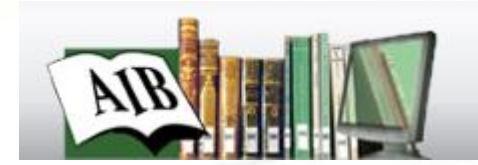


# AIB - Virtual Lab 2.2

*Digitalizzazione sostenibile*

*Digitalizzazione delle raccolte [ SDG 9-11]*

Klaus Kempf [C] Gabriele Lo Piccolo [F]



AIB - Virtual Lab 2.2

# Workflow Solutions

---

How we plan a project?

**ScanGate™ / ScanFlow™**

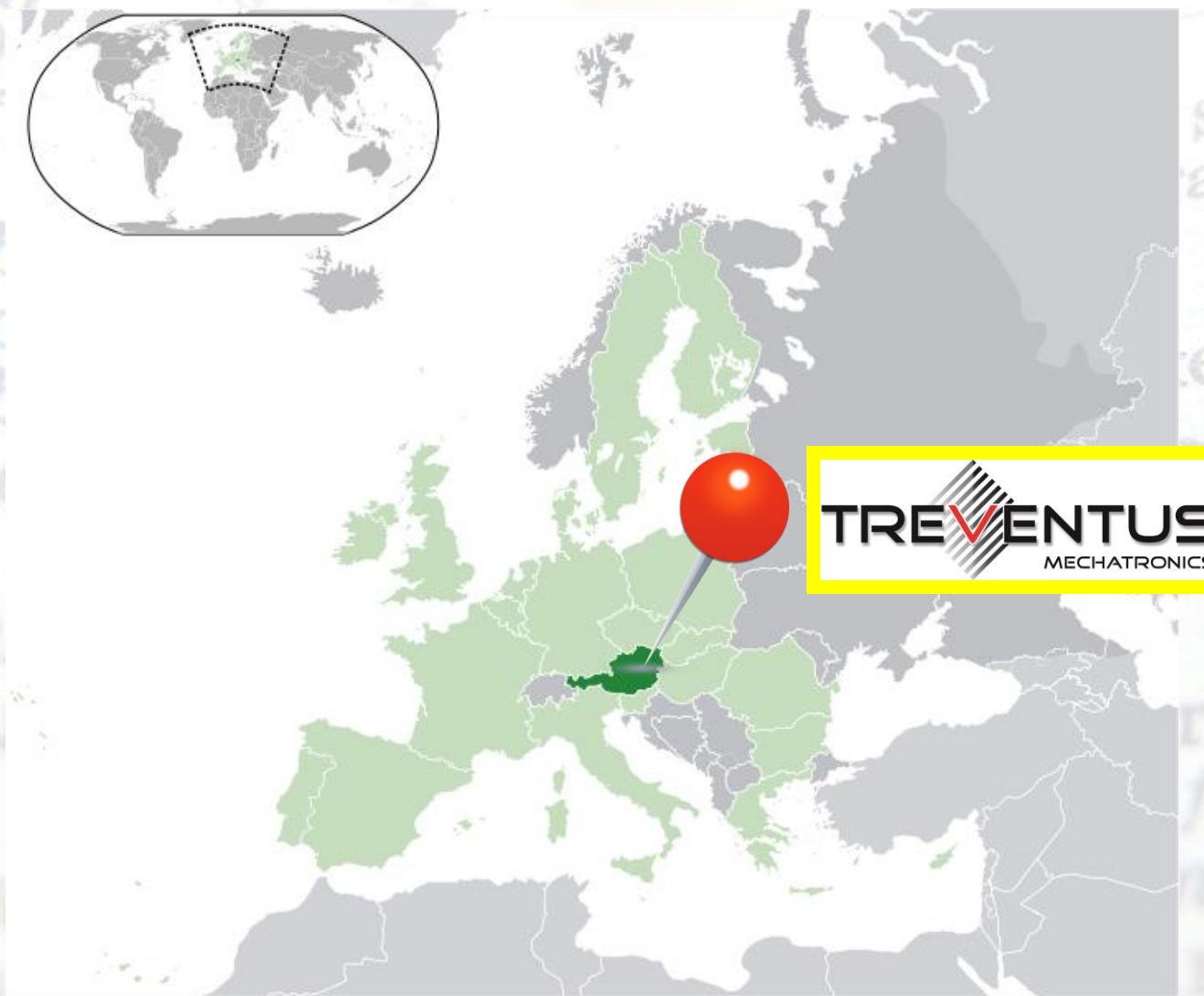


by Stephan Tratter

TREVENTUS

# TreVentus Mechatronics – Vienna

---



# Spin OFF – Vienna University of Technology



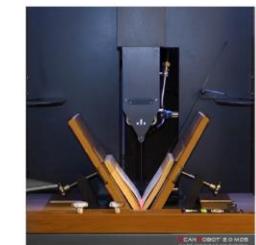
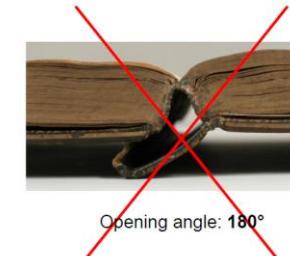
scan

## SCANROBOT® 2.0 MDS MASS DIGITIZATION SYSTEM



- Unique V-shape
- Automatic page turning
- Fast and productive
- Gentle
- Robust

V-Shape = careful to books



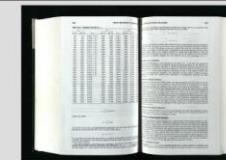
Opening angle: 60°

flat, distortion free, single

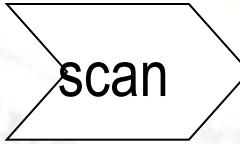
Planetary scanner



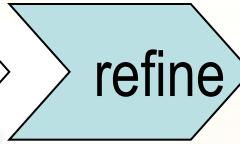
ScanRobot®



TREVENTUS



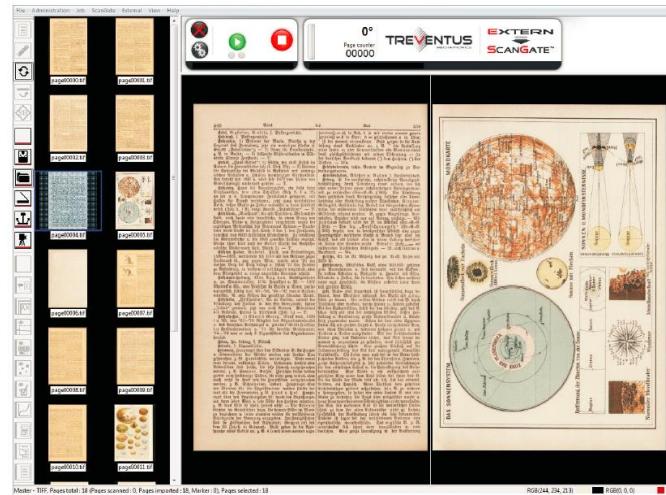
scan



refine

# SCANGATE™

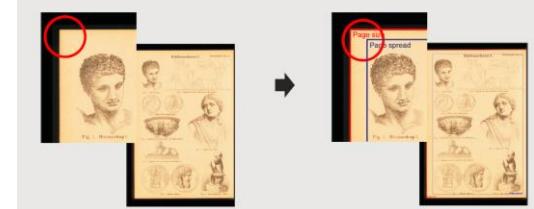
Capturing, Processing & Management Software



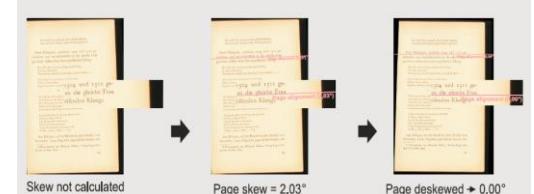
- Easy and intuitive
- Various scanner-interfaces
- Realtime preview
- Comprehensive functions
- Metadata interface



## Automatic border recognition



## Deskewing

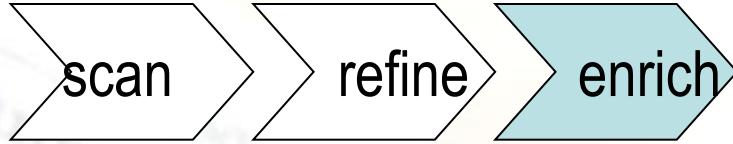


## Cropping



*...and lots more...*

**TREVENTUS**



scan

refine

enrich



TREVENTUS

# OCR SOLUTIONS

Software for automated text recognition

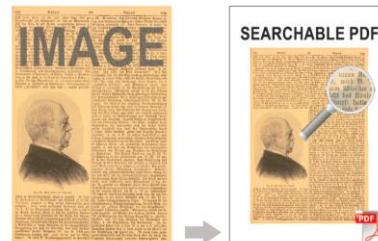


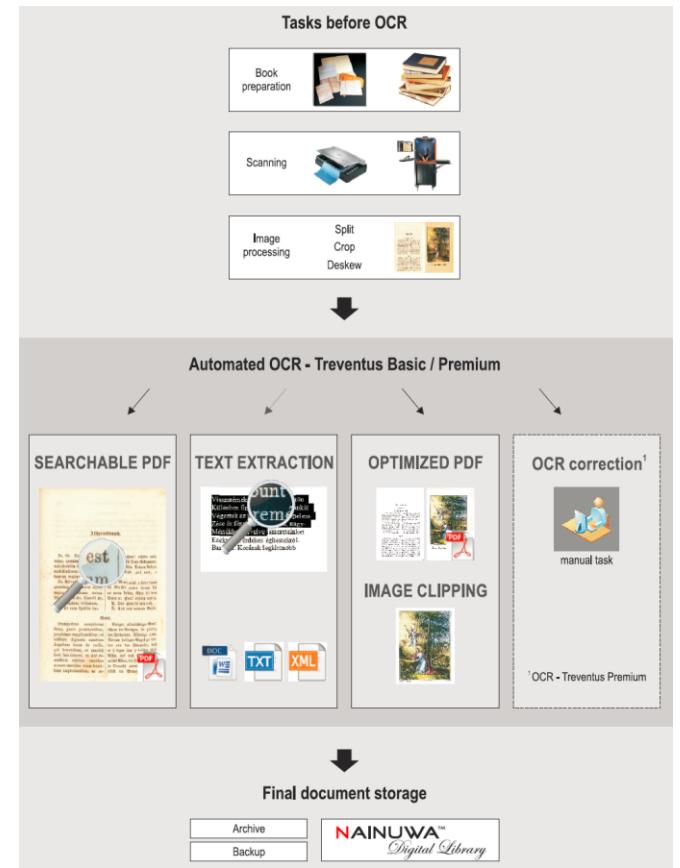
IMAGE CLIPPING



TEXT EXTRACTION



- Highly automated
- Direct integration
- Flexible package size
- >190 languages
- dictionary support



**TREVENTUS**

Specify

scan

refine

enrich

workflow

# SCANFLOW™ Professional

Workflow & Management Software



- Manage
- Organize
- Automize
- Highly flexible
- Multiple Outputs & lot more...

## Multiple workflows

Workflow 1



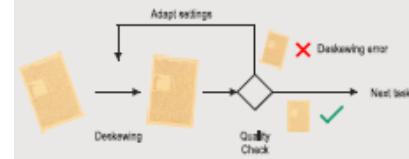
Workflow 2



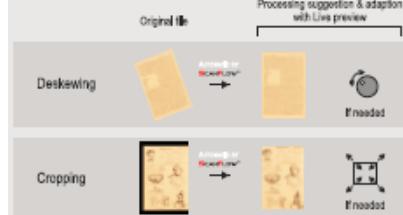
Workflow 3



## Undo function (e.g. Deskewing)



## Processing suggestion & Live preview



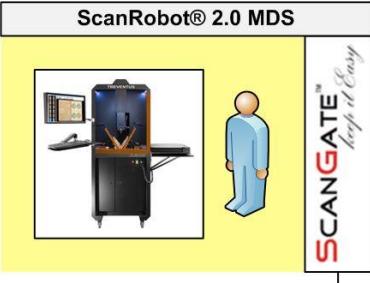
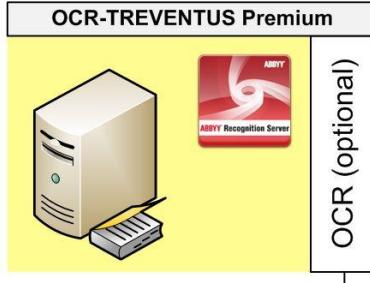
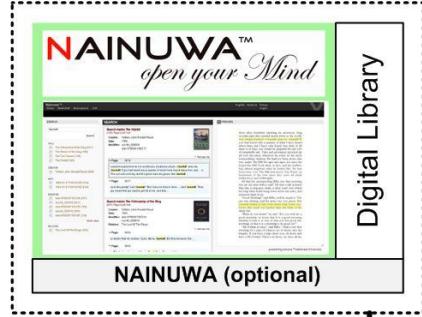
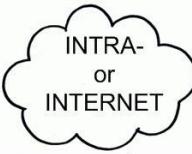
# Concept: ScanFlow® WF

