



Post-scan workflow & content experience solution

UBSM 90th anniversary, May 24th 2016 | Belgrade

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
- Founded in 1976 , 50+ employees (Germany, Romania, US)
- Participating in US research project:
 - Library of Congress (2004), NDNF specification
- Participating in European research projects:
 - METAe – The Metadata Engine (2000 – 2003)
 - ENP – Europeana Newspapers Project (2012 – 2015)
- Mass digitisation projects:
 - The British Library (books, 2007 – 2008)
 - Dutch National Library (newspapers, 2008 – 2012)

Portfolio

Consulting

Successful projects start with careful listening.

CCS has one of the best expertise in digitisation programs and project management.

We analyze your requirements and offer a customized workflow design based on your architecture and standards.

Technology

Providing cutting-edge technology to meet the requirements of the market.



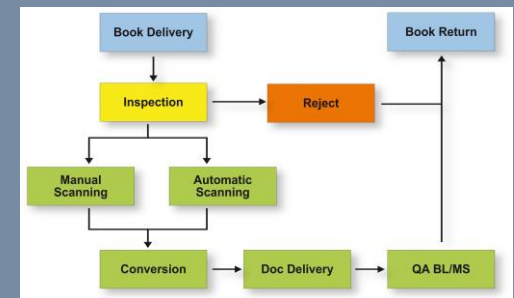
Content
Experience
Solutions

Digitisation Services

Processing more than 2 million pages every month in the lead libraries of the world.

Best expertise in applied integrated digitisation workflow. Dealing with most valuable items.

Newspapers / Books
Journals / Magazines



docWorks | Selected References

docWorks I

National Library of Norway
National Library of Finland
National Library of Luxemburg
National Library of Latvia
National Library of Estonia
National Library of Slovakia
National Library of Poland

National Library of Vietnam
National Library of Australia

University of California, Riverside
University of Texas, Austin
Library of Congress, Washington DC
Harvard University, Cambridge
Princeton University, Princeton
National Library of Medicine, Bethesda
Washington State Library, Olympia
Library of Virginia, Richmond
Queens Library, New York
Cleveland Public Library
Indiana State Library, Indianapolis
J. Paul Getty Trust, Los Angeles
National Library of Trinidad & Tobago

docWorks II

Digital Divide Data, USA
Backstage Library Works, USA
Hudson Microimaging, USA
brightsolid., UK
LETA, Latvia
CD Imaging, Singapore
Contentra Technologies, India

More than 100 million book pages
More than 20 million newspaper pages

digitizationServices

British Library, London
Royal Library of the Netherlands
National Library of Luxemburg
National Library of Finland
National Library of Norway
Royal Library of Denmark
National Library of Latvia
National Library of Austria
Wellcome Library, London

FAZ, Germany
Axel Springer Verlag, Germany

Library of Congress, Washington DC
Washington State Library, Olympia
University of California, Riverside
University of Minnesota, Minneapolis
Michigan State University, East Lansing
National Library of Australia
National Library of Singapore
National Library of New Zealand

