



THE NATIONAL ARCHIVES
SERVICE OF FINLAND

READ- Recognition and Enrichment of Archival Documents

Vili Haukkovaara
Development manager
National Archives of Finland
vili.haukkovaara@arkisto.fi



READ

13.9.2016

1



In brief

- Basics
- The big picture
- Goals
- NAF tasks and vision of architecture



Basics

- The READ project
 - Aims to build a new virtual research platform which will revolutionize access to digital handwritten material
 - In collaboration with technical experts, archives, researchers and public users



Basics

- Project funded by European Commission
 - Part of Horizon2020 program
 - e-infrastructure Call for Virtual Research Environments
 - Successor to the tranScriptotium project



Basics

- 1/2016-6/2019
- 8,2mill. € budget
- 13 partners from all over Europe
 - +10 associated partners (MoU partners like Munch museum)



Basics

- Partners

University of Innsbruck (co-ordinator)

University College London

Vienna University of Technology

Leipzig University

National Archives Finland

Ecole Polytechnique Federale de Lausanne

National Center for Scientific Research “Demokritos”

Universitat Politecnica de Valencia

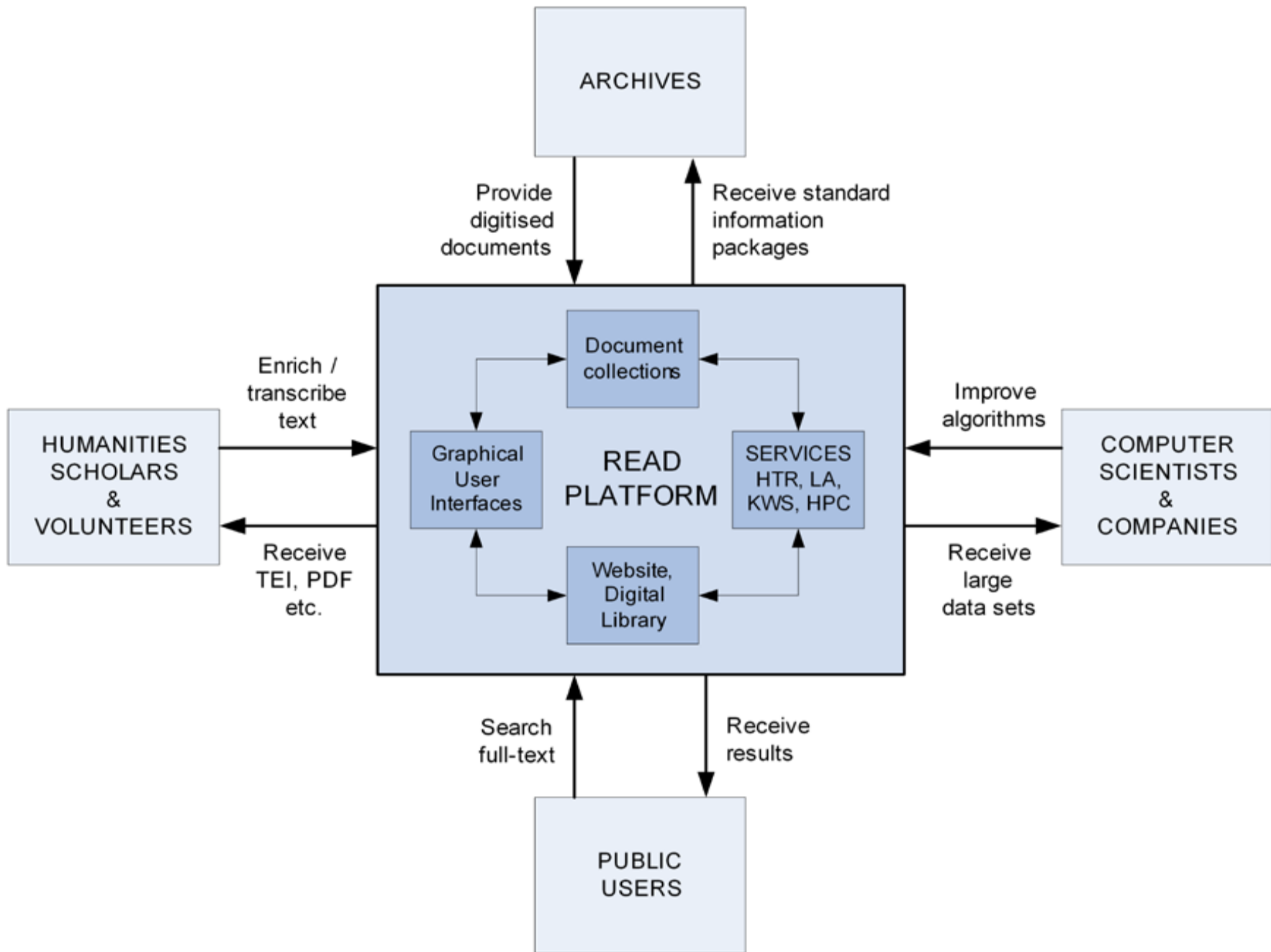
University of London Computer Centre

University of Rostock

Xerox Research Centre Europe

Zurich State Archives

Passau Diocesan Archives



Goals

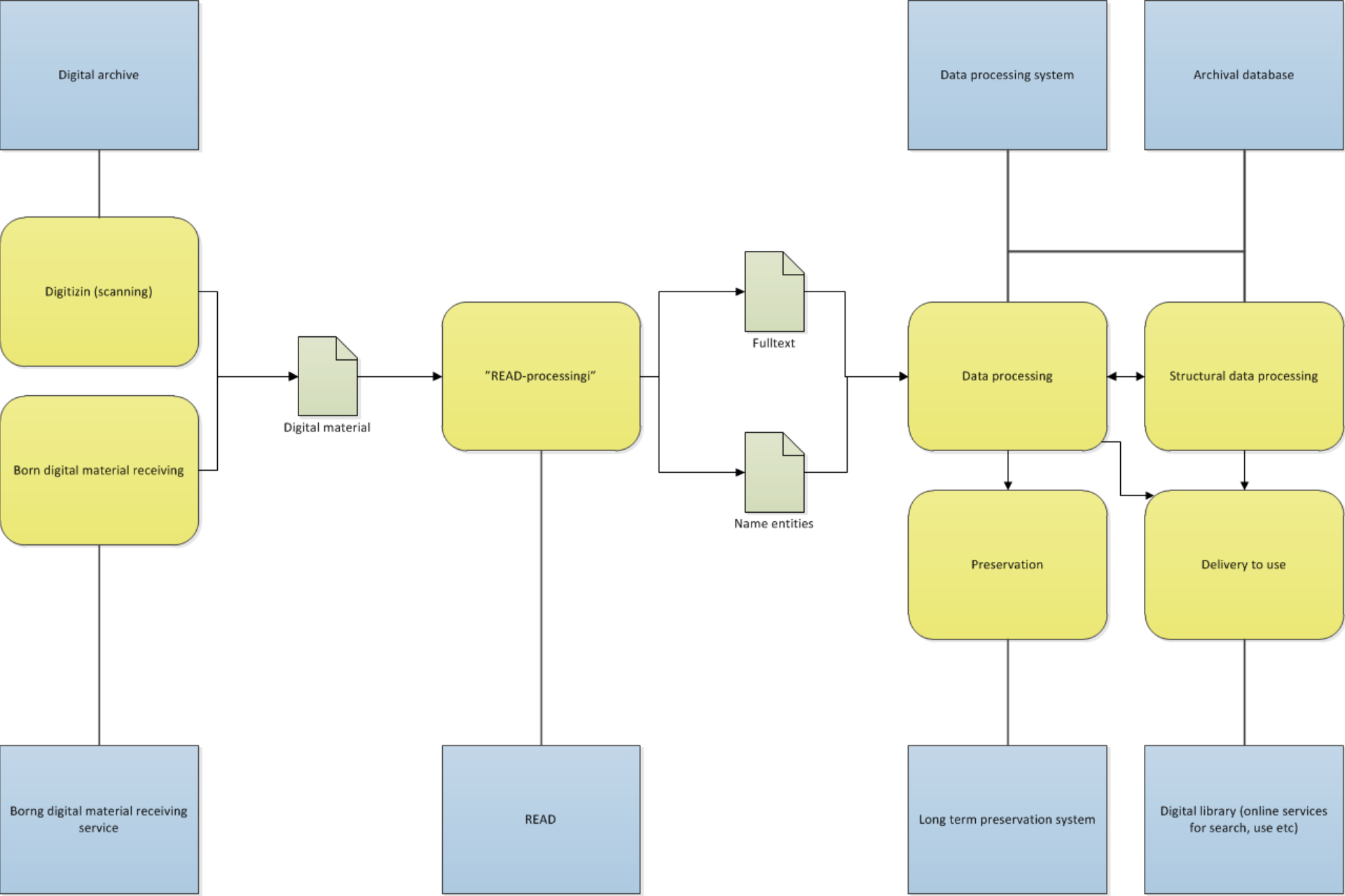
- Improve the whole workflow cycle
 - Layout Analysis, Document Understanding
 - Handwritten Text Recognition
 - Exploration, Keyword Spotting
 - User involvement
- Collect large data sets for research purposes
 - Millions of image files will be collected and be made available to the research community

Goals

- Integration of archival collections
 - To get archives involved
- Humanities scholars
 - Transkribus as a standard platform for transcriptions
- Create sustainable service
 - READ-platform will exist after the project

NAF tasks

- To provide large amounts of data
 - Digital material
 - Transcriptions
 - But how?
 - Involve archives, researchers and public users
- Test the whole system with large amounts of data



13.9.2016

11

Want to know more?

- The project
 - <http://read.transkribus.eu/>
- Transkribus software (desktop client)
 - <https://transkribus.eu/Transkribus/>
- The precursor project (transScriptorium)
 - <http://transcriptorium.eu/>