---

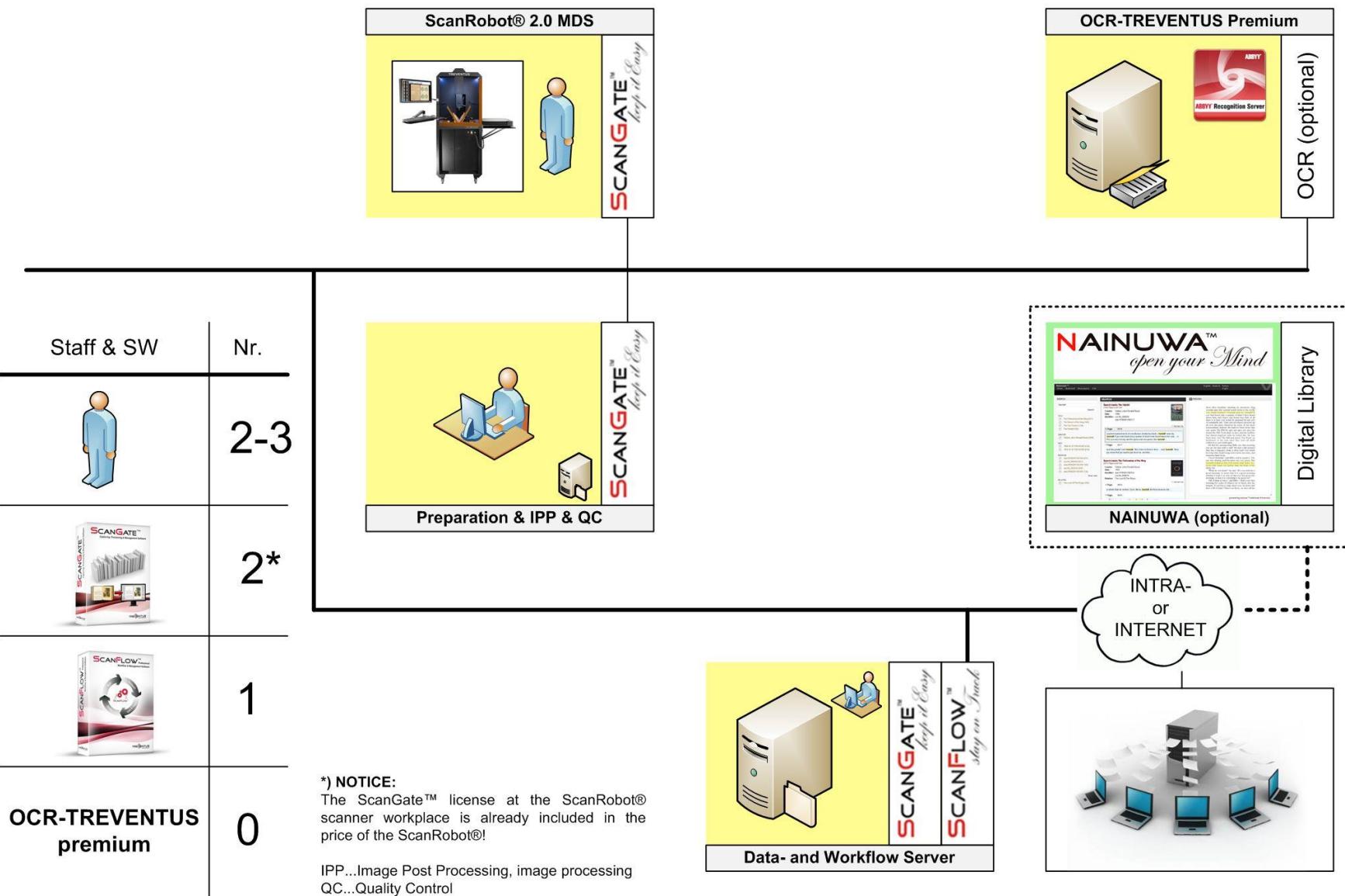
# ScanFlow® Workflows

The concept about ScanRobot® + Workflow solution

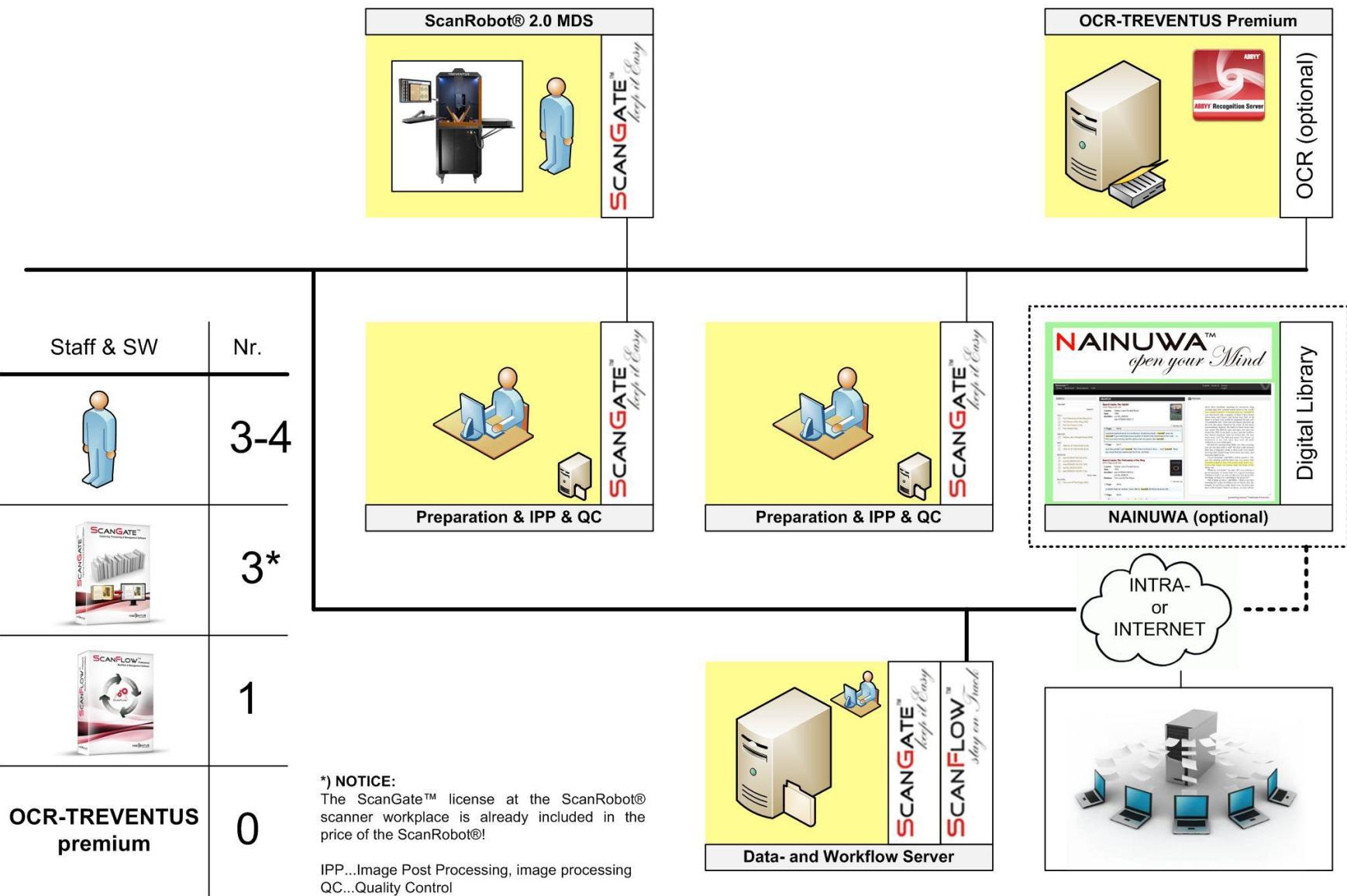
# General overview: Digitization Center Project – Infrastructure A (SF0)

		<b>ScanRobot® 2.0 MDS</b>	<b>OCR-TREVENTUS Premium</b>
			
<b>Staff &amp; SW</b>	<b>Nr.</b>		
	1		
	0*		
	0		
<b>OCR-TREVENTUS premium</b>	<b>0</b>		
<b>*) NOTICE:</b> The ScanGate™ license at the ScanRobot® scanner workplace is already included in the price of the ScanRobot®!		 <b>NAINUWA</b> open your Mind	
IPP...Image Post Processing, image processing QC...Quality Control		 INTRA- or INTERNET 	

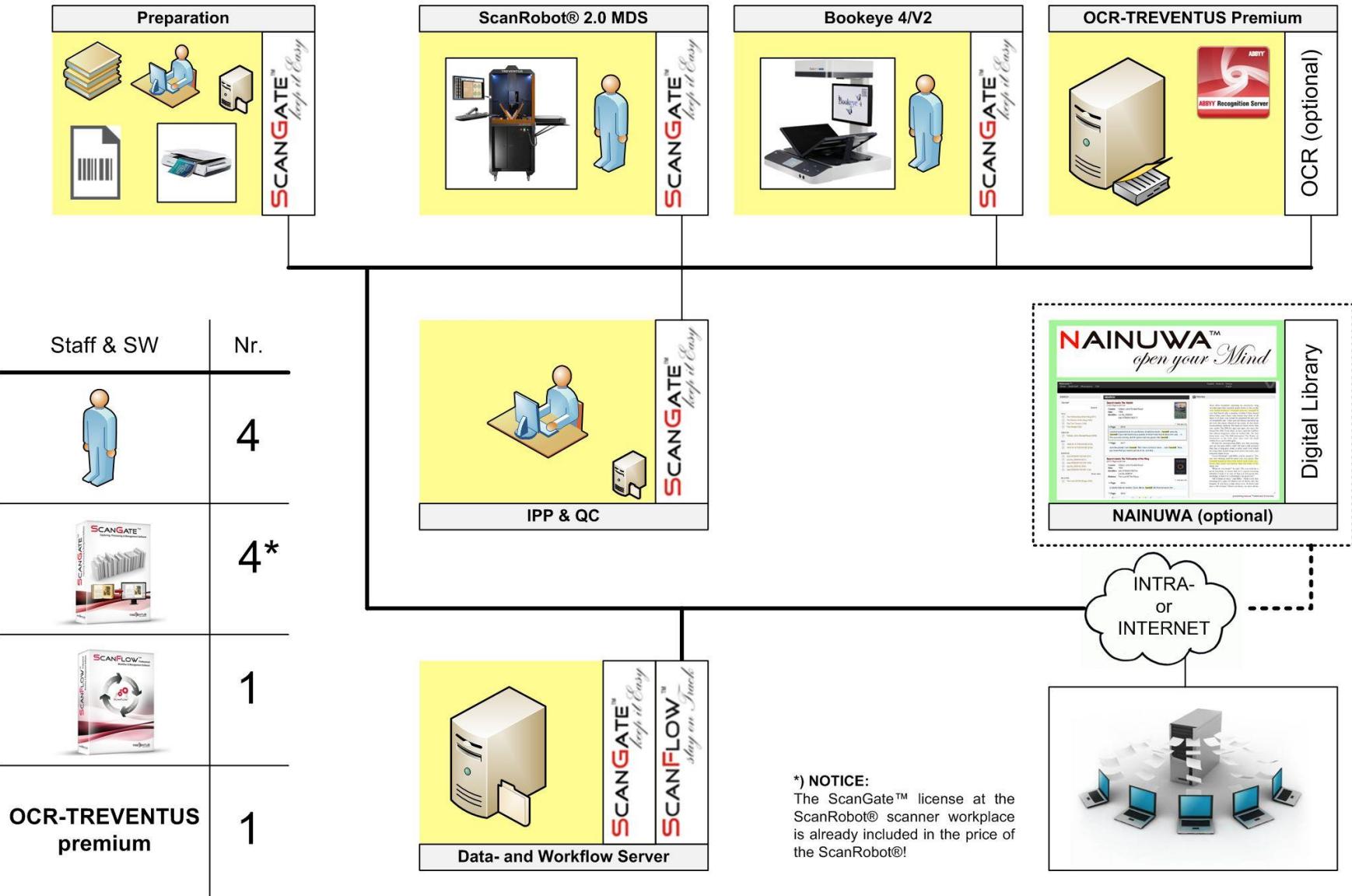
# General overview: Digitization Center Project – Infrastructure B (SF2)



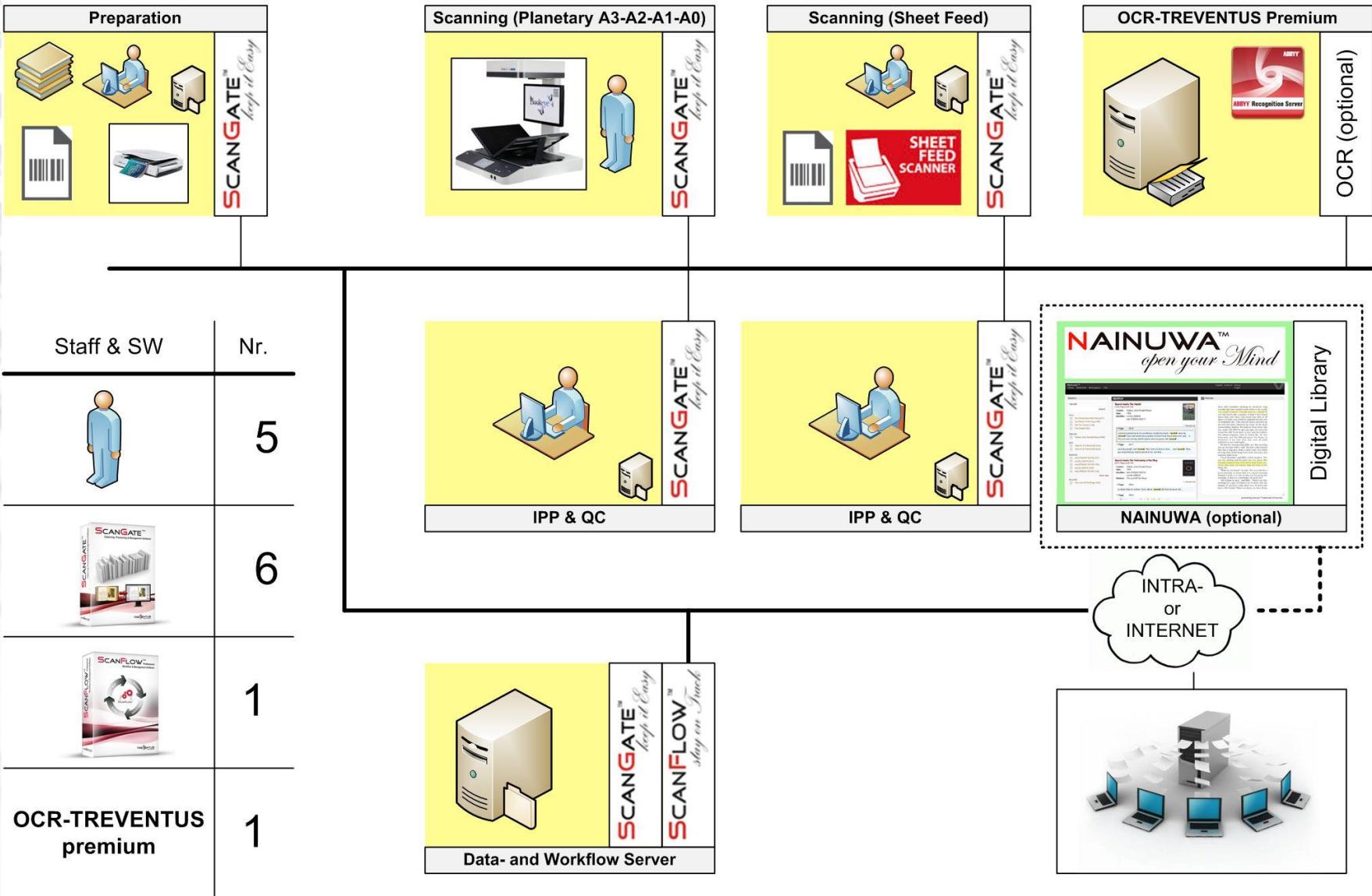
# General overview: Digitization Center Project – Infrastructure B (SF3)



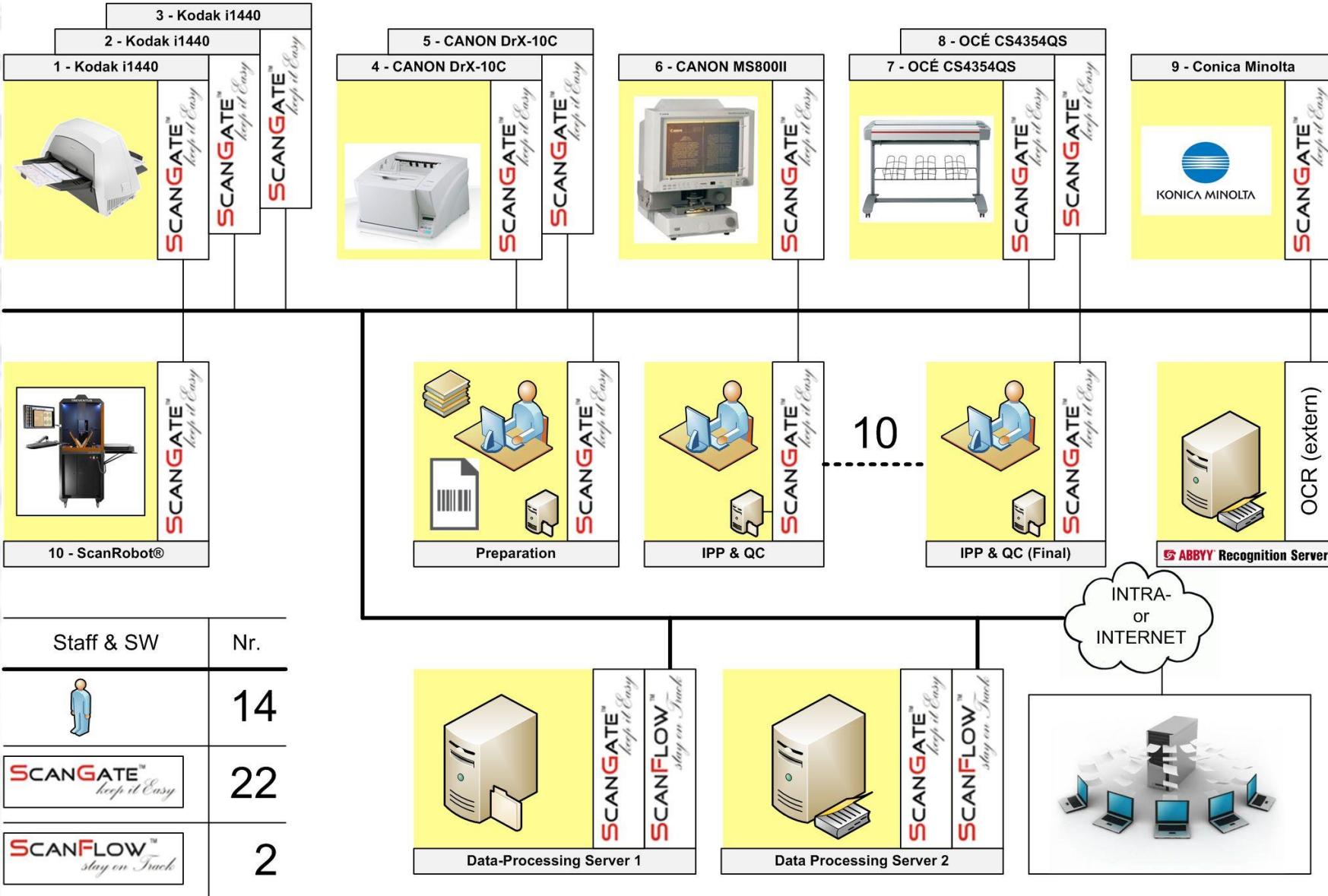
# General overview: Digitization Center Project – Infrastructure C (SF4)



# General overview: Digitization Center Project – Infrastructure D (SF6)



# General overview: Digitization Project (SF22)



Specify

scan

refine

enrich

workflow

access



# NAINUWA™ DIGITAL LIBRARY SYSTEM

The homepage of the NAINUWA Digital Library System. It features a dark blue background with a scenic sunset over mountains. Key statistics are displayed: 6,025 VOLUMES INDEXED, 98,426 PAGES INDEXED, and 349 KILOMETERS OF SEARCHABLE TEXT. Navigation links include LOGIN, REGISTER, and ENGLISH language selection. A search bar with an Advanced Search option is also present.

- FAST search&find
- On every browser
- Responsive design
- Gorgeous zooming
- Highlighting & lot more

A collage of screenshots and diagrams illustrating the NAINUWA system. It includes:

- A vertical list of search filters: Collection, Type, Creator, Subject, Publisher, Contributor, Date, Type, Identifier, Language, Coverage, Coverage Spatial, Google Info Place, Rights.
- Browser icons for Chrome, Firefox, Internet Explorer, and Opera.
- A mobile device displaying a responsive version of the library's website.
- A sidebar titled "MY BOOKSHELF" with sections for "MY BOOKMARKED PAGES", "MY BOOKMARKED VOLUMES", and "MY CUSTOM VOLUMES".
- Language icons for Français, Español, English, Italiano, and Deutsch.
- A detailed view of a book entry for "Goody Two Shoes" with social media sharing options (Facebook, Twitter, Google+).
- A process flow diagram showing a document being converted from an IMAGE to a SEARCHABLE format, with TEXT EXTRACTION as the final step.
- A close-up of a scanned page from a book with a magnifying glass icon highlighting text.
- A snippet of text from a scanned book page with a blue circle around a portion of the text.