More than 25 million book pages
More than 15 million newspaper pages

Creating a Digital Library



Select



Scan



Convert



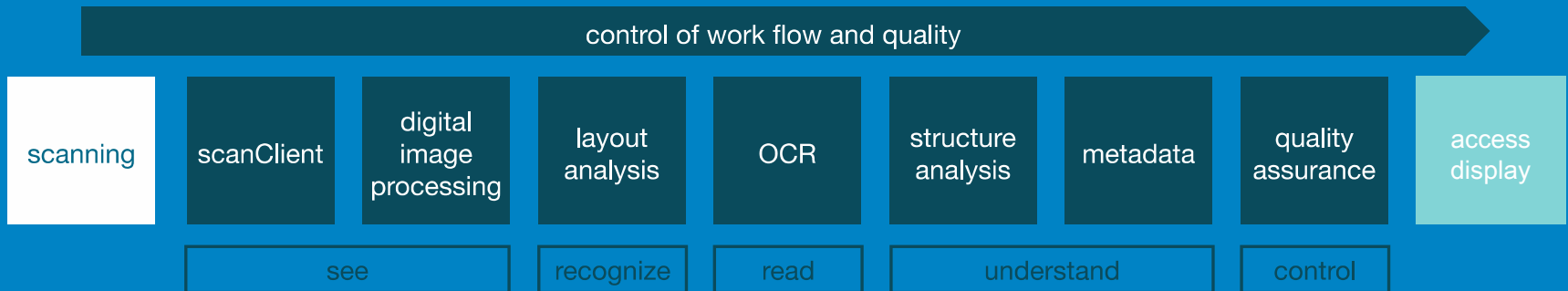
Archive



Display

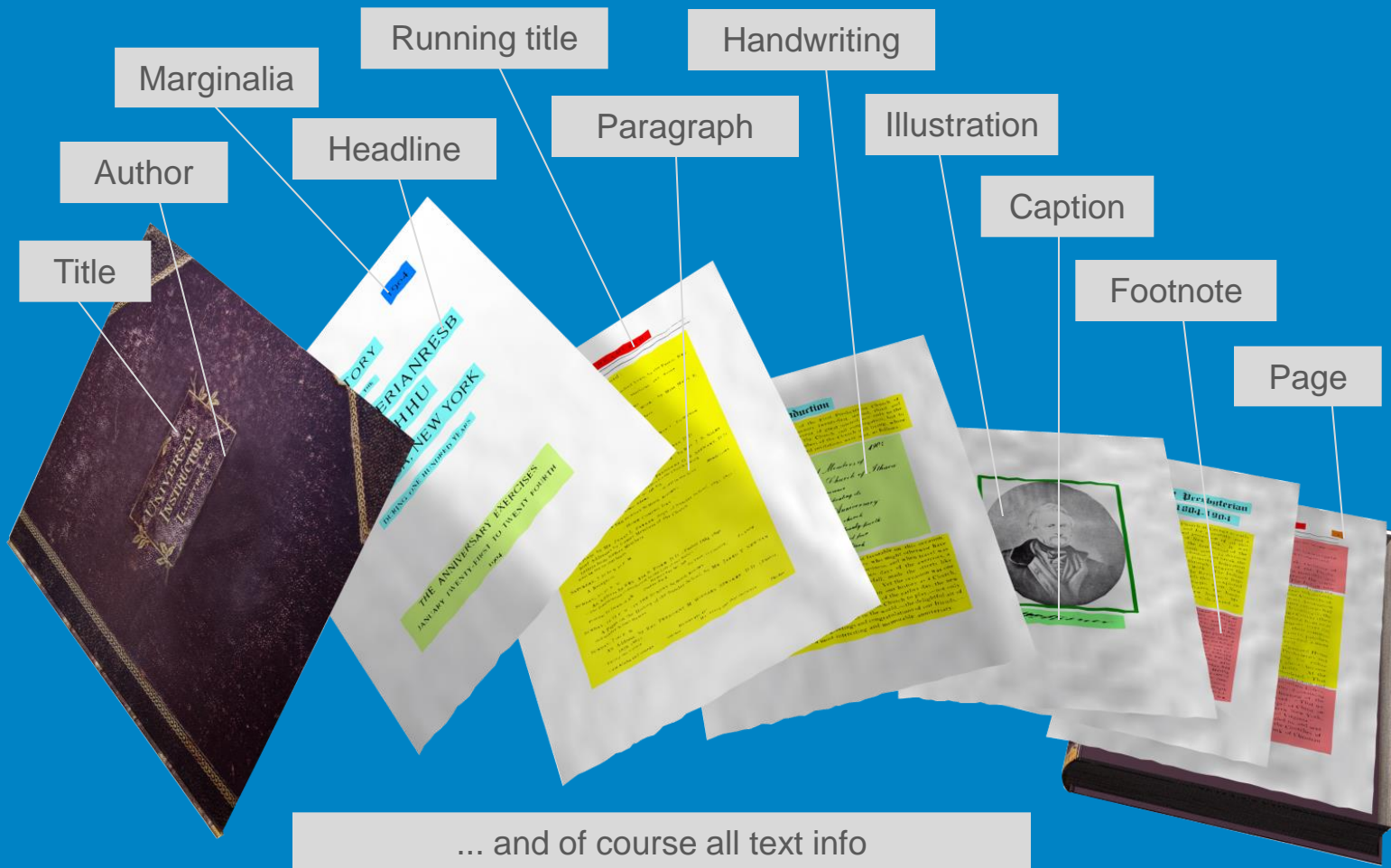
What is docWorks

- docWorks is a continuous digitisation workflow software including automated and interactive quality assurance options for every processing step.
- It's a highly scalable solution serving small, medium and large scale digitisation projects.



