succeed★ DELIVERABLE SUBMISSION SHEET

To: Marcel Watelet   (Project Officer)

DG Communications Networks, Content & Technology
Creativity Unit
EUFO 01/196A- European Commission
L-2920 Luxembourg

From: Support Action Centre of Competence in Digitisation

Project acronym: Succeed   Project Number: 600555

Project Manager: Rafael C. Carrasco Jiménez

Project Coordinator: Universidad de Alicante

The following deliverable:

Deliverable title: Training materials

Deliverable number: D3.4

Deliverable date: 30/12/2013

Partners responsible: INL

Status: ☒ Public   ☐ Restricted   ☐ Confidential

is now complete

☒ It is available for your inspection

☐ Relevant descriptive documents are attached

The deliverable is:

☒ A document

☒ A website   Url: http://www.digitisation.eu/training/succeed-training-materials

☐ Software

☐ An event

☐ Other

Sent to Project Officer: **Sent to functional mailbox:**   **On date:**

Marcel.Watelet@ec.europa.eu   CNECT-ICT-600555@ec.europa.eu   30/12/2013

D3.4 Training materials

Succeed

26/12/2013

Deliverable number	D3.4	Start: 7	Due: 10	Actual: 12
Deliverable name	Training materials			
Internal/External	External			
Activity type	R			
Participant	UA, **INL**, IAIS, PSNC			
Estimated person months for this deliverable	3 PM			
Dissemination level[1]	PU			

Document history

Revisions					
Version	Status	Author	Date	Changes	
0.1	Draft	Sebastian Kirch	05.12.2013		
0.2	Draft	Katrien Depuydt, Isabel Martínez, Sebastian Kirch	09.12.2013		
0.3	Pre-final	Sebastian Kirch	20.12.2013	Integrated all comments & suggestions	
1.0	Final	Isabel Martínez, Rafael Carrasco, Katrien Depuydt	26.12.2013	Integrated comments and suggestions from internal supervisors and project coordinator.	

Approvals				
Version	Date of approval	Name	Role in project	Signature
0.3		Stefan Pletschacher, Faye Fornasier	Supervisor D3.4	

Distribution			
This document was sent to:			
Version	Date of sending	Name	Role in project
0.1	05.12.2013	WP3 partners	
0.2	10.12.2013	Stefan Pletschacher, Faye Fornasier	Supervisor D3.4
0.3	20.12.2013	All WP3 participants	
1.0	26.12.2013	All WP3 participants	
1.0	30.12.2013	Marcel Watelet (P.O), EC	

1 PU Public; RP Restricted to other programme participants (including Commission Services); RE Restricted to a group specified by the consortium (including Commission Services); CO Confidential, only for members of the consortium (including the Commission Services)

1. OVERVIEW

This deliverable is part of WP3. This work package will support the validation of digitisation tools, linguistic tools and resources created by either commercial parties or research and development programs by fostering their transference for exploitation to libraries and other cultural heritage organisations. In particular, a selection of the tools compiled in the Survey of Tools[2] (deliverable 3.1) will be made available to the community and the partners involved will provide assistance for the adaptation of the tools to specific domains and languages.

To foster the take-up of tools, 4 internal libraries and 9 external libraries will evaluate digitisation tools of their choice within the Succeed project. The objective of deliverable D3.4 is to provide documentation and training material for the tools to be validated. These materials must help the participating libraries to take-up the tools in their productive environment. Therefore, the training material for each tools needs to include the following topics:

- How to install the tool.
- How to integrate the tool into existing workflows and applications.
- How to operate and customize the tool.

For improved usability and easier continuous update, the training material is distributed on-line as web pages (in HTML format) and also as standard learning objects (in SCORM format). This report, therefore, only lists the tools and some basic information about their associated material. The actual content can be found at [http://www.digitisation.eu/training/succeed-training-materials].

---

[2] Overview on available tools for text digitisation: http://succeed-project.eu/publications/available-tools/index-succeed

2. METHODOLOGY

The production of training material proceeded in five consecutive steps. The dates when each step was performed are given in parentheses.