# How to turn books → digital libraries?

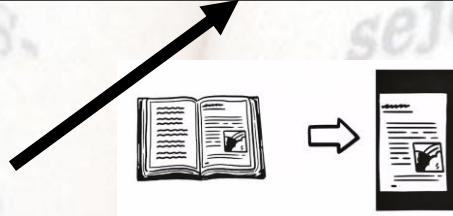
---



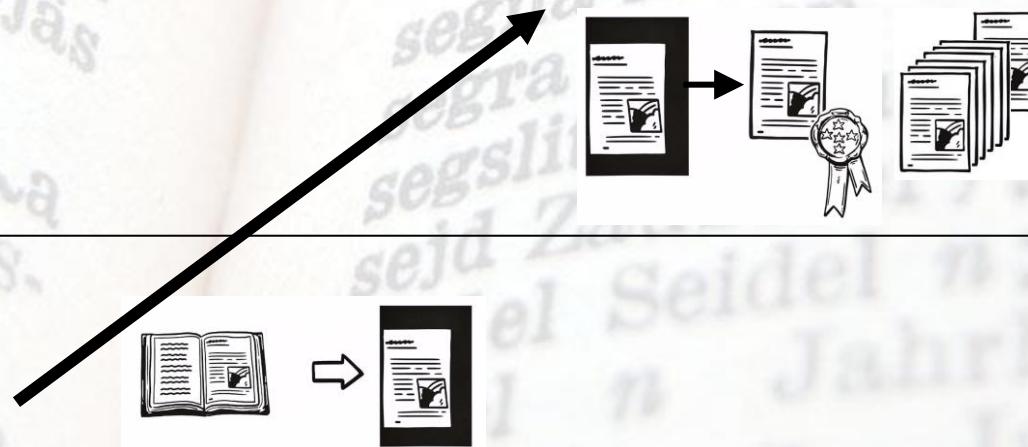
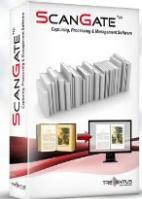
# From document → digital library



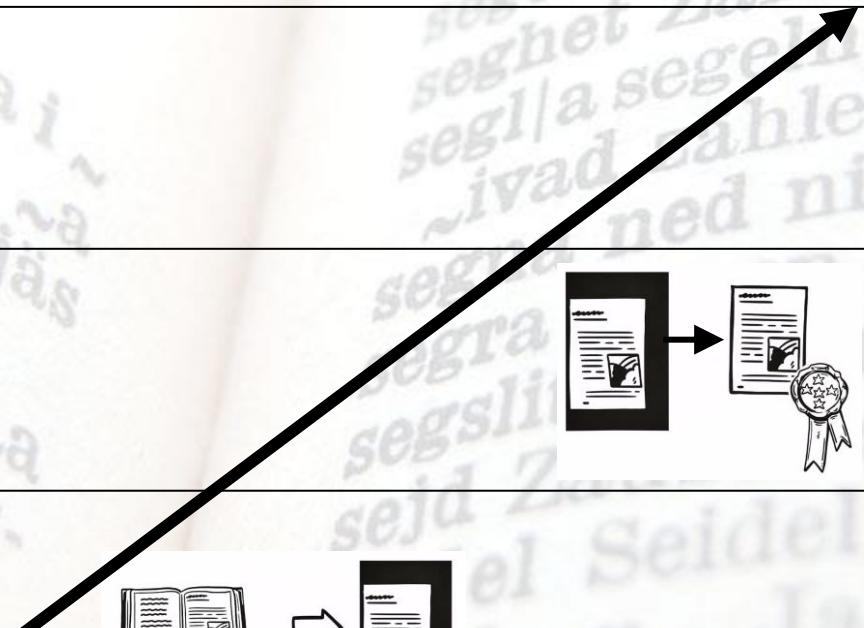
# From document → digital library



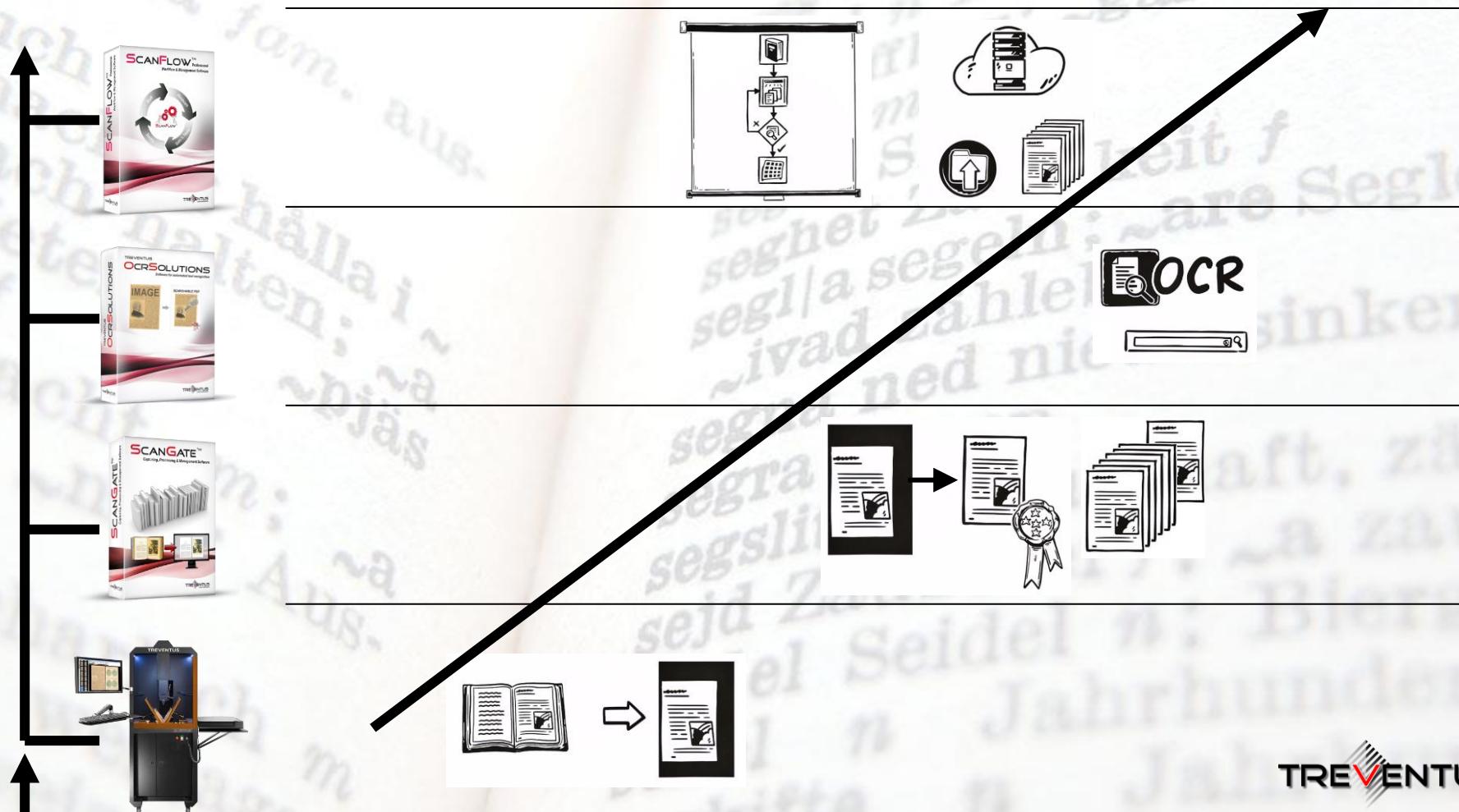
# From document → digital library



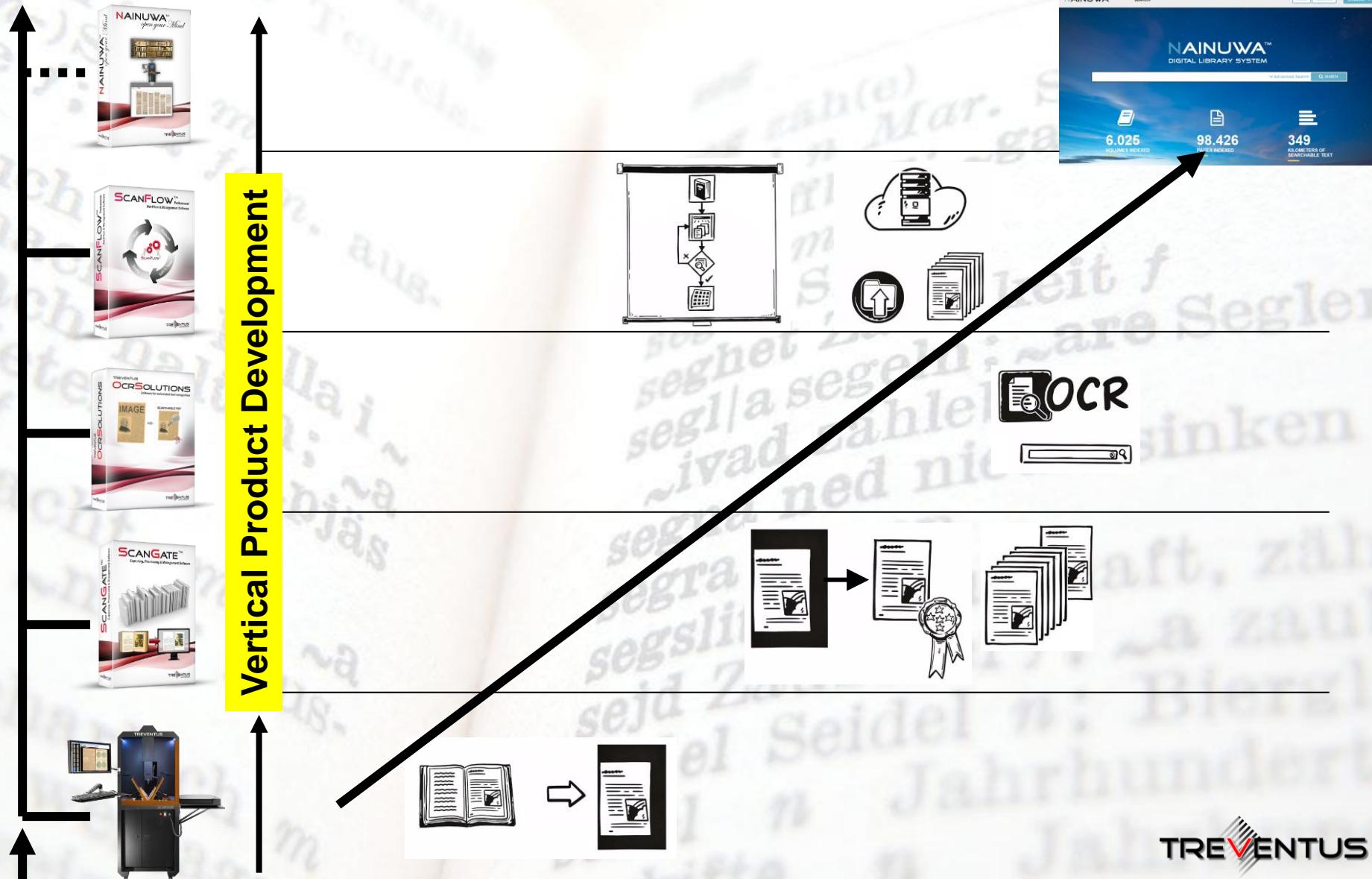
# From document → digital library



# From document → digital library



# From document → digital library





# TREVENTUS

---

## History & Awards

# From the idea...

---



Prof. Zagler in 2003 had an idea!

...but put them in a drawer!

# ...to the realization: 2004 – KickOFF

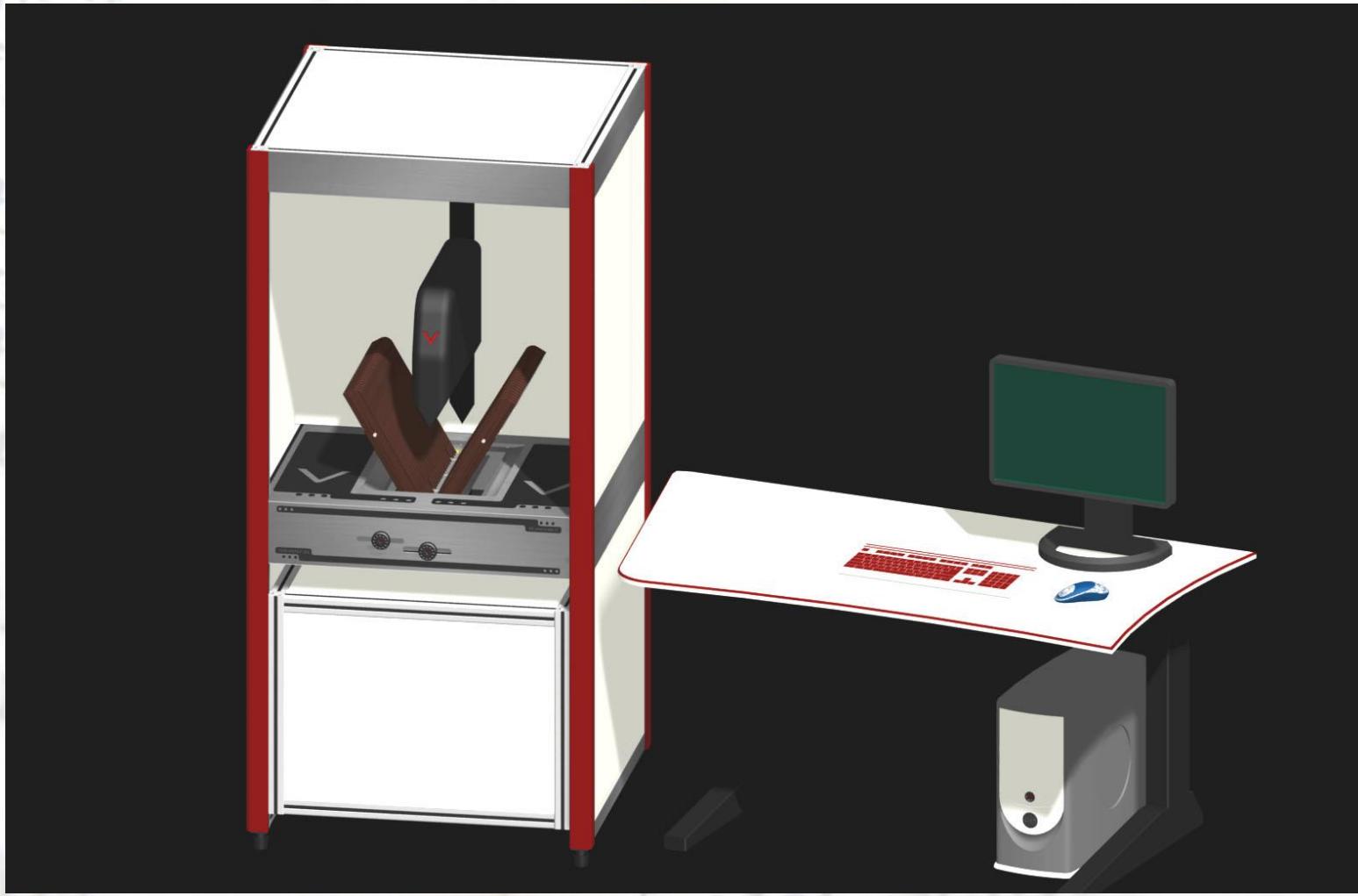
---

- Assessment
  - Markus Barth
  - Christoph Bauer
  - Stephan Tratter



# The Scanner: a 3D-Model

---



# History: Treventus 2004 - 2018

---

- 2004:** KickOFF to the pilot project „**ScanRobot®**“
- 2005:** proof of concept / first prototype & Patent
- 2006:** TU Wien SpinOFF: **TREVENTUS Mechatronics**
- 2007:** „**ICT Grand Prize**“ of the EU (CeBIT)  
200.000 € prize money
- 2007:** **ScanRobot® SR30x & ScanGate™**  
+ strategic partnerships
- 2009:** **ScanFlow™ / OCR Solutions**
- 2010:** launch of **ScanRobot® 2.0 MDS**
- 2016:** launch of **Nainuwa™**
- 2022:** > **ScanRobots®** in over 65 countries worldwide  
> 100 digitization solutions → **ScanFlow™**



# 2007: European ICT Innovation Prize



ENTUS

# Awards

2008, March	 Vienna economy furtherance fonds	<b>Best young entrepreneur</b>  TREVENTUS is awarded in the category "best young entrepreneur" for the automatic bookscanner <b>ScanRobot®</b> .
2007, March		<b>European ICT grand prize</b>  TREVENTUS wins EU's "grand prize" for ICT (endowed: EUR 200,000)! The automatic Bookscanner <b>ScanRobot®</b> convinced the jury!  <i>"Information and communications technology opens up a world of new possibilities – new ways of working and new ways of living!"</i> said Information Society and Media Commissioner <b>Viviane Reding</b> as she awarded the prizes in Hannover on the <b>CeBIT 2007</b> .
2006, June		<b>Theodor Kery – Innovation prize</b>  TREVENTUS wins the " <b>Theodor Kery - Innovation prize</b> ", organized by the Burgenlandstiftung.
2005, December		<b>Genius Innovation Award</b>  TREVENTUS wins the Austrian <b>Genius Award 2005</b> .

# about Prof. Zagler

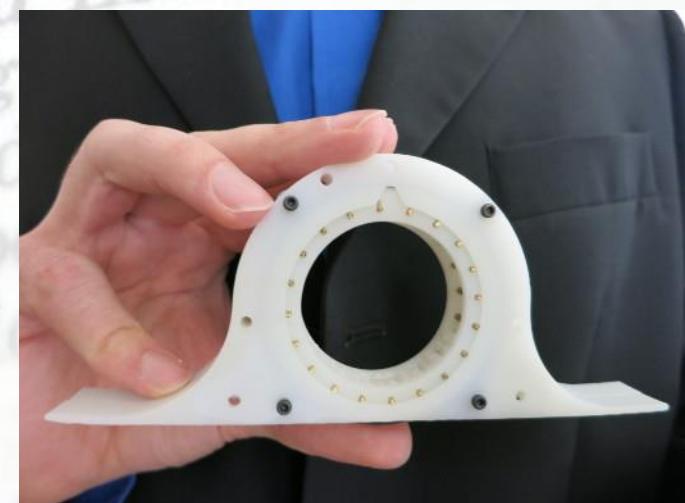
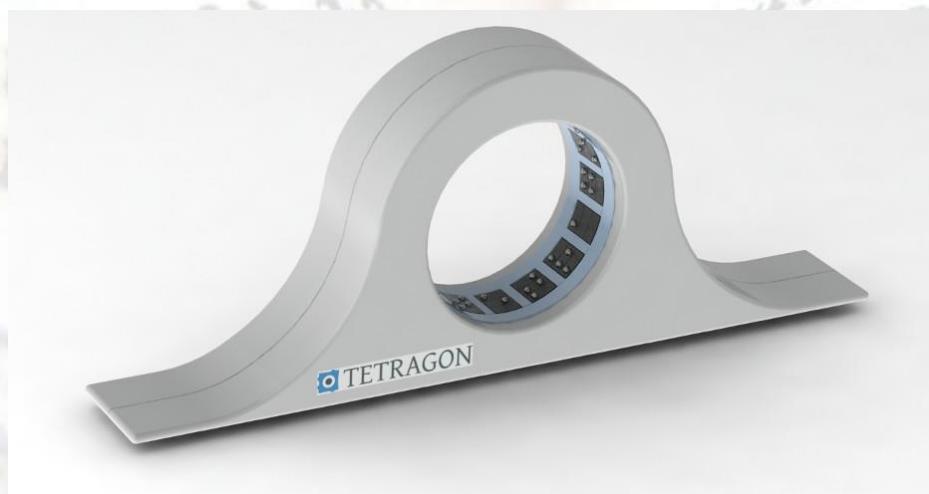
---



- Professor & Entrepreneur
- Vienna University of Technology
  - AAT - Center for applied assistive technologies
- Treventus (2004)
- Others...
- Tetragon (2016) → Braille Ring



 TETRAGON





# TREVENTUS

---

status quo – market situation

# 8th IT-days in Stuttgart 2009

---



At the 8th IT-meeting in Stuttgart we have made a personal demo for:

- Chancellor of Germany Mrs. **Angela Merkel**
- Federal Minister for Economics and Technology Mr. **Brüderle**
- Director of the German National Library Mr. **Altenhöner**

# ALA 2010: Washington DC



Library of Congress: VIP guide

# Extract of references

---

- Libraries
- Archives
- Industry
- Bank & financial institutions
- Editorial & Printing houses
- Universities
- Research Centres
- Digi Service Provider
- Military & Intelligence Agencies
- Ministries
- Parliament Libraries
- Courts (high court)
- Hospitals
- Patent offices

# Installations

---

- Installations from
  - Deep Siberia – Brasil Jungle
- A project in Middle East since 4 years
  - 29 ScanRobots
  - 800.000 pages/day



Biggest solution of automatic  
bookscanners worldwide!!

