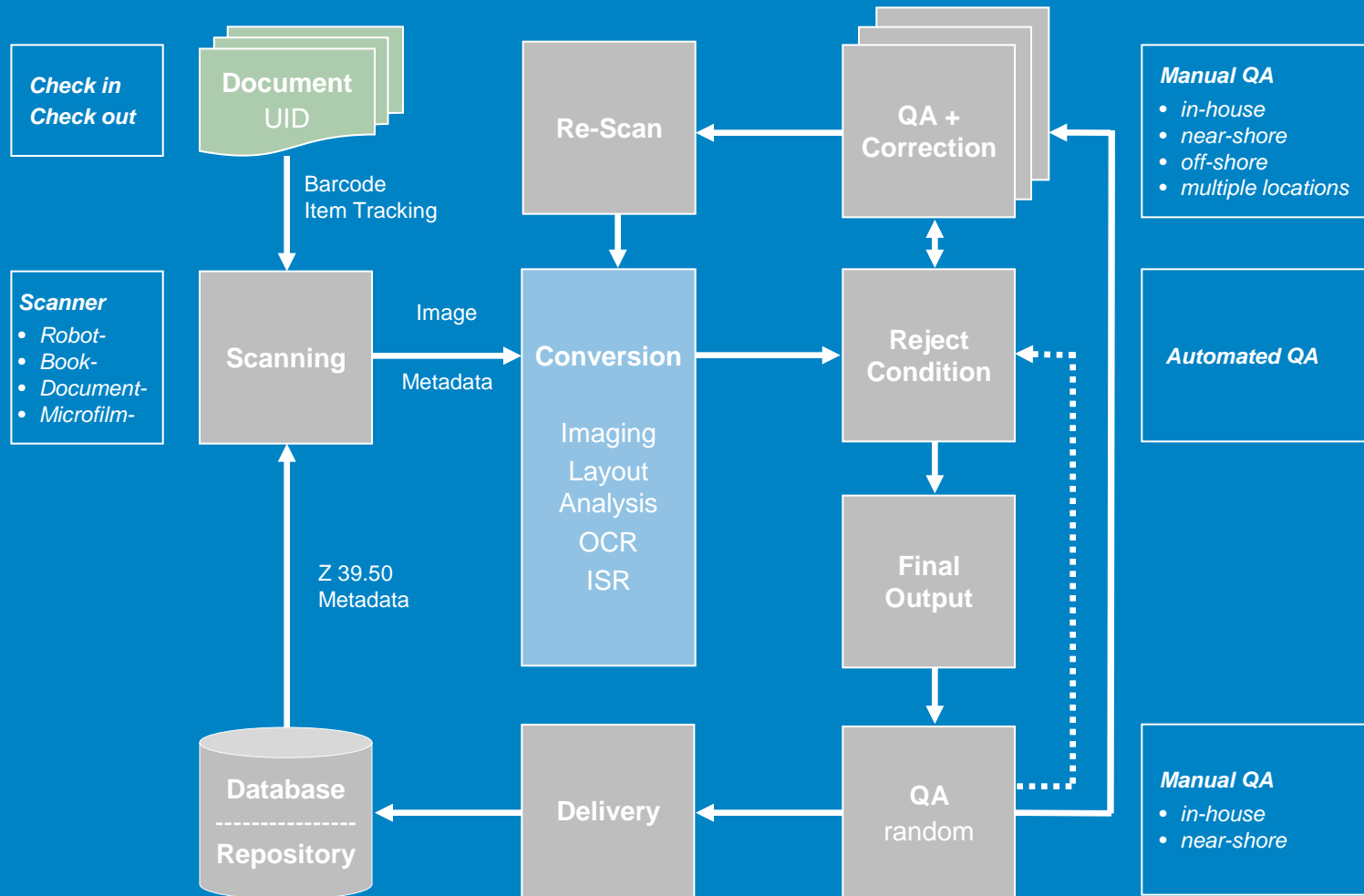
* ENMAP = Europeana Newspapers Mets Alto Profile

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



Possible conversion scenarios

- Conversion at library (on-site)
- Conversion at library (on-site), QA outsourced to service provider via internet transfer (remote QA solution)
- Conversion near/-off-shore at service provider incl. basic QA, final QA at the library via remote QA solution
- Conversion at service provider

Remote QA at library

QA on-site
@ Library

INPUT

Internet

Offshore
Processing
@ CCS

RQA

GUI



Services



Storage



dW Share

POOL

Master

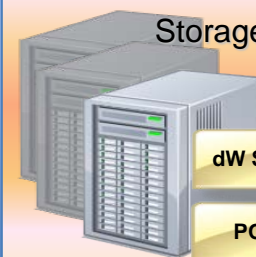
GUI



Services



Storage



dW Share

POOL

IN

OUT

OUTPUT



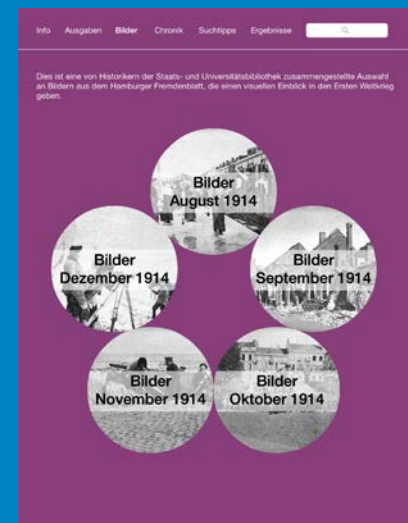
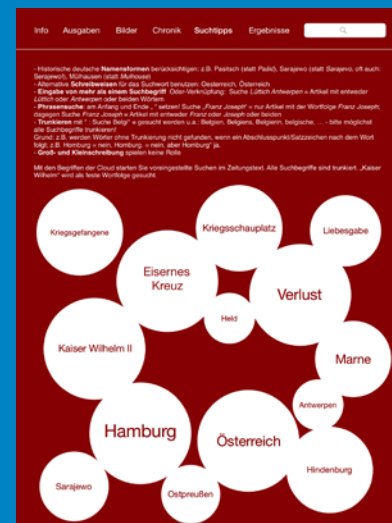
CES – Content Experience Solutions – Touchpad

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>



CES – Content Experience Solutions – MagicBox



See video on Vimeo:
<http://vimeo.com/108877847>

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

www.content-conversion.com