- According to the required output features processing steps can be activated and deactivated.

Structure Analysis | Monograph

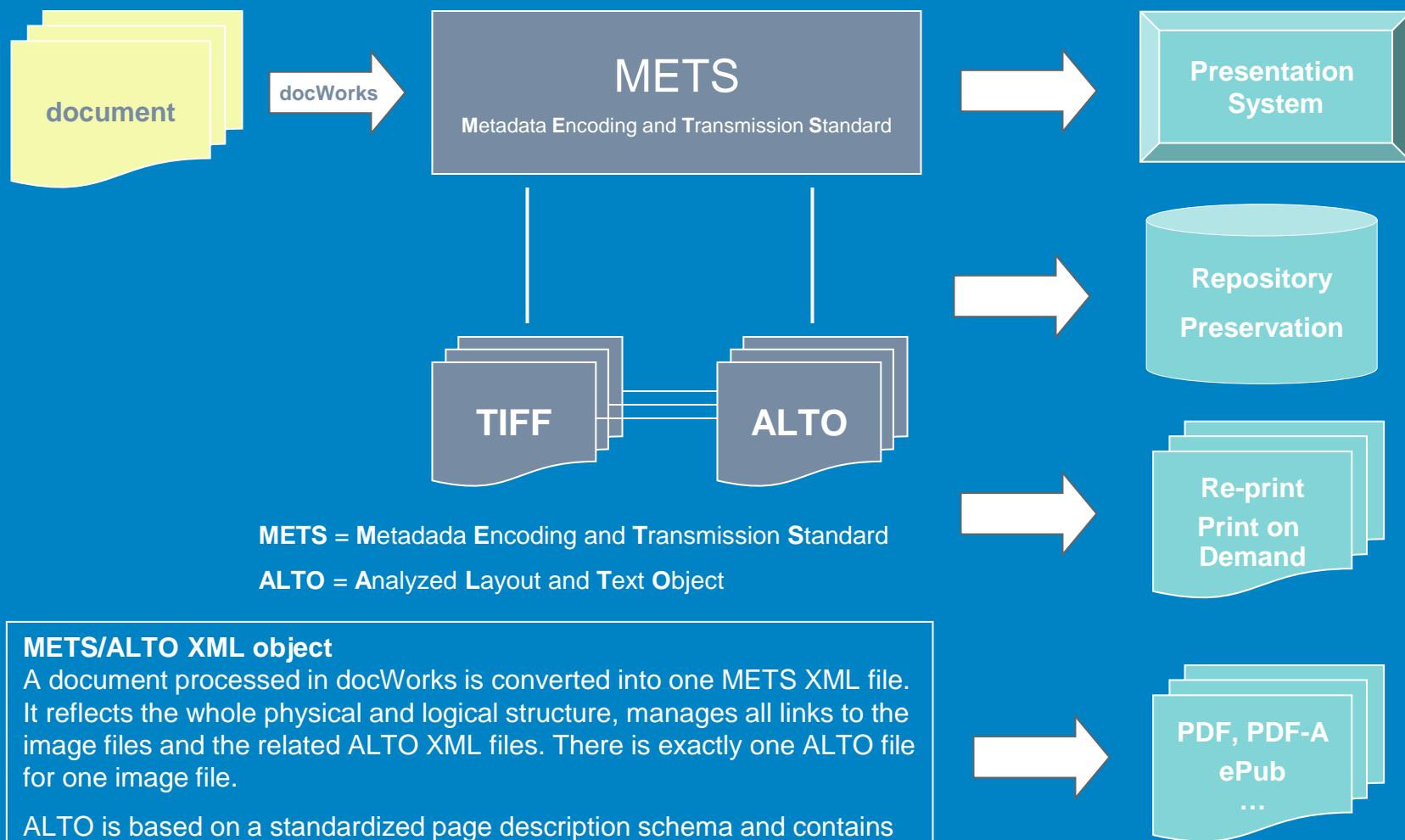


Structure Analysis | Newspaper

- 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)



Output | METS/ALTO package



METS/ALTO XML object

A document processed in docWorks is converted into one METS XML file. It reflects the whole physical and logical structure, manages all links to the image files and the related ALTO XML files. There is exactly one ALTO file for one image file.