# References: ScanRobot® and ScanFlow®



國家圖書館  
National Central Library



UNSW  
SYDNEY



UNIVERSITÉ DU  
LUXEMBOURG



Australian  
National  
University



UMEÅ UNIVERSITY



Canon



The University of Dublin



National Library  
of Sweden



UNIVERSITÀ DI PISA

Εθνική Βιβλιοθήκη της Ελλάδος  
National Library of Greece

...and many more...

## Lab 2.2: digitalizzazione delle raccolte (digitalizzazione sostenibile)

- **Abstract:** La digitalizzazione è preservazione del patrimonio della Biblioteca dall'eccessiva manipolazione e consultazione del manufatto.  
Una digitalizzazione consapevole mette in campo le nuove tecnologie per fronteggiare il deterioramento dovuto al tempo ed è fondamentale adottare strumenti consoni al patrimonio.  
Tutto in campo informatico muta velocemente e con questo i costi di hardware e software si sono moltiplicati. Tuttavia esistono soluzioni open source che consentono di utilizzare diverse fonti di dati e diversi strumenti modalità e qualità di visualizzazione.  
L'adozione di tecnologie (come software di catalogazione o scanner) sono direttamente legate alla dimensione del patrimonio di cui la Biblioteca dispone.
- **SubArgomento:** La valorizzazione, cioè la presentazione dei dati e la loro messa a disposizione, viene dominata dal principio che la biblioteca digitale, e tutto ciò che si realizza nel suo ambito, deve essere user friendly, anzi proprio user driven per cui i servizi e la loro presentazione ed offerta devono essere pensati dal punto di vista dell'utente.  
Aspetto tecnico: La tecnologia IIIF, per la visualizzazione ed elaborazione dei dati, offre non solo nuove possibilità di visualizzazione, ma anche nuove forme di elaborazioni, per esempio la possibilità di formulare delle glosse/commenti e etc.

# Sostenibilità?

---

# Intro - Workflows

# Difference: Digitization – Mass Digitization

---

## mass-digitization



workflow  
logistic  
organization



scanning  
image treatment  
text recognition  
meta data



# What means “MASS” in “mass digitization”?

## digitization

1 book	→ 400 pages	8 GByte
10 books	→ 4.000 pages	80 GByte
100 books	→ 40.000 pages	800 GByte
1.000 books	→ 400.000 pages	8 TByte
10.000 books	→ 4 Mio pages	80 TByte

## mass digitization

Scale from “digitization” to “mass digitization” means  
“the digital industrial revolution”!

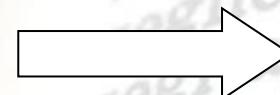
- split the whole work in small pieces (tasks)
- and organize them (→ Taylorism).

# Introduction I – The idea

---



Physical  
Material



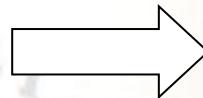
Digital  
Library

# Introduction II – Scanning

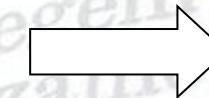
---



Physical  
Material



scanning



Digital  
Library

# Introduction III – Digitizing



inhomogenous

exceptions

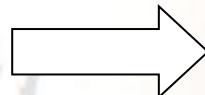
staff

preparation

quality control

scanning

Physical  
Material



digitizing

Digital  
Library

exceptions

output formats

storage

OCR

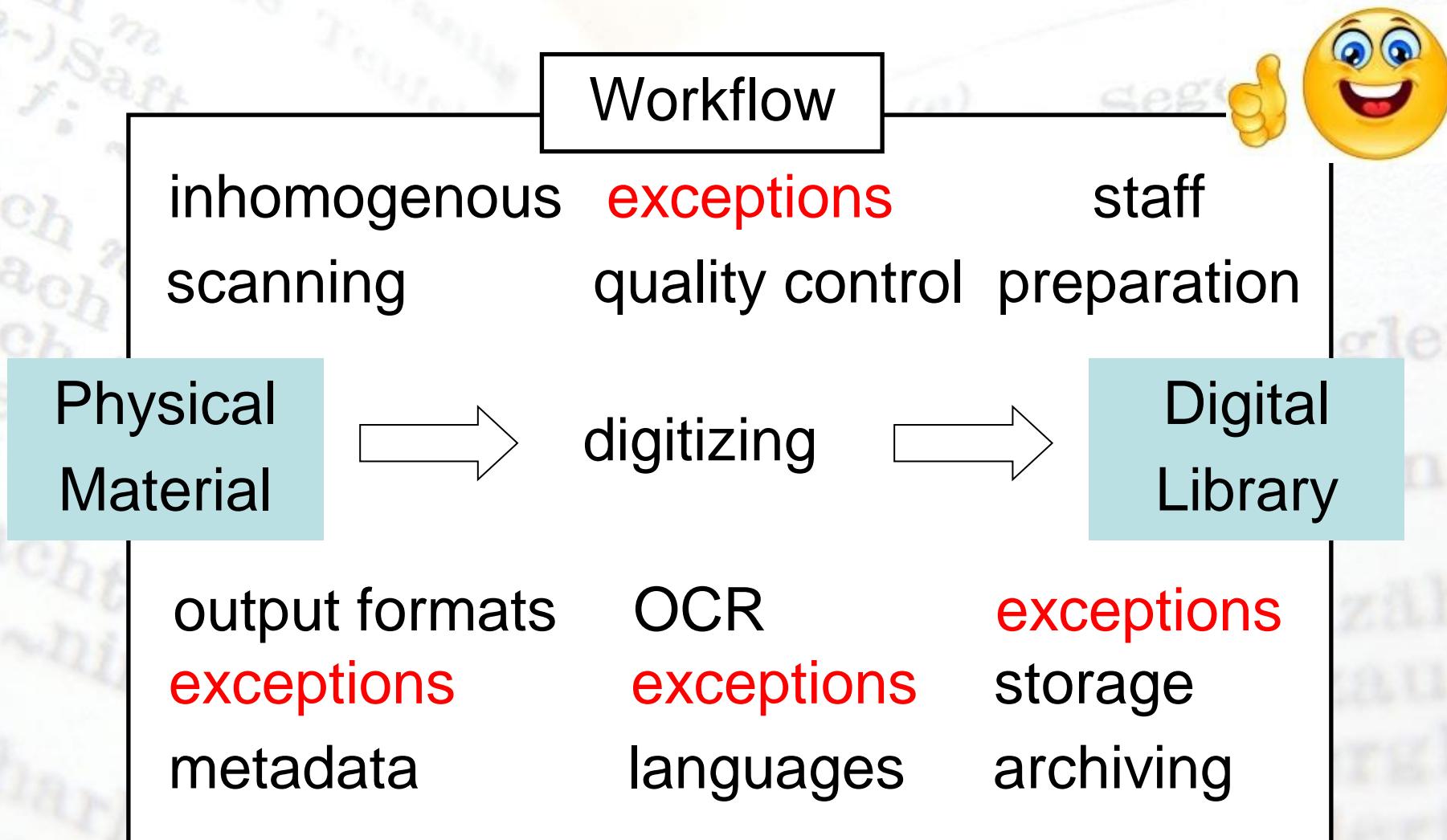
metadata

archiving

languages

exceptions

# Introduction IV – Workflow



# Argomento: Digitalizzazione Sostenibile

---

## Sostenibilità?

# STEPS – DIGITIZATION

Example

Tasks													
Preperation	Automatic Scanning	Manual Scanning	IPP & QC	Enrichment / OCR	Final QC / Storage								
Working Steps	<u>Logistic &amp; Preparation</u> <ul style="list-style-type: none"> <li>Books from shelf</li> <li>Prepare books</li> <li>Create Jobs (in ScanGate™)</li> <li>Insert Metadata</li> <li>Process-Slip Barcode</li> <li>Books back 2 shelf</li> </ul>	<u>Scanning (Automatic)</u> <ul style="list-style-type: none"> <li>Open Scan-Job</li> <li>Scan Book</li> <li>QC: Completeness</li> <li>Re-Scan Book</li> </ul>	<u>Scanning (Manual)</u> <ul style="list-style-type: none"> <li>Open Scan-Job</li> <li>Book - Page - Map</li> <li>QC: Completeness</li> <li>Re-Scan Book</li> </ul>	<u>IPP-Quality Control (QC)</u> <ul style="list-style-type: none"> <li>Apply Functions: <ul style="list-style-type: none"> <li>De-Skew</li> <li>Cropping</li> <li>Homogenize</li> <li>Binarize</li> <li>Import Files</li> <li>Export to formats</li> <li>Etc.</li> </ul> </li> <li>Control Results / Preview / Check</li> <li>Return Books back to preparation desk</li> </ul>	<u>Enrichment / OCR</u> <ul style="list-style-type: none"> <li>OCR <ul style="list-style-type: none"> <li>188 Languages</li> <li>Export to PDF</li> <li>Export to Word</li> <li>Verification*</li> <li>Processing</li> </ul> </li> </ul> <p>→ Add/Change Metadata</p>	<u>Final QC / Storage</u> <ul style="list-style-type: none"> <li>Final Check</li> <li>Store Derivates</li> <li>Store Masters</li> </ul>							
FILES	<ul style="list-style-type: none"> <li>Job Files (Scanjobs)</li> <li>Images: Master Files</li> </ul>	<ul style="list-style-type: none"> <li>Job Files (Scanjobs)</li> <li>Images: Master Files</li> </ul>	<ul style="list-style-type: none"> <li>Job Files (Scanjobs)</li> <li>Images: Master Files</li> </ul>	<ul style="list-style-type: none"> <li>Job Files (Scanjobs)</li> <li>Images: Master Files</li> <li>Images: Work Files</li> </ul>	<ul style="list-style-type: none"> <li>Job Files (Scanjobs)</li> <li>Images: Master Files</li> <li>Images: Work Files</li> <li>Metadata</li> <li>OCR Results</li> </ul>	<ul style="list-style-type: none"> <li>Job Files (Scanjobs)</li> <li>Images: Master Files</li> <li>Images: Derivates</li> <li>Metadata</li> <li>OCR Results</li> </ul>							

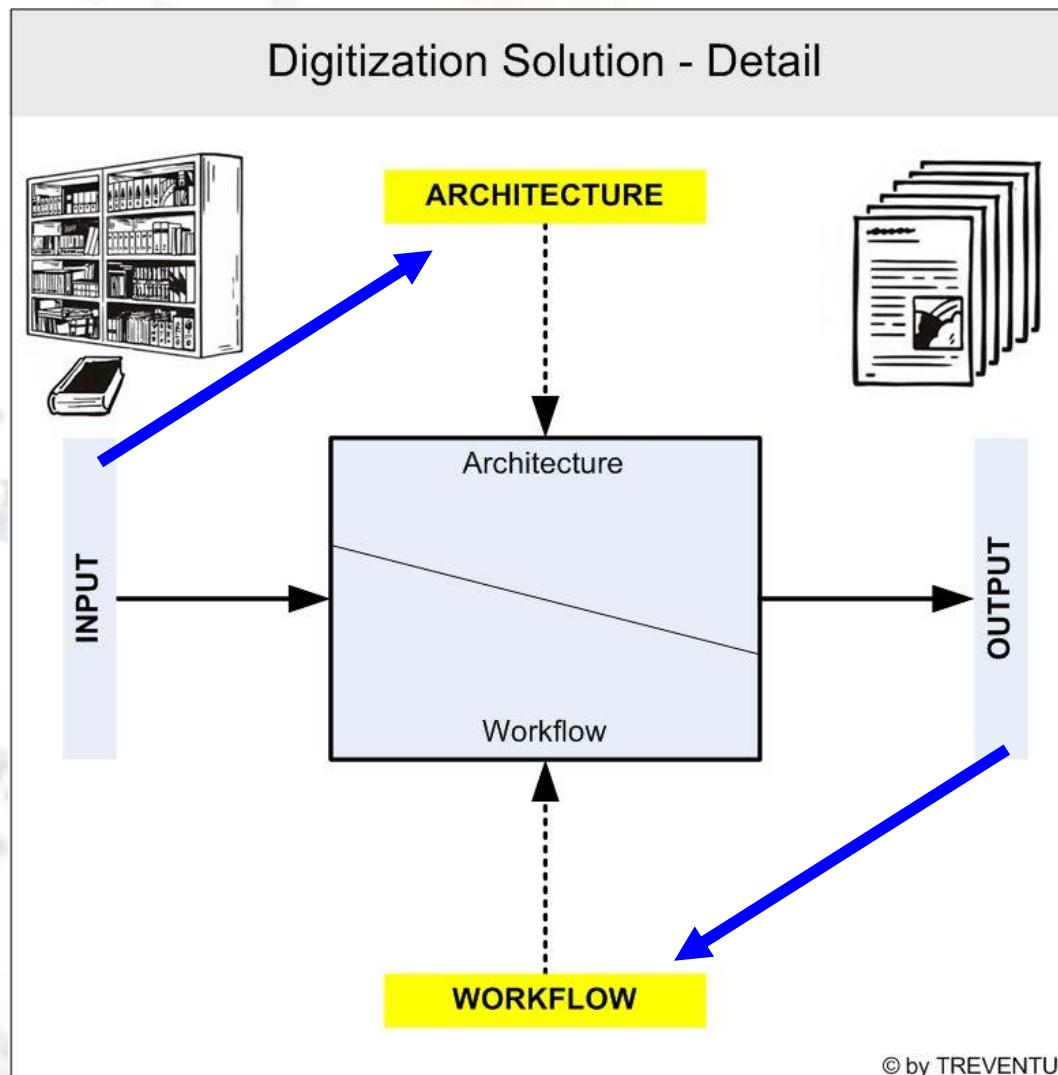
# Paragone

---

- Dinner for 2 → 2
- Friday Night Party → 40
- Wedding → 500

# Intro

# Look at A Digitization Project

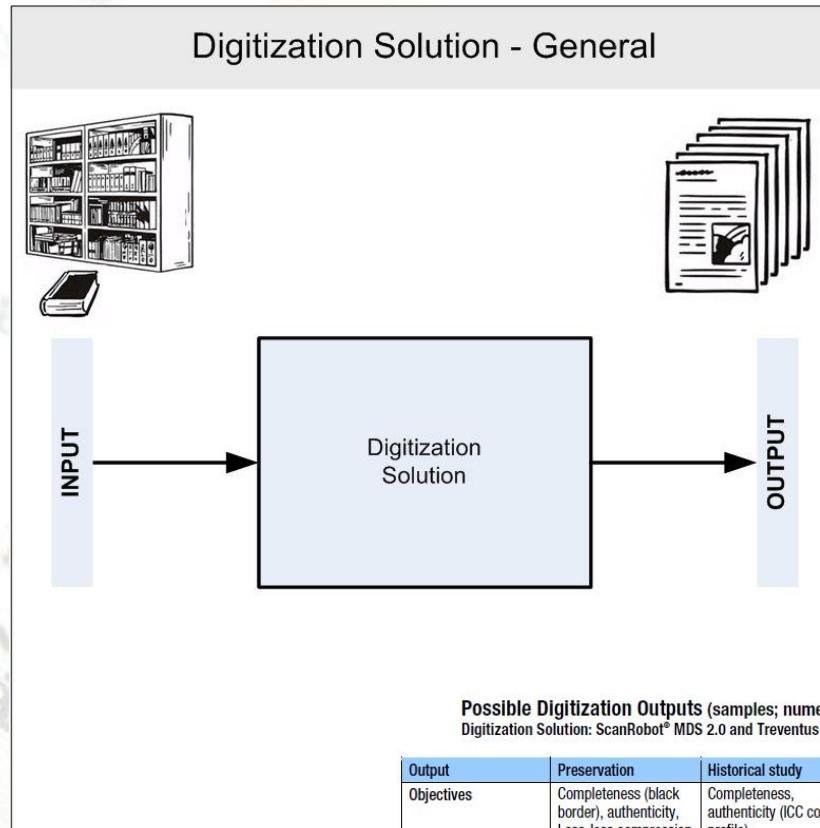


# How do we build up solutions?

Which GOAL?



(→ Why digitization?)



Which Output?

Which documents?