1. **Document Structure (July 2013)**
   During a workshop held at Fraunhofer IAIS on July 9th, the content of the training material documents was discussed. The WP3 partners agreed then that each tool needs:

   - Installation instructions explaining where to get and how to install a tool on the supported platforms and operating systems.
   - Additional documentation explaining how to configure, customize and integrate a tool into an existing digitisation environment.
   - A quick start guide for basic usage which should allow the reader to start using the tool with little effort
   - In some cases, examples of how to use the tool, additional help like screenshots or links and advanced usage instructions.
   - Information about licensing and usage conditions.

   This content is therefore typically structured into several sections: introduction, requirements and licensing, basic and advanced usage. However, since the tools are heterogeneous, neither the proposed structure of the training documents nor a single learning design were strictly enforced.

   Additionally, the documentation distributed by the tool providers, when available online, was referenced rather than included in order to maintain the content up-to-date.

2. **Tool Selection (August – October 2013)**
   Before the production of training material could be started, the participating libraries had to choose which tools they wanted to evaluate within the Succeed project based on which tools fit their current digitization workflow and which requirements they had to improve this workflow. Libraries were supposed to select the tools by the end of September. However it took some libraries as long as end of October to select the tools, which is the main reason why this deliverable was slightly delayed with respect to the schedule in the Description of Work.

3. **Distribution of Work (October 2013)**
   For each tool selected by the libraries, one WP3 partner was designed to produce the training material based on the previously discussed structure. The selection was done by the WP leaders according to the experience and knowledge the partners already had with the selected tools.

4. **Creation and Compilation of Training Material (October – December 2013)**
   Each partner then produced the material assigned, creating one separate document per tool. The final versions of these documents were then aggregated by the WP3 leaders to produce the online training content for this deliverable.

5. **Distribution of Training Material (December 2013)**
   After the training material had been finalized it was distributed by the partners to the libraries using the respective tools.

The training material can be found following the link for each tool in section 3. Making the training material available online has some advantages:
- training material can be easily maintained up to date,
- training material is more accessible,
- and, since the content is open, it promotes the feedback by the community.

3. TOOLS

The following table presents the tools that were selected by the libraries. For each of these tools there is a corresponding training material document aggregating the documentation for a tool.

Tool	Tool Group	Responsible WP3 partner	Selected by
Abbyy Finereader	Text Recognition	INL	University Library of Leuven; Biblioteca Virtual Miguel de Cervantes
Alchemy API	Text Processing	INL	Library of the University of Granada
COBaLT	Text Processing	INL	Biblioteca Virtual Miguel de Cervantes
DBpedia Spotlight	Text Processing	INL	Bibliothèque nationale de France
Document Deskewer	Image Processing	IAIS	National Library of Finland; University Library of Darmstadt
Evaluation Tool for OCR	Text Recognition	UA	Bibliothèque nationale de France
Gimp	Image Processing	IAIS	General Historical Library of Salamanca
ImageMagick	Image Processing	IAIS	Wielkopolska Biblioteka Cyfrowa; University Library of Bratislava
JHOVE2	Misc. Utilities	PSNC	Wielkopolska Biblioteka Cyfrowa
Lexicon Webservice	Text Processing	INL	Biblioteca Virtual Miguel de Cervantes; Bibliothèque nationale de France; Koninklijke Bibliotheek; The British Library
NE Attestation tool	Text Processing	INL	University Library of Antwerp
NERT	Text Processing	INL	University Library of Leuven; Koninklijke Bibliotheek
Newspaper segmentation & Korrektor	Layout Analysis	IAIS	National Library of Finland; University Library of Darmstadt
NLTK (NE)	Text Processing	INL	University Library of Antwerp; Koninklijke Bibliotheek
Omnipage	Text Recognition	PSNC	General Historical Library of Salamanca

Tool	Tool Group	Responsible WP3 partner	Selected by
Page Curl Correction	Image Processing	IAIS	Biblioteca Virtual Miguel de Cervantes
Scan Tailor	Image Processing	PSNC	Wielkopolska Biblioteka Cyfrowa; Wroclaw University Library; University Library of Bratislava; Library of the University of Granada
Stanford (NE)	Text Processing	INL	University Library of Antwerp
Tesseract OCR	Text Recognition	PSNC	Wroclaw University Library