ALTO is based on a standardized page description schema and contains all information of a page (print space, margins, coordinates, OCR results).

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 original items, their digital copies and related digital content?

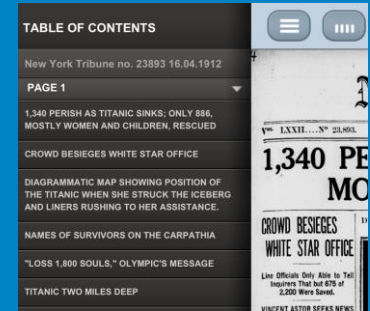
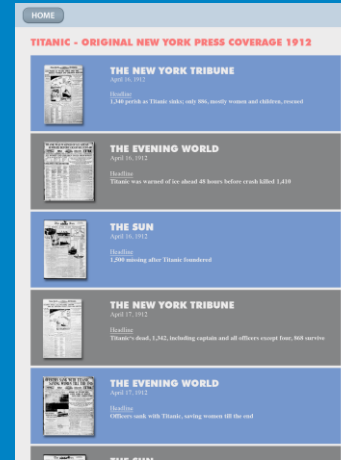
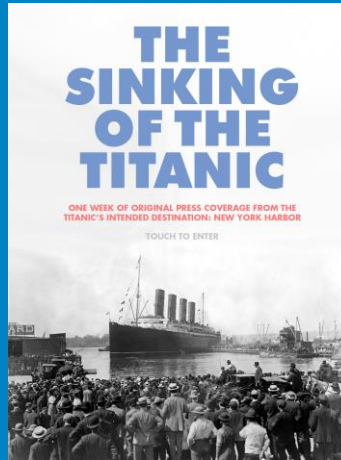
Apps (iOS, Android)



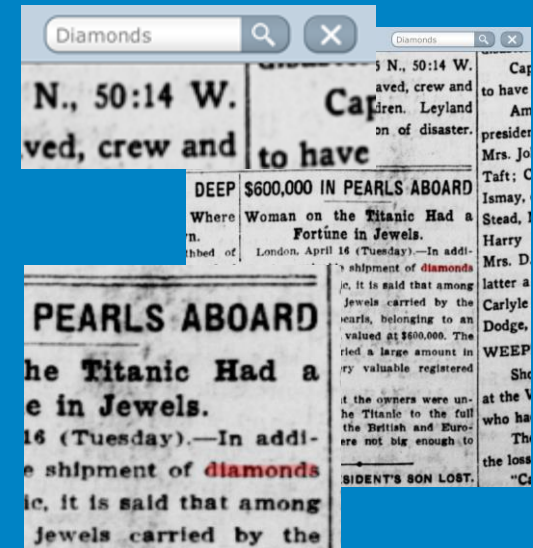
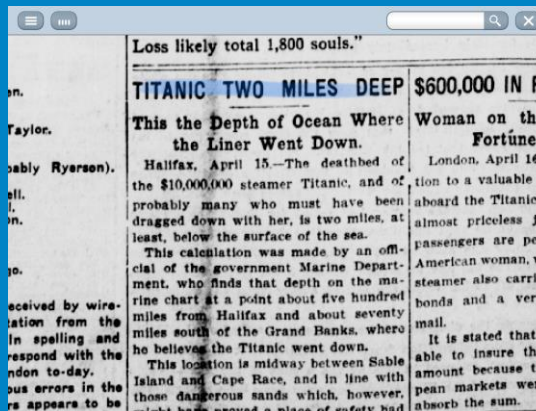
Titanic NY: More than 6,000 downloads

iPad App Titanic

- Attractive title page
- Table of contents (TOC) for issue selection
- TOC for article selection
- Article display with automated zooming and headline highlighting



- Text retrieval and article selection from hit list
- Article display and keyword highlighting
- Titanic iPad App:



<https://itunes.apple.com/de/app/titanic-new-york-press/id654662489?l=en&mt=8>

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:

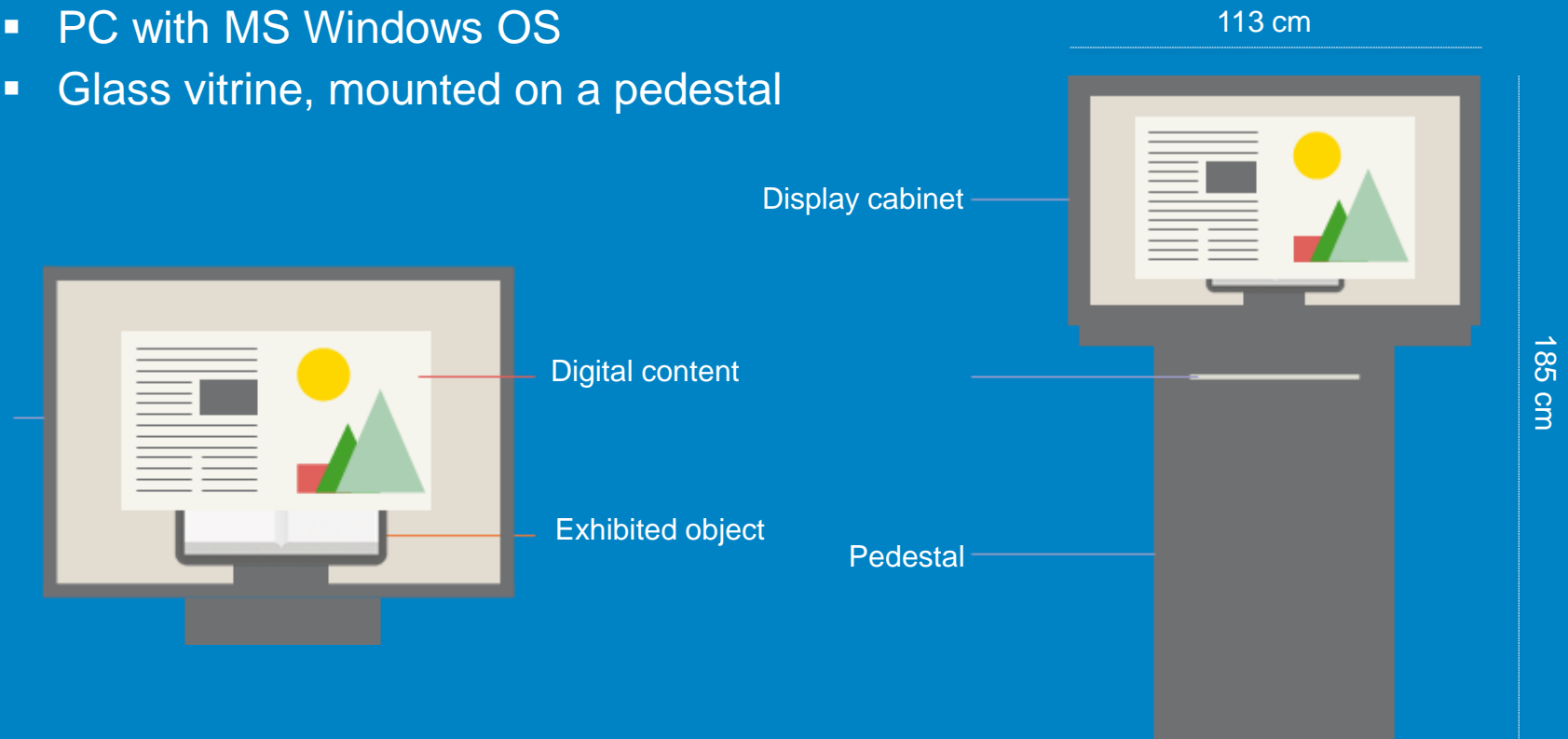
<https://vimeo.com/108877847>

<https://vimeo.com/140187203>

The MagicBox (II)

The MagicBox technology consists of the following parts:

- 47 inch transparent LCD display with an Infrared touch panel
- Application software to showcase:
 - books, magazines, newspapers, pictures, videos and 3D models
- PC with MS Windows OS
- Glass vitrine, mounted on a pedestal



The MagicBox (IV)

Product history:

- 2014 – prototype shown at a few exhibitions in Europe
- Product launch in May 2015
- References:
 - National Library Vietnam (2 boxes, June 2015)
 - Datacom IT, Australia (May 2015)
 - Saudi Arabia (June 2015)
 - University Library Belgrade, Serbia (November 2015)
 - Leica Camera AG, Germany (November 2015)
 - Irish Distillers, Ireland (November 2015)
 - Cleveland Public Library, USA (November 2015)
 - Gold Coast City Council Library, Australia (January 2016)
 - Lancaster University Library, UK (April 2016)

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