Output	Preservation	Historical study	Information retrieval A	Information retrieval B	Information retrieval C
Objectives	Completeness (black border), authenticity, Loss-less compression	Completeness, authenticity (ICC color profile)	Good reading experience (digital); centred pages, contrast	Very good reading experience, high contrast, Low file size	Very good reading experience, high contrast, Images not changed, Image database possible
Sample Image					

# GOAL: Why Digitization is done?

---

- Preservation
  - 4 elements (water, fire, air, earth (quake))
  - War
  - Stolen items (burglary)
  - Material itself (iron ink, acid damaged paper...)
- Access & Research
  - + Images + Metadata + Full Text + Enrichments
- Share
- Re-Use in digital way
  - E.g. Publishers: Out of print books → Portfolio

JPG

TIFF



JPG

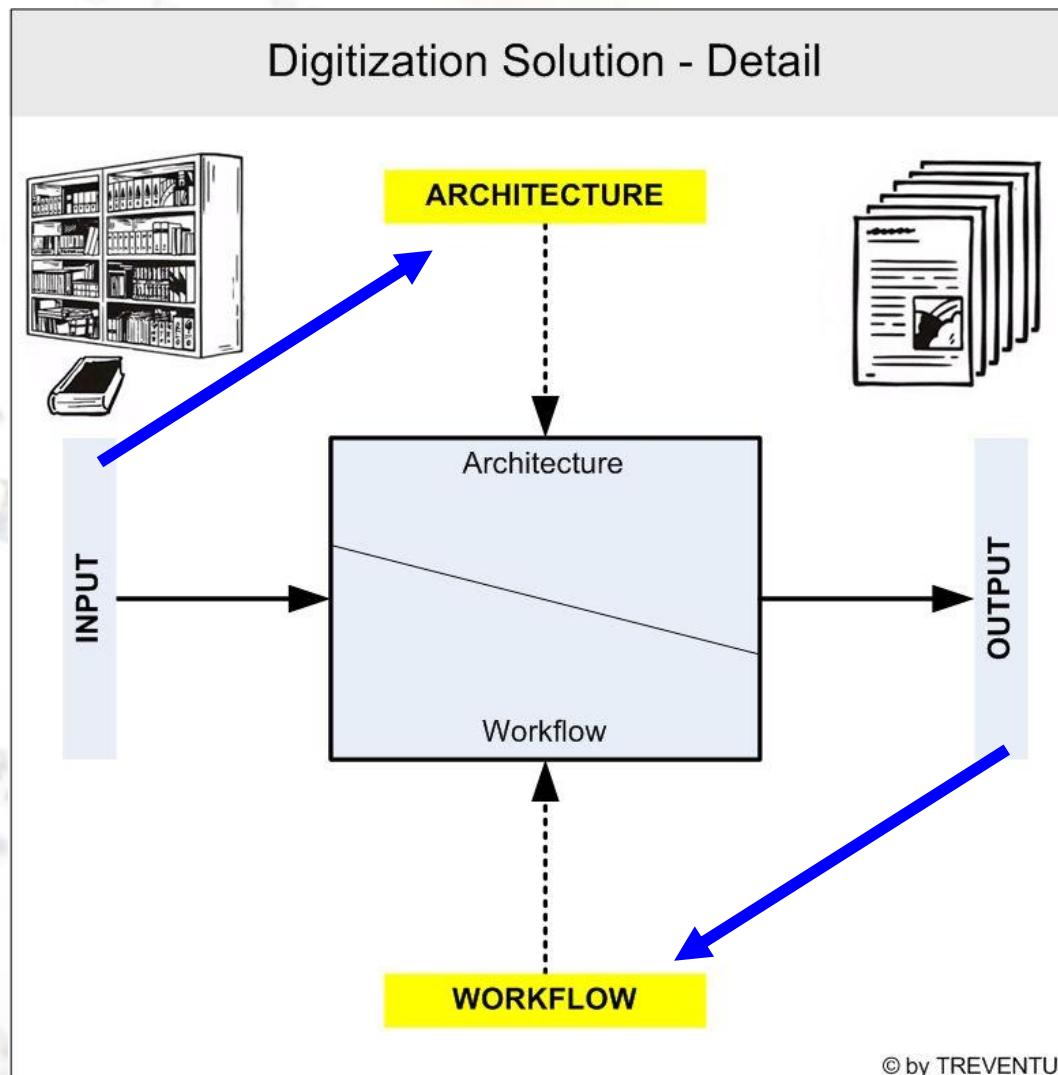


# Pianificare „Progetti di Digitalizzazione“

---

1. Obiettivo
2. Materiale
3. Risultato
4. Soluzione
5. Dati e spazio
6. Lavoro e team
7. Tracking e Reporting

# Look at A Digitization Project

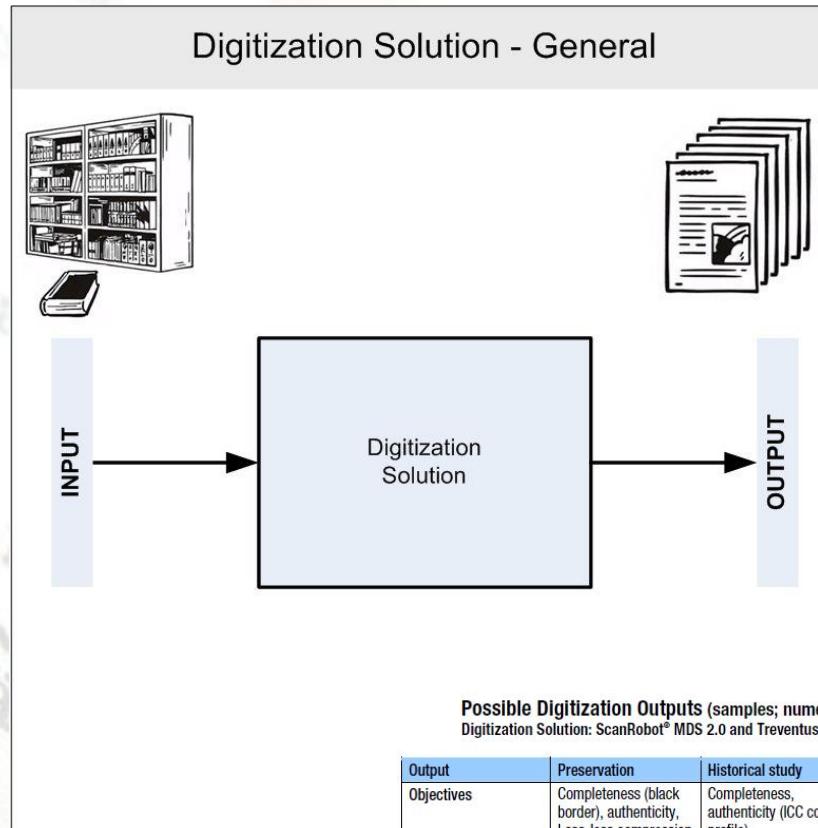


# How do we build up solutions?

Which GOAL?



(→ Why digitization?)



Which Output?

Which documents?

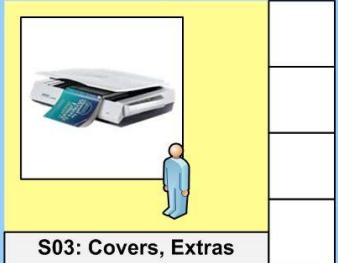
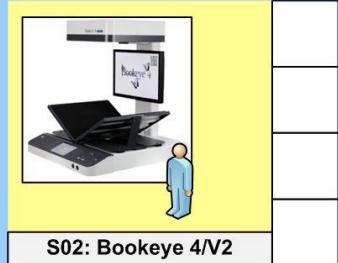
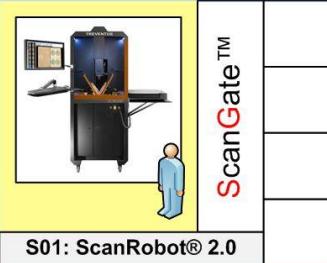


Output	Preservation	Historical study	Information retrieval A	Information retrieval B	Information retrieval C
Objectives	Completeness (black border), authenticity, Loss-less compression	Completeness, authenticity (ICC color profile)	Good reading experience (digital); centred pages, contrast	Very good reading experience, high contrast, Low file size	Very good reading experience, high contrast, Images not changed, Image database possible
Sample Image					

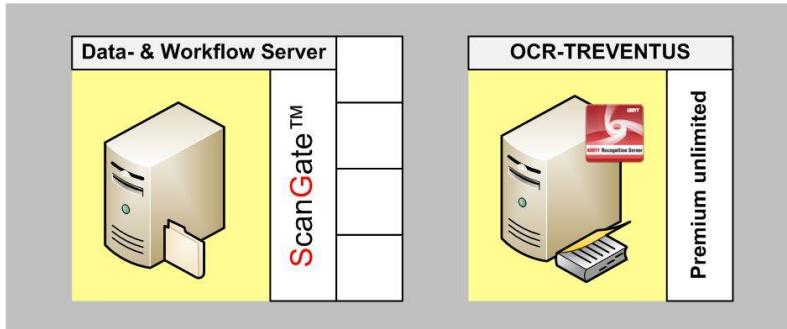
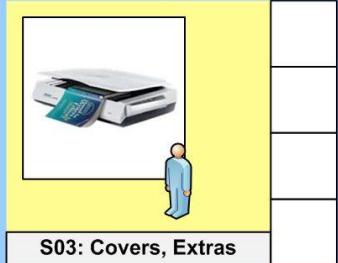
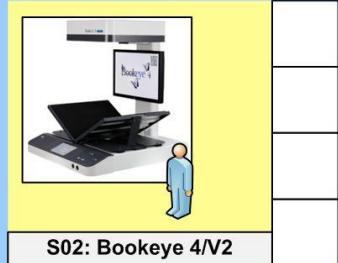
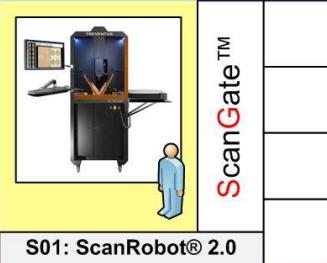
# 1.) Architettura

Pianificazione die posti di lavoro

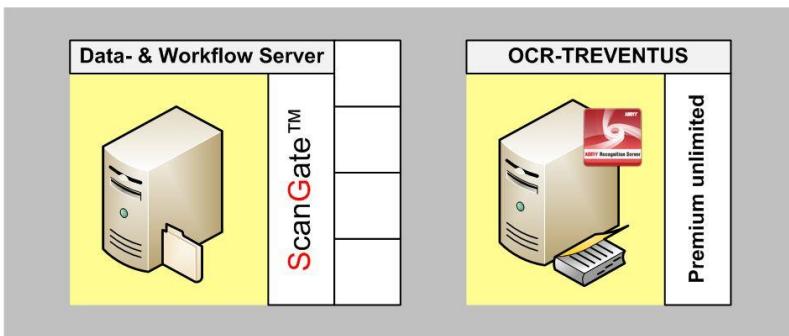
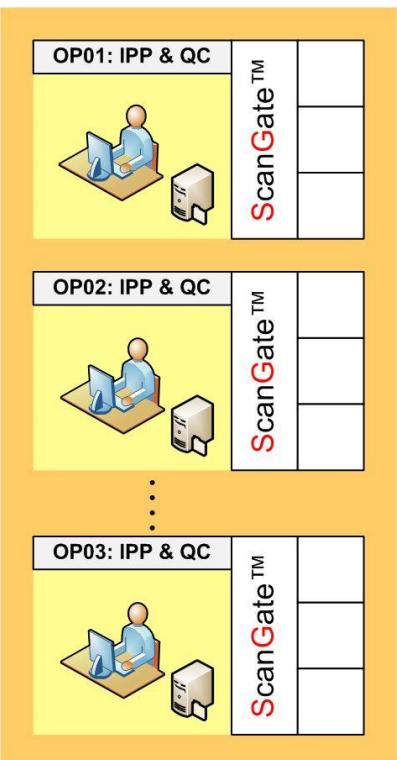
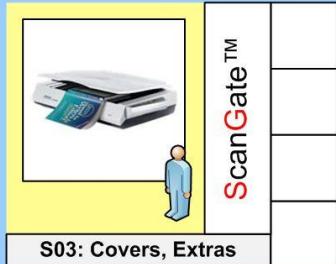
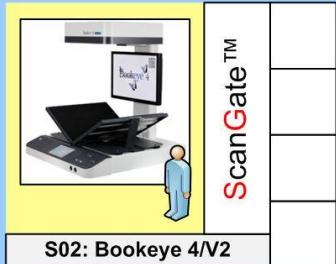
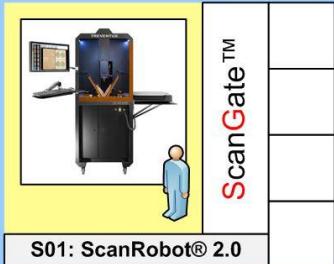
## General overview: Digitization Center Project – Infrastructure C (SF6)



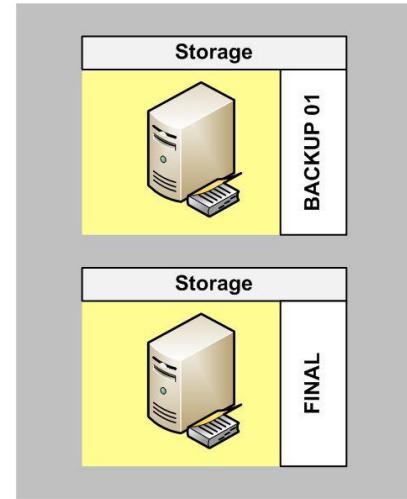
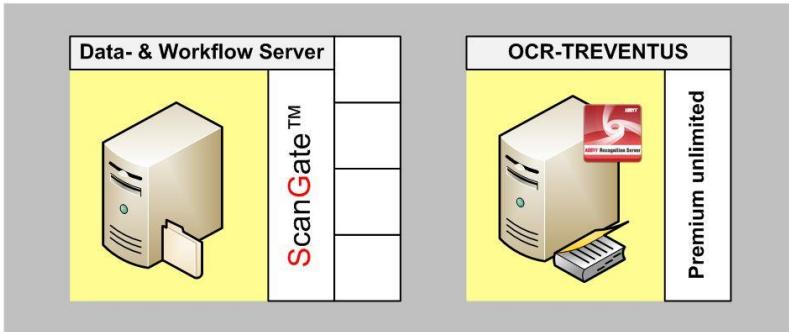
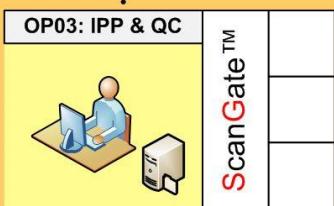
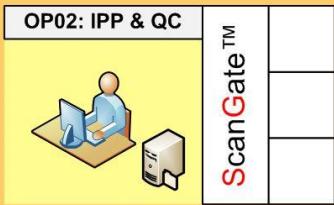
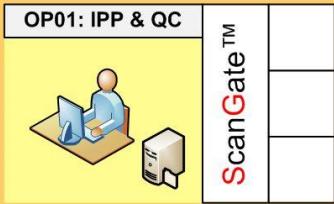
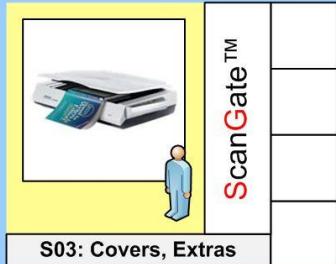
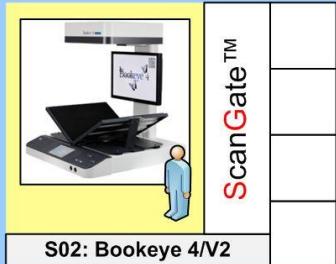
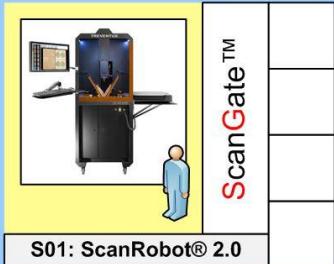
## General overview: Digitization Center Project – Infrastructure C (SF6)



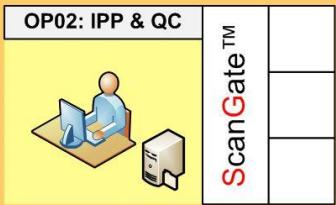
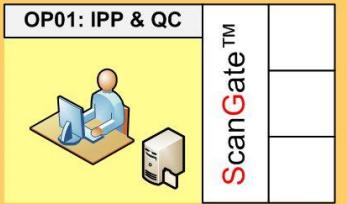
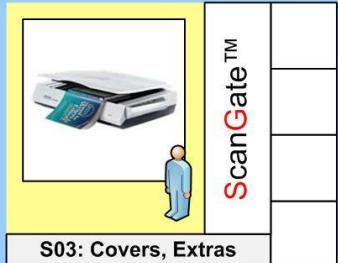
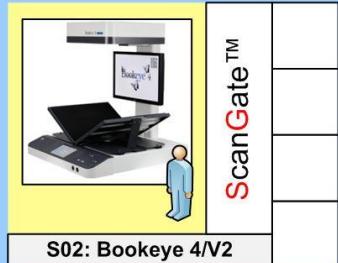
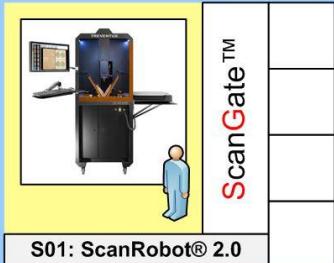
# General overview: Digitization Center Project – Infrastructure C (SF6)



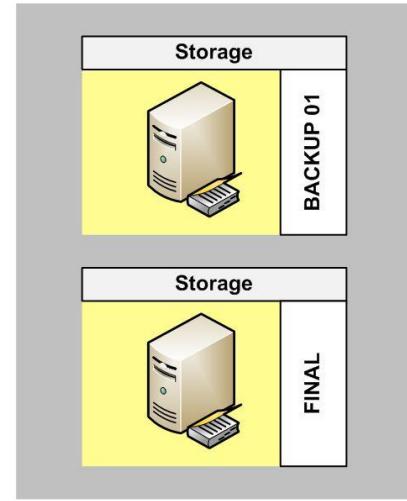
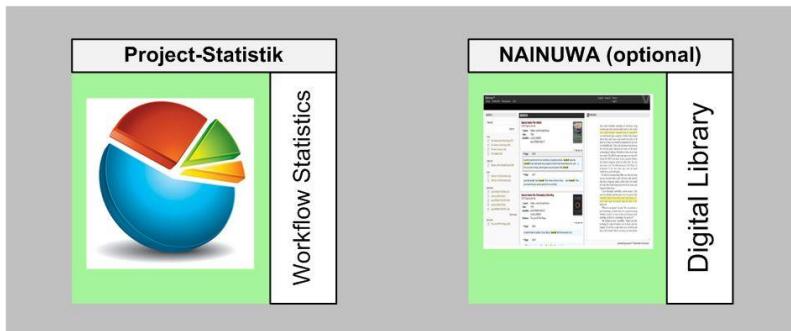
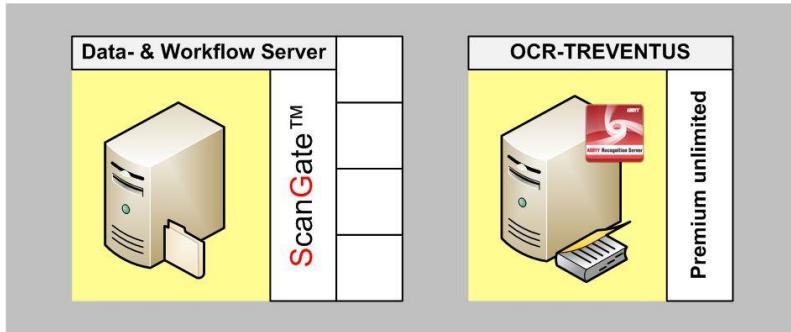
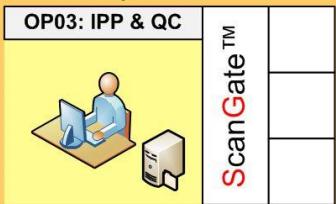
# General overview: Digitization Center Project – Infrastructure C (SF6)



## General overview: Digitization Center Project – Infrastructure C (SF6)

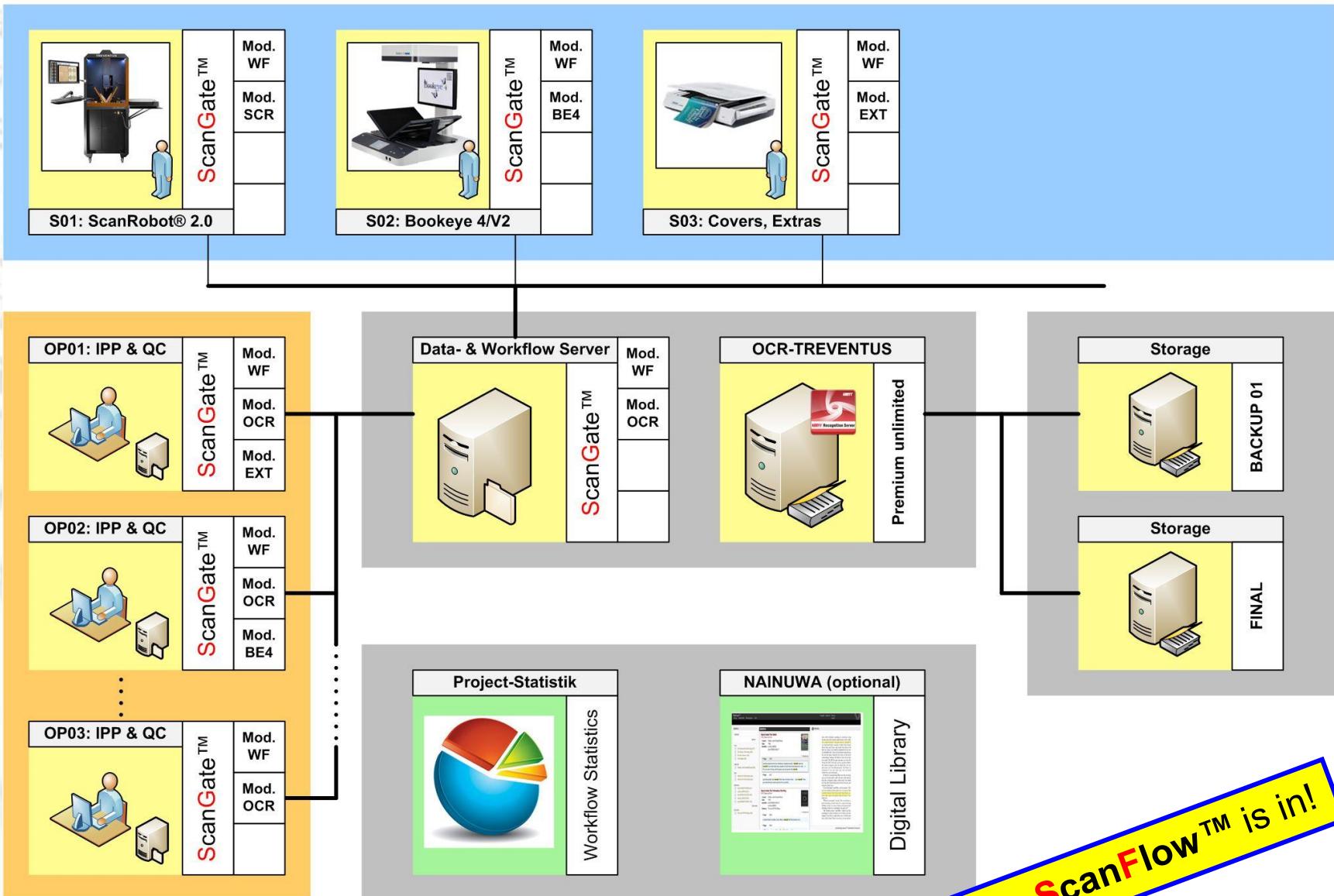


...

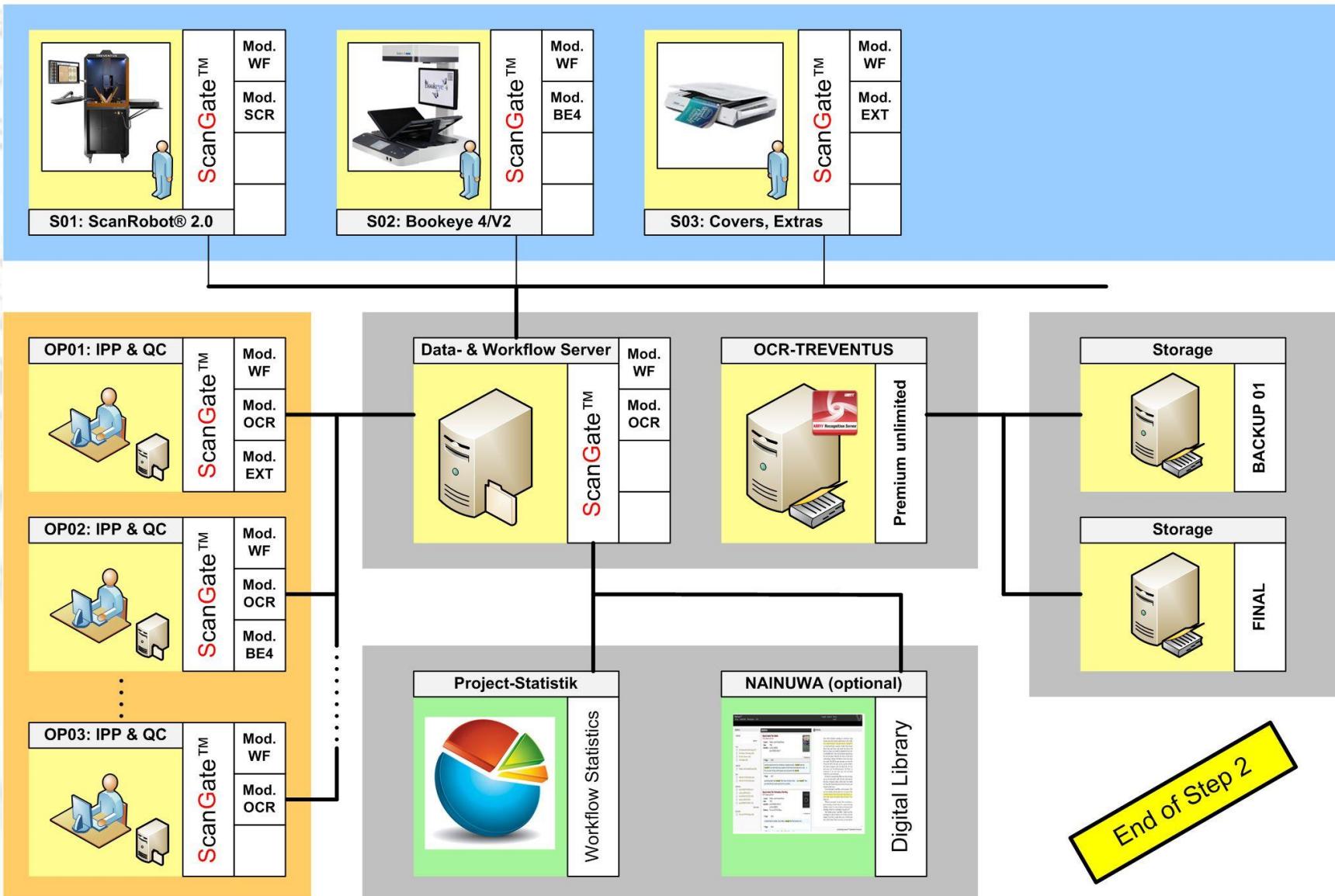


End of Step 1

# General overview: Digitization Center Project – Infrastructure C (SF6)



# General overview: Digitization Center Project – Infrastructure C (SF6)



# 2.) Output

Result – Layout – Formats

# Output Layout

## Possible Digitization Outputs (samples; numerous variations and combinations possible)

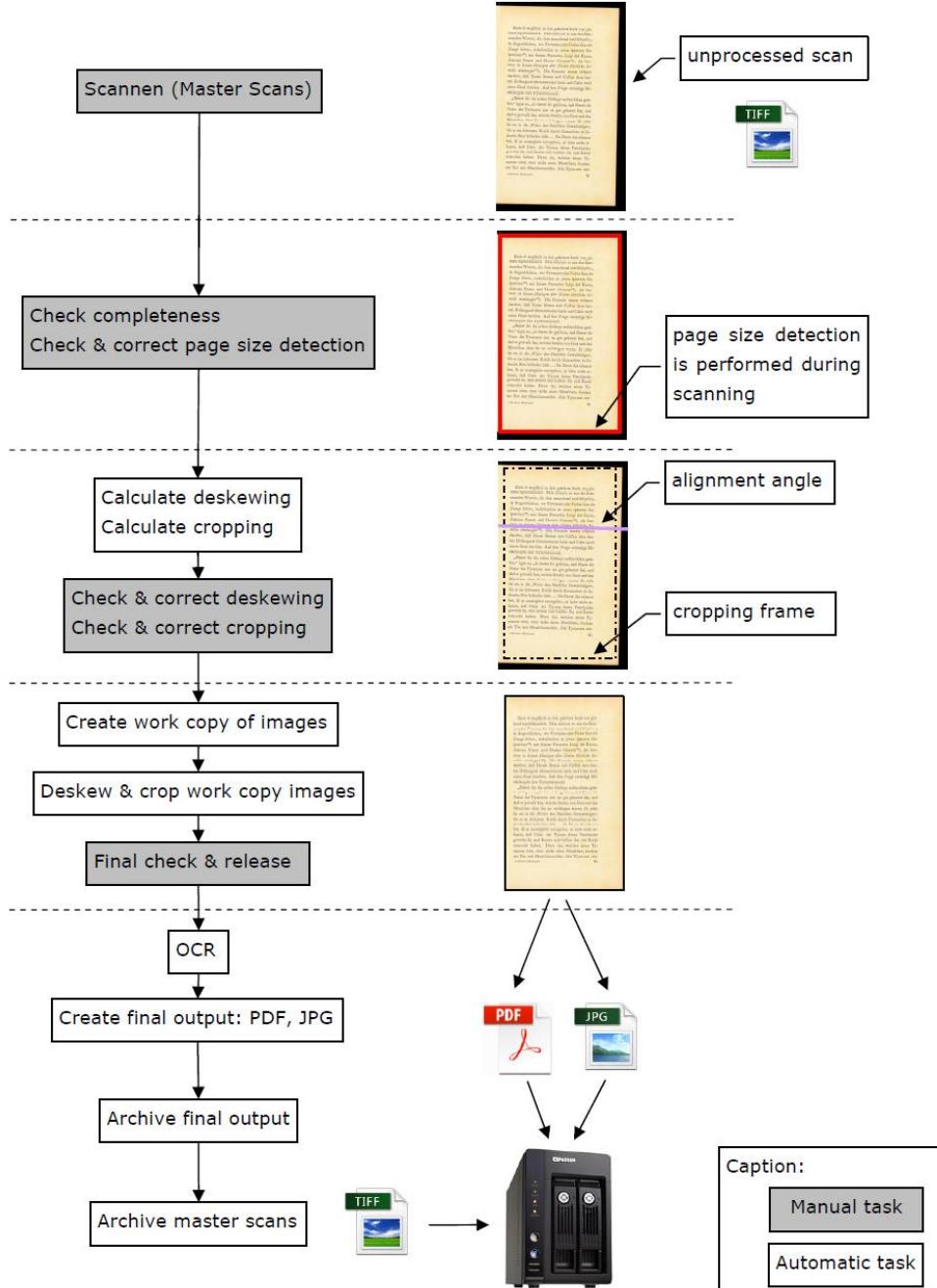
Digitization Solution: ScanRobot® MDS 2.0 and Treventus Workflow solution

Output	Preservation	Historical study	Information retrieval A	Information retrieval B	Information retrieval C
Objectives	Completeness (black border), authenticity, Loss-less compression	Completeness, authenticity (ICC color profile)	Good reading experience (digital); centred pages, contrast	Very good reading experience, high contrast, Low file size	Very good reading experience, high contrast, Images not changed, Image database possible
Sample Image					

Understanding of the required result files:

1. Output Layout (see above)
2. Output Format (e.g. jpg, tiff, pdf...)
3. Output Files Naming
4. Output OCR results
5. Storage Location

## Schematic Flowchart of the Workflow

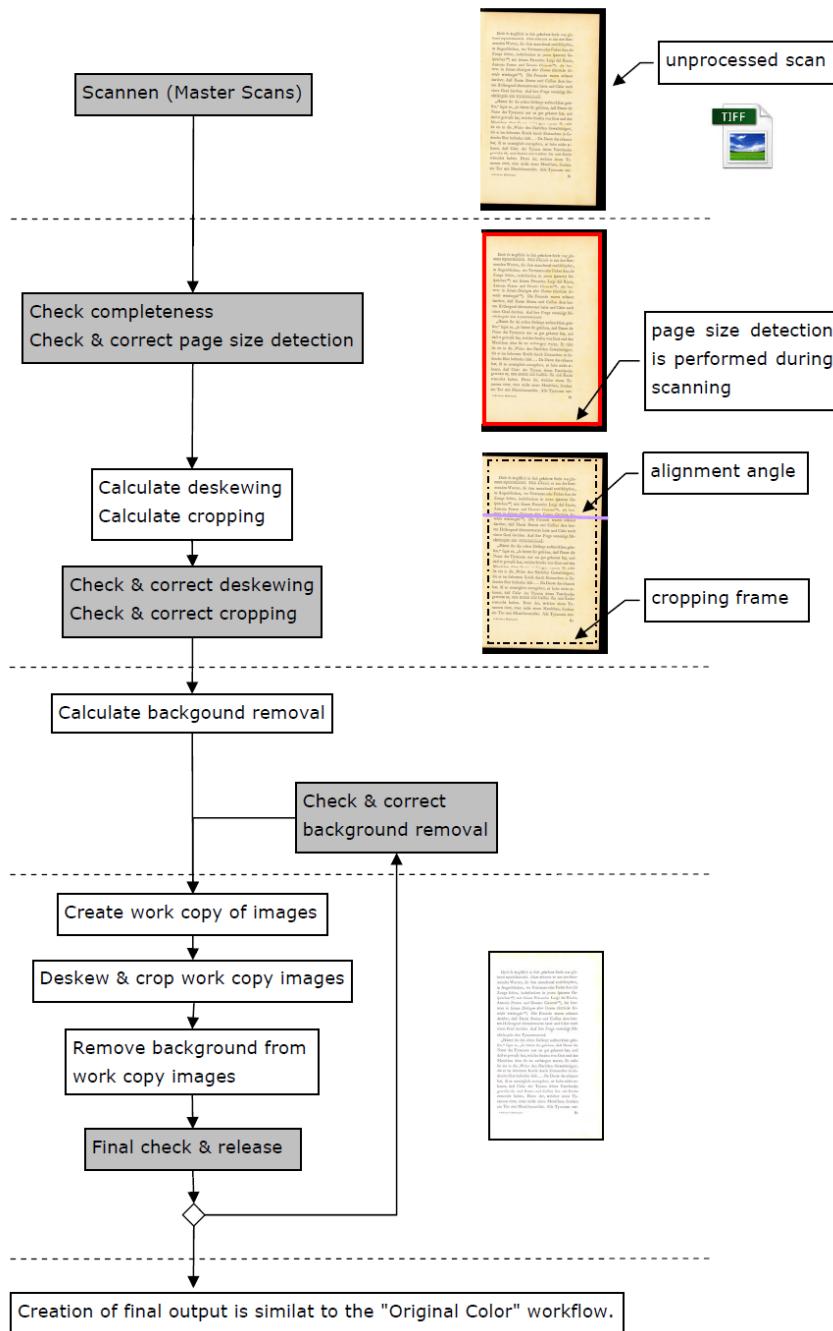


# Workflow Example

- Specification of the output

## „Workflow Specification for the ScanRobot (Original Color)“

## Schematic Flowchart of the Workflow



# Workflow Example

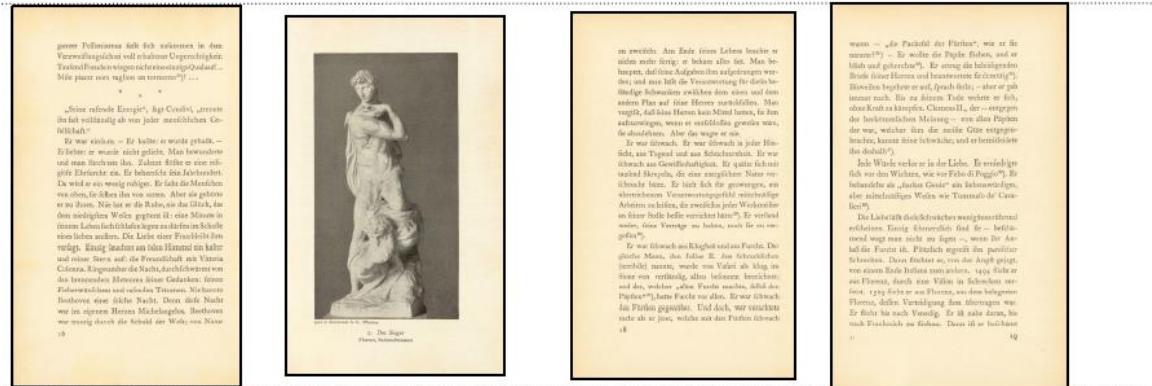
- Specification of the output

**„Optimized Output  
(Background Color Removed)“**

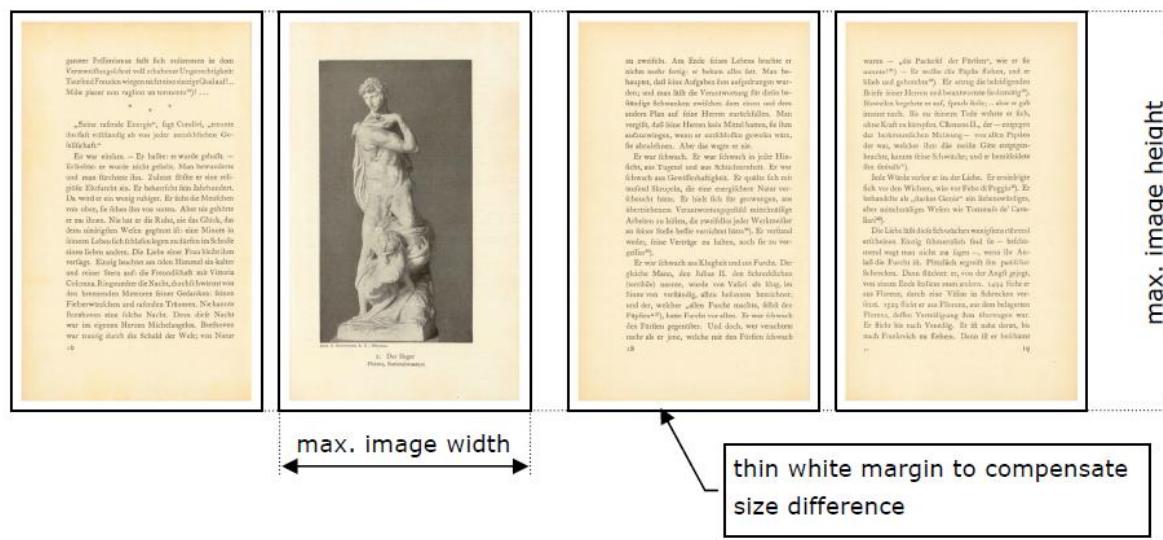
# Extrapolation

## Concept of Extrapolation

### uneven sized images after deskewing & cropping



### even sized images after extrapolation



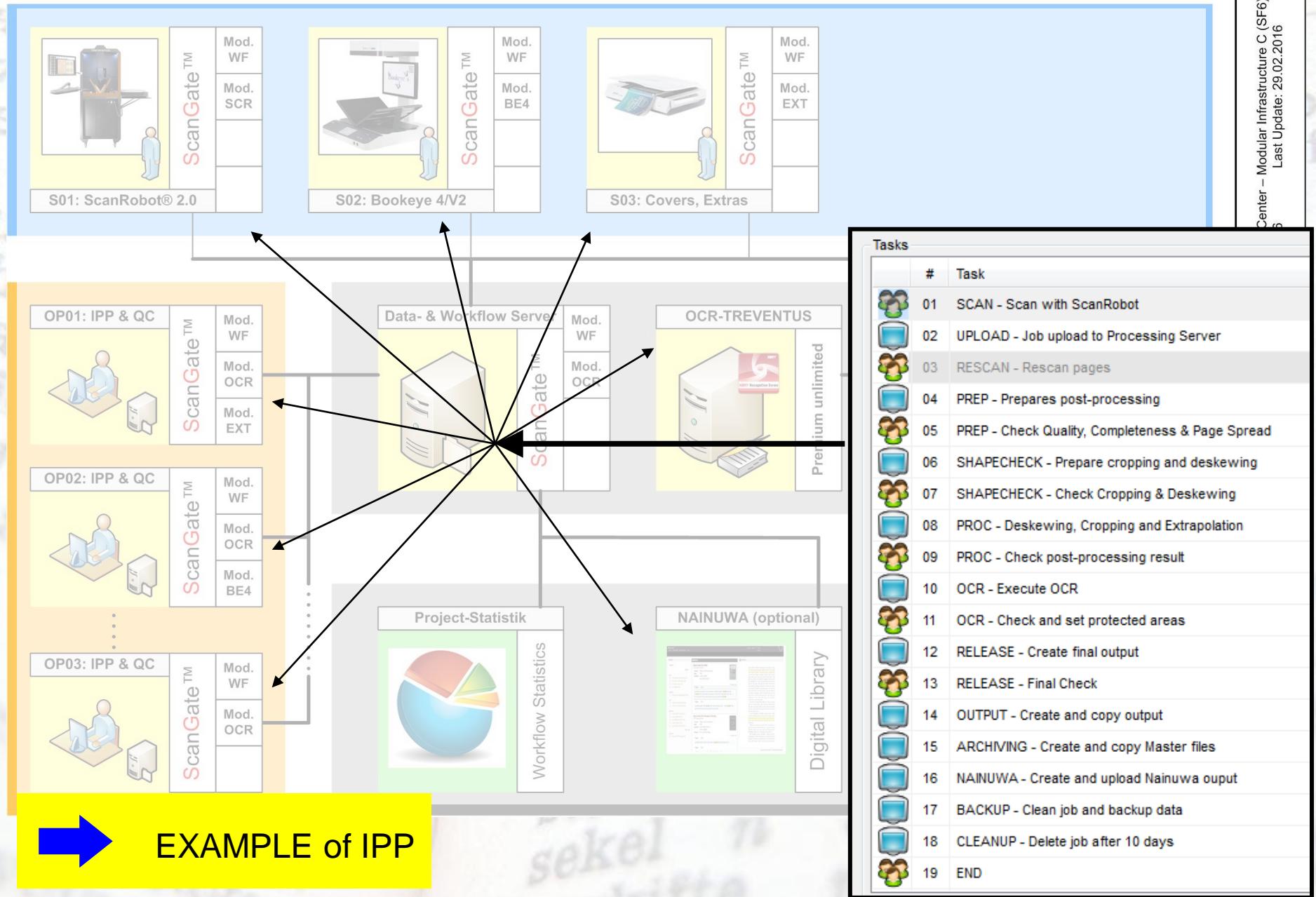
# Output Layout → Workflow for IPP

Tasks	
#	Task
01	SCAN - Scan with ScanRobot
02	UPLOAD - Job upload to Processing Server
03	RESCAN - Rescan pages
04	PREP - Prepares post-processing
05	PREP - Check Quality, Completeness & Page Spread
06	SHAPECHECK - Prepare cropping and deskewing
07	SHAPECHECK - Check Cropping & Deskewing
08	PROC - Deskewing, Cropping and Extrapolation
09	PROC - Check post-processing result
10	OCR - Execute OCR
11	OCR - Check and set protected areas
12	RELEASE - Create final output
13	RELEASE - Final Check
14	OUTPUT - Create and copy output
15	ARCHIVING - Create and copy Master files
16	NAINUWA - Create and upload Nainuwa ouput
17	BACKUP - Clean job and backup data
18	CLEANUP - Delete job after 10 days
19	END

Specification of working tasks

1. Manual Tasks
2. Automatic Tasks
3. Storage Locations

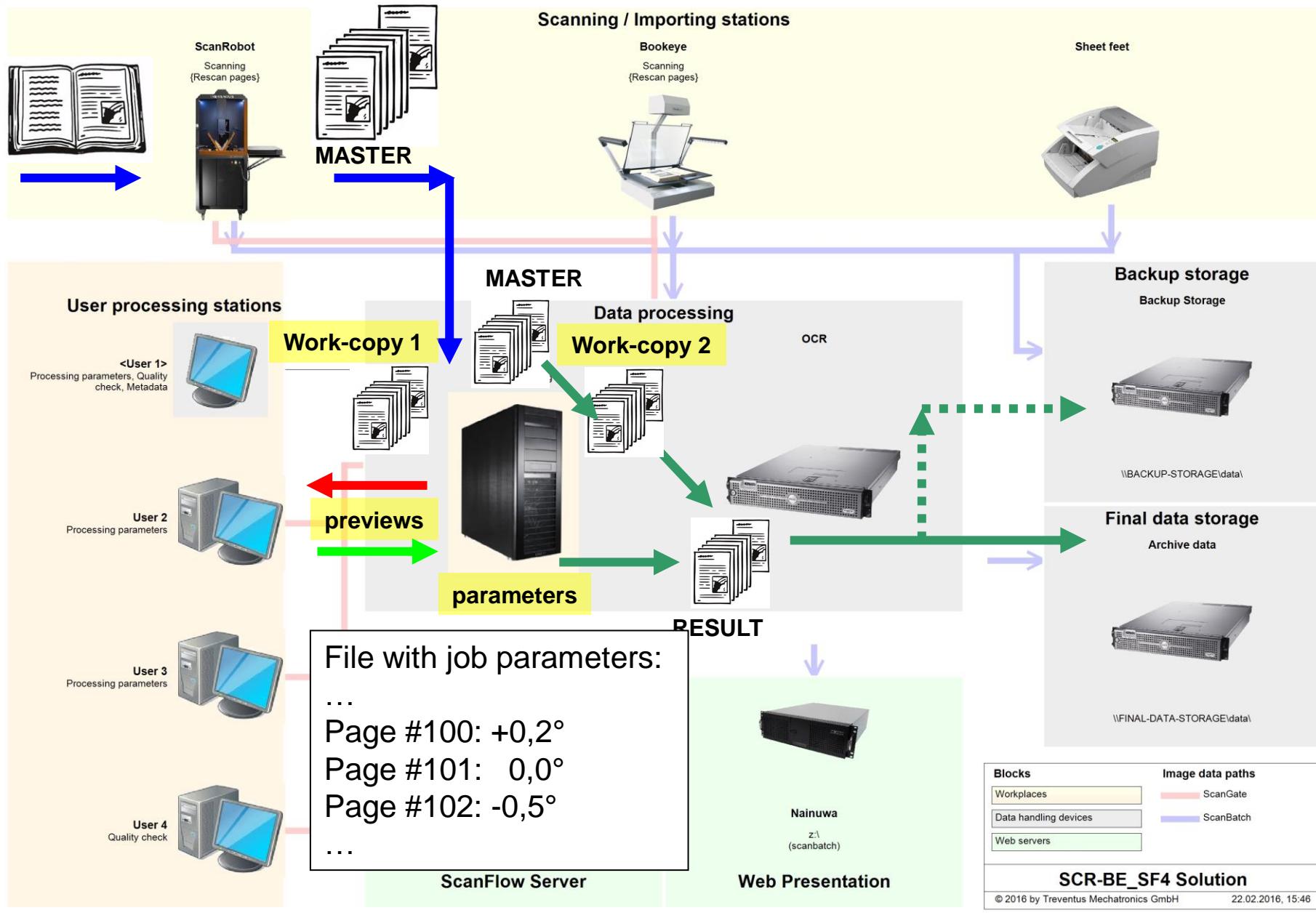
# General overview: Digitization Center Project – Infrastructure C (SF6)



# 3.) Data-Flow

Work-Flow & Data-Flow

# Sample Infrastructure

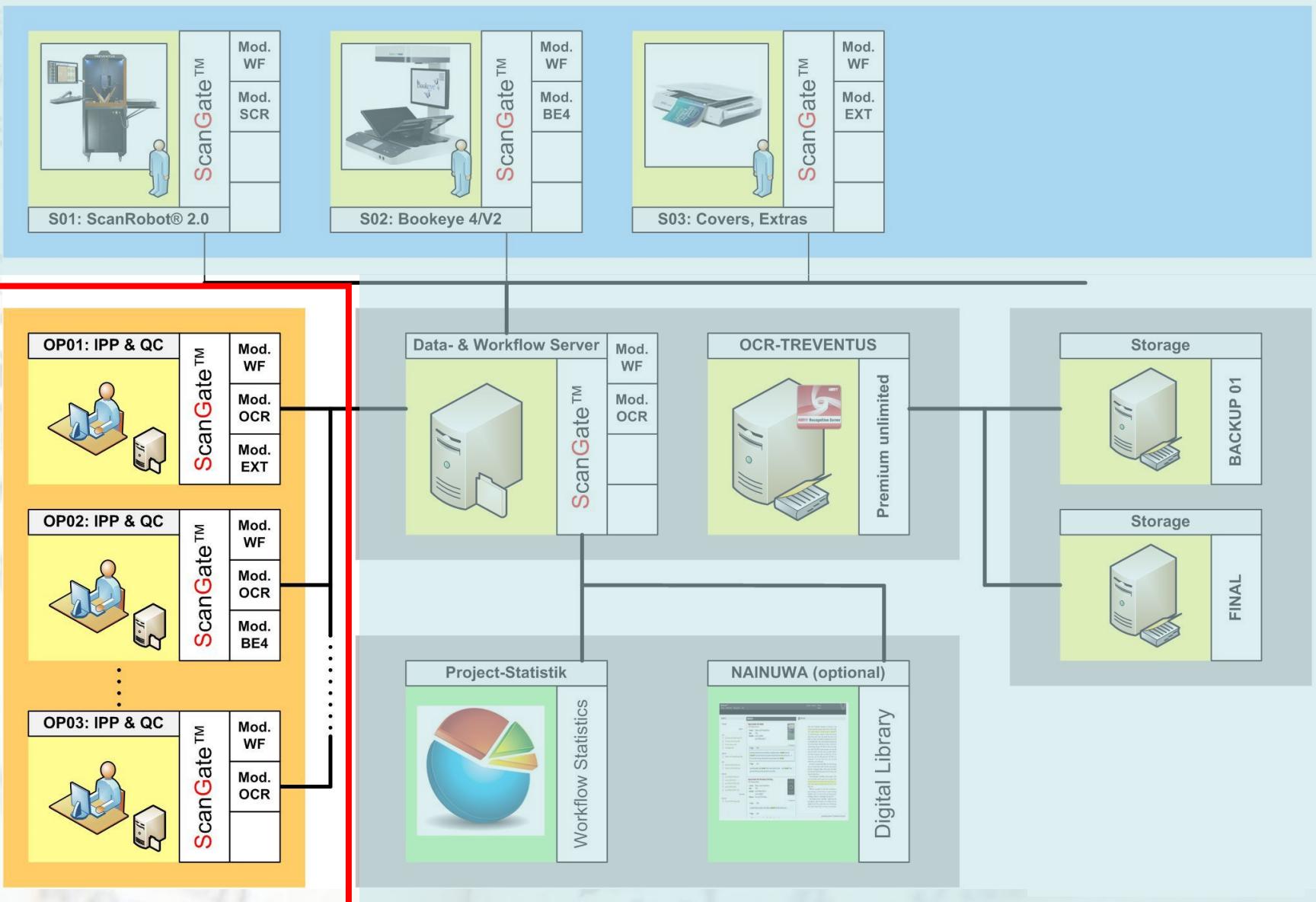


---

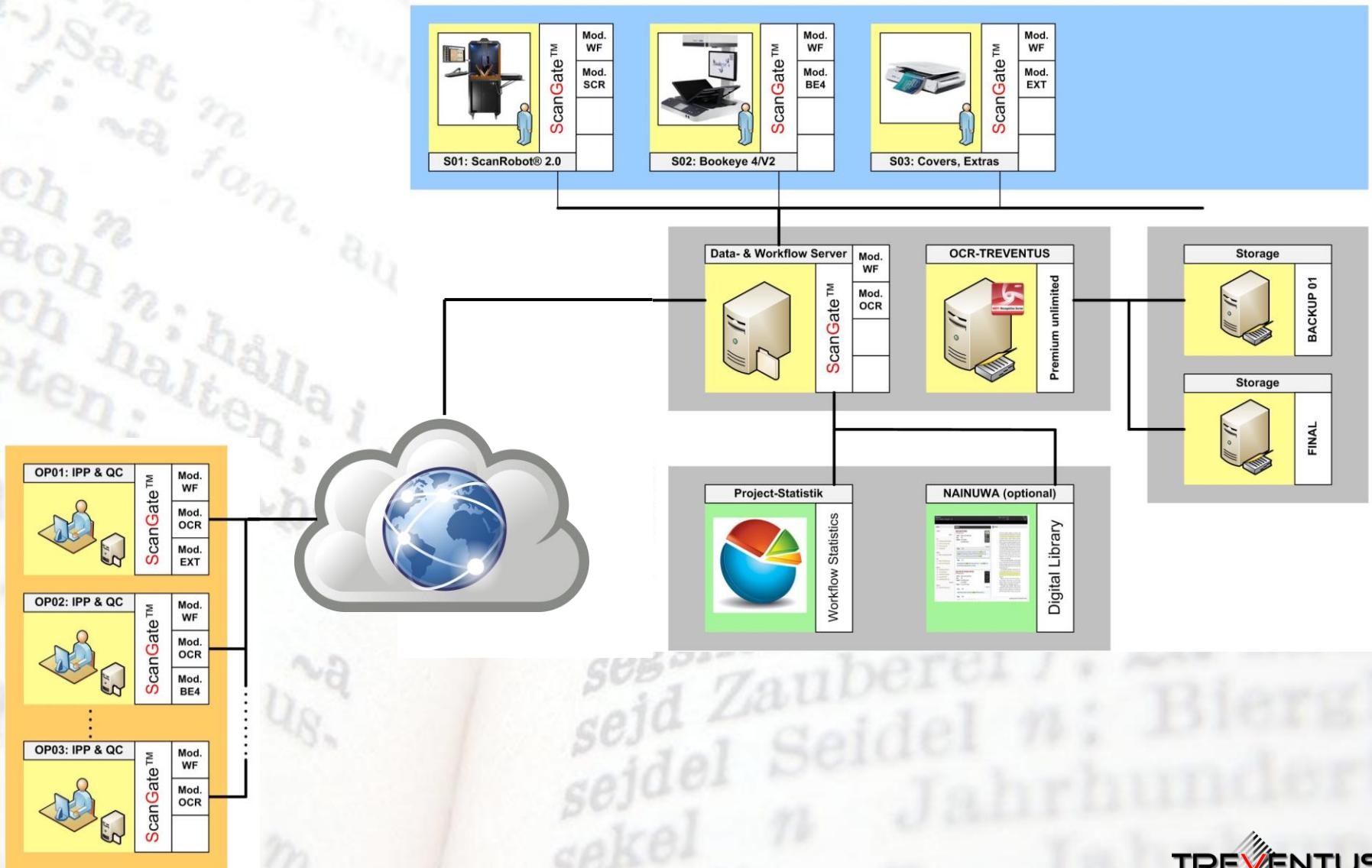
# 4.) WAN-Solution

WAN...Wide Area Network → Distributed Work

# General overview: Digitization Center Project – Infrastructure C (SF6)



# WAN - Solution



---

# 5.) Planing & Simulation

Simulation of digitization solutions with **ScanGate®**

# Simulation Results I

## General evaluation

### Project timings

Project duration	<b>33 days - 1103 h</b>
Optimization	Balanced
Scanning	482.8 h
Processing	415.8 h
Idle	203.4 h

### Operator timings



### Pending jobs



# Simulation Results II

Books_001												
#	Task name	Pages	Data location	Files size	Instance type	Instance name	Executor name	Execution start time	Entries	Execution end time	Execution	Task duration
1	(1) SCAN - Scan with ScanRobot (+ special pages ...)	300		9.0 GB	Scanning station						0 %	10 min 00 sec
2	RESTART 2 - Remove parameters: cropping, page c ...	300		9.0 GB	Data processing							1 sec
3	RESTART 1 - Clean Job, parameters are kept	300		9.0 GB	Data processing							1 sec
4	IMPORT - Create Master JPEG for Import	300		10.5 GB	Data processing							5 min 00 sec
5	PREP - Calculate Borders & Background	300		10.5 GB	Data processing							3 min 30 sec
6	UPLOAD - Job upload	300		10.5 GB	Data processing							5 min 50 sec
7	(9) RESCAN - Rescan with ScanRobot (+ special ...)	300	User 1	10.5 GB	Scanning station							1 min 00 sec
8	RESCAN - Remove cropping frame, background colo ...	300	User 1	10.5 GB	Data processing	User 1 [sb]	scanbatch					1 sec
9	PREPROCESSING - Calculate Borders & Background	300	User 1	10.5 GB	Data processing	User 1 [sb]	scanbatch					5 min 00 sec
10	(4) PREPARE - Check Page Spread (blue)	300	User 1	10.5 GB	User processing station							15 min 00 sec
11	IMGSHAPE - Set Cropping larger than individual ...	300	User 1	10.5 GB	Data processing	User 1 [sb]	scanbatch					1 sec
12	IMGSHAPE - Set Cropping smaller than individual ...	300	User 1	10.5 GB	Data processing	User 1 [sb]	scanbatch					1 sec
13	IMGSHAPE - Calculate Black Border for QC (all p ...)	300	User 1	10.5 GB	Data processing	User 1 [sb]	scanbatch					1 sec
14	(5) SHAPECHECK - Check Deskeweing & Cropping	300	User 1	10.5 GB	User processing station							5 min 00 sec
15	ENHANCEMENT - Set light Homogenisation	300	User 1	10.5 GB	Data processing	User 1 [sb]	scanbatch					1 sec
16	PREVIEW - Clear Work Images	300	User 1	10.5 GB	Data processing	User 1 [sb]	scanbatch					1 sec
17	PREVIEW - Create Work JPG (resolution original)	300	User 1	12.0 GB	Data processing	User 1 [sb]	scanbatch					5 min 00 sec
18	PREVIEW - Set Default Extrapolation for Page Sp ...	300	User 1	12.0 GB	Data processing	User 1 [sb]	scanbatch					1 sec
19	PREVIEW - Deskew - Crop - Extrapolate - Rotates ...	300	User 1	12.0 GB	Data processing	User 1 [sb]	scanbatch					5 min 00 sec
20	PREVIEW - Process Image Pixel functions QC	300	User 1	12.0 GB	Data processing	User 1 [sb]	scanbatch					5 min 00 sec
21	PREVIEW - Remove Extrapolation (all pages)	300	User 1	12.0 GB	Data processing	User 1 [sb]	scanbatch					1 sec
22	(7) RELEASE - Final Check	300	User 1	12.0 GB	User processing station							5 min 00 sec
23	OUTPUT - Clean Job	300	User 1	9.0 GB	Data processing	User 1 [sb]	scanbatch					1 sec
24	OUTPUT - Renumber pages (non arabic)	300	User 1	9.0 GB	Data processing	User 1 [sb]	scanbatch					1 sec
25	OUTPUT - Job upload to cascading server	300	User 1	9.0 GB	Data processing	User 1 [sb]	scanbatch					5 min 00 sec
26	RESTART OUTPUT PROCESSING - Clean Job	300	Cascade / OCR	9.0 GB	Data processing	Cascade / OCR [sb]	scanbatch					1 sec
27	OUTPUT - Create Work TIFF	300	Cascade / OCR	18.0 GB	Data processing	Cascade / OCR [sb]	scanbatch					5 min 00 sec
28	OUTPUT - Set Default Extrapolation for Page Spr ...	300	Cascade / OCR	18.0 GB	Data processing	Cascade / OCR [sb]	scanbatch					1 sec
29	OUTPUT - Deskew - Crop - Extrapolate - Rotate ( ...	300	Cascade / OCR	18.0 GB	Data processing	Cascade / OCR [sb]	scanbatch					5 min 00 sec
30	OUTPUT - Create JPG for OCR	300	Cascade / OCR	19.5 GB	Data processing	Cascade / OCR [sb]	scanbatch					3 min 30 sec
31	OUTPUT - Clean OCR folder	300	Cascade / OCR	19.5 GB	Data processing	Cascade / OCR [sb]	scanbatch					1 sec
32	OUTPUT - OCR Roman Type (single page PDF)	300	Cascade / OCR	19.5 GB	Data processing	Cascade / OCR [sb]	scanbatch					3 min 30 sec
33	OUTPUT - Process Image Pixel functions	300	Cascade / OCR	19.5 GB	Data processing	Cascade / OCR [sb]	scanbatch					5 min 00 sec
34	OUTPUT - Remove Extrapolation (all pages)	300	Cascade / OCR	19.5 GB	Data processing	Cascade / OCR [sb]	scanbatch					1 sec
35	OUTPUT - Create JPG Derivatives (color)	300	Cascade / OCR	19.5 GB	Data processing	Cascade / OCR [sb]	scanbatch					3 min 30 sec
36	OUTPUT - Clean PDF folder	300	Cascade / OCR	19.5 GB	Data processing	Cascade / OCR [sb]	scanbatch					1 sec
37	OUTPUT - Exchange images in PDF (color)	300	Cascade / OCR	19.5 GB	Data processing	Cascade / OCR [sb]	scanbatch					1 sec
38	OUTPUT - Clean PDF combined folder	300	Cascade / OCR	19.5 GB	Data processing	Cascade / OCR [sb]	scanbatch					1 sec
39	OUTPUT - Combine PDF	300	Cascade / OCR	19.5 GB	Data processing	Cascade / OCR [sb]	scanbatch					1 sec
40	ARCHIVE - Compress Master TIFF (LZW)	300	Cascade / OCR	19.5 GB	Data processing	Cascade / OCR [sb]	scanbatch					5 min 00 sec
41	ARCHIVE - Clean Job	300	Cascade / OCR	9.0 GB	Data processing	Cascade / OCR [sb]	scanbatch					1 sec
42	ARCHIVE - Archive Job	300	Cascade / OCR	9.0 GB	Data processing	Cascade / OCR [sb]	scanbatch					5 min 00 sec

---

# 6.) Workflow - Visual

Monitoring of digitization solutions

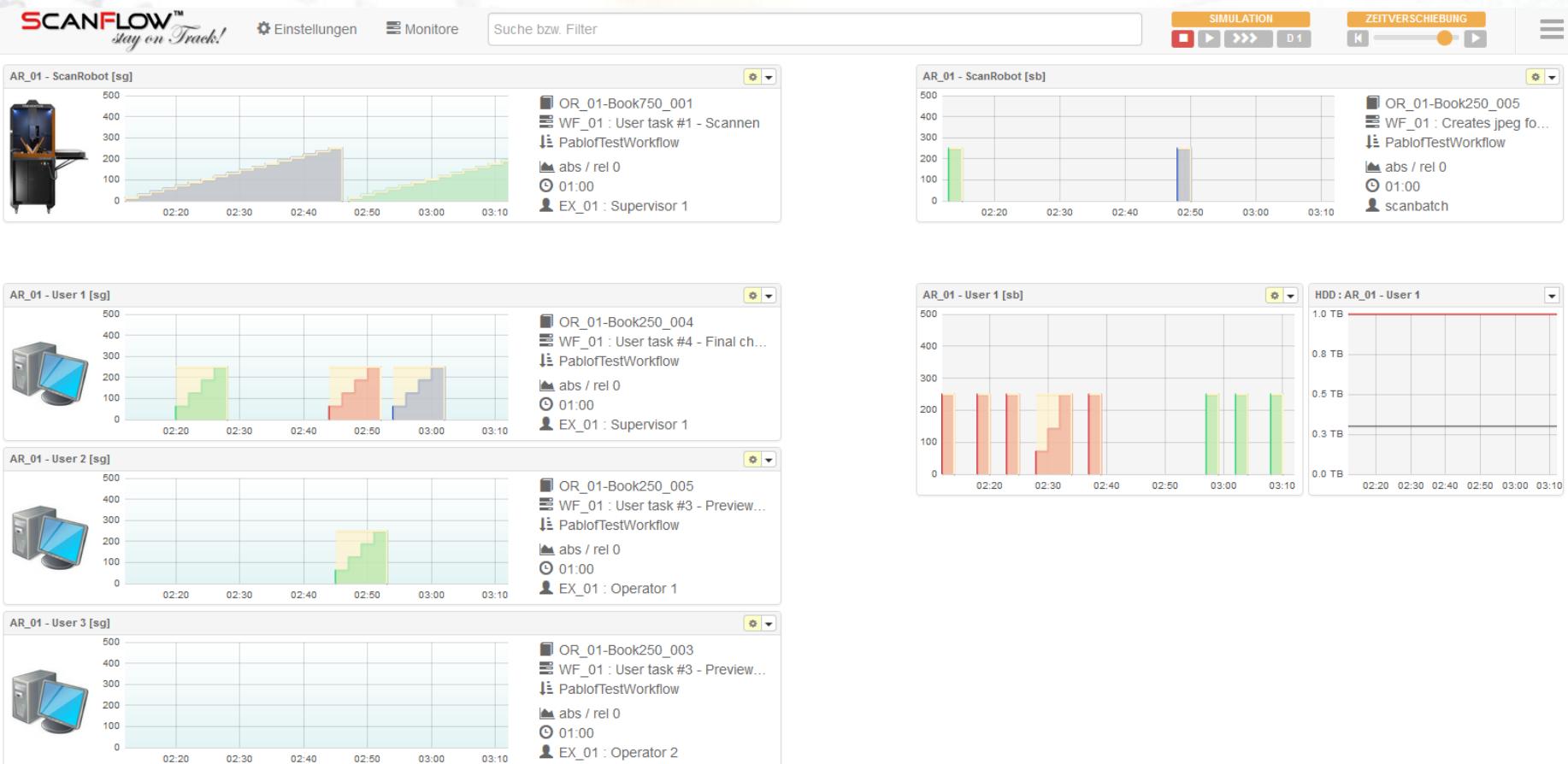
With modules of **ScanFlow™ Server**

# Workflow Visualization

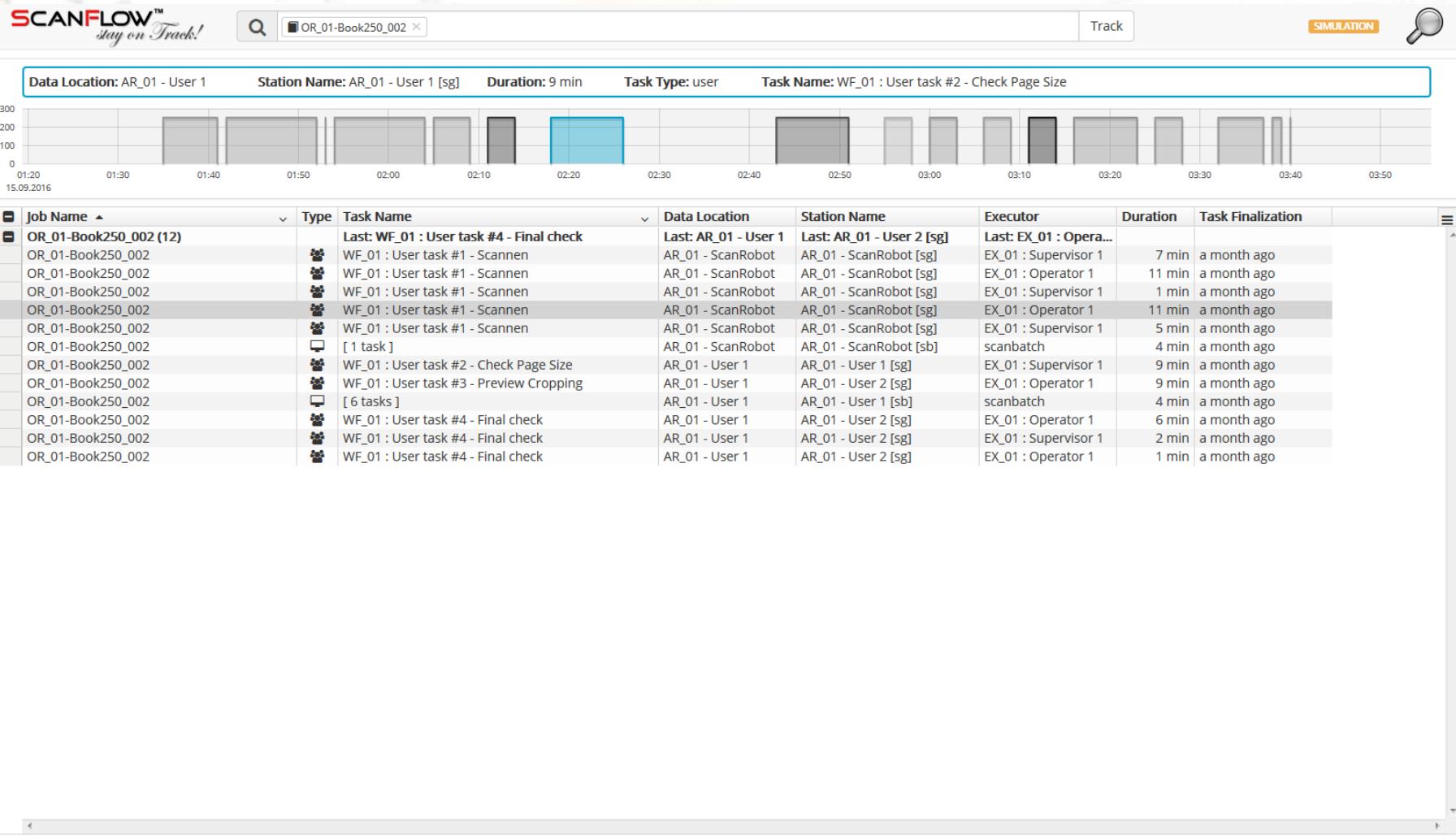
---

- Visualization of digitization solutions  
With modules of **ScanFlow™ Server**
- Explanation:
  - Workflow = Motor in the CAR
  - Visualization = Tacho of the CAR

# Workflow Server: Monitoring Module



# Workflow Server: Tracking Modul



---

# 7.) Conclusion

Sostenibilitá?

# Argomento: Digitalizzazione Sostenibile

---

## Sostenibilità?

# Argomento: Digitalizzazione Sostenibile

---

- Data
  - Formati
  - Metadati...
- Soluzione
  - Hardware
  - Software
  - Processes (= Workflow)
  - Project – Plan

*Digital change  
is a journey,  
not a destination.  
Let's travel together!*





Thank you!

[www.treventus.com](http://www.treventus.com)

[tratter@treventus.com](mailto:tratter@treventus.com)



# TREVENTUS

---

Company & Production

# ScanRobot® production line I



# ScanRobot® production line II



TREVENTUS manufacturing plant