

Task Brief 3.1.5

Task Brief

WP Number and Title	WP3	Technological implementation
Task Number and Title	T3.1	Technological implementation (Overall Coordination)
Subtask Number and Title	T3.1.5	Development and adaptation of specific tools; implementation and adaptation of taxon finder and name recognition tools; improvement and implementation of OCR techniques.
Task Leader and Contact	ATOS	roger.essoh@atosorigin.com ¹
Sub task Leader and Contact	-	-
Related Deliverable Number and Title	D3.5	Technical architecture status and progress report with particular focus on the development of the German prototype (M 12)
	D3.6	Release of German prototype (M 18)
	D3.7	Key components documented for output of D3.5 e.g. BHL-Europe Portal, OCR demonstrators, distributed storage model, etc. (M 24)
Start Date	M6 1.10.2009	End Date M36 30.04.2012

1 Partners involved

Partner	PM	Contact (email)	Obligation
02 NHM		a.smales@nhm.ac.uk	WP3 Leader – Coordination
05 AIT	3	kochw@ait.co.at	Integration of ontology into METS editor.
06 ATOS		Roger.essoh@atosorigin.com	Task Leader
23 Sp2000		pscoones@wiley.com	Input to SIP creation
26 MOBOT		Chris.Freeland@MOBOT.ORG	Input on providing metadata structures for serials and books.

¹ Minutes Tech Workshop: http://bhl.wikispaces.com/file/view/BHL-E_TechWorkshop.pdf [14 May 2009]

Task Brief 3.1.5

Partner	PM	Contact (email)	Obligation
28 UEBER		dobratz@cms.hu-berlin.de	

2 Objectives

This sub task needs to create tools that are able to compose Submission Information Packages (SIP) that are compatible with the BHL Europe system. The composition of SIPs includes the creation of digital representations of physical objects and the description of these. This covers integration of controlled vocabulary and images of pages and the management of these.

3 Methodology

To complete this sub task, the interfaces definition to Ingest needs to be done first. As soon as the interface is defined, tools that compose SIPs can be created. The affected OAIS modules are shown in Figure 3-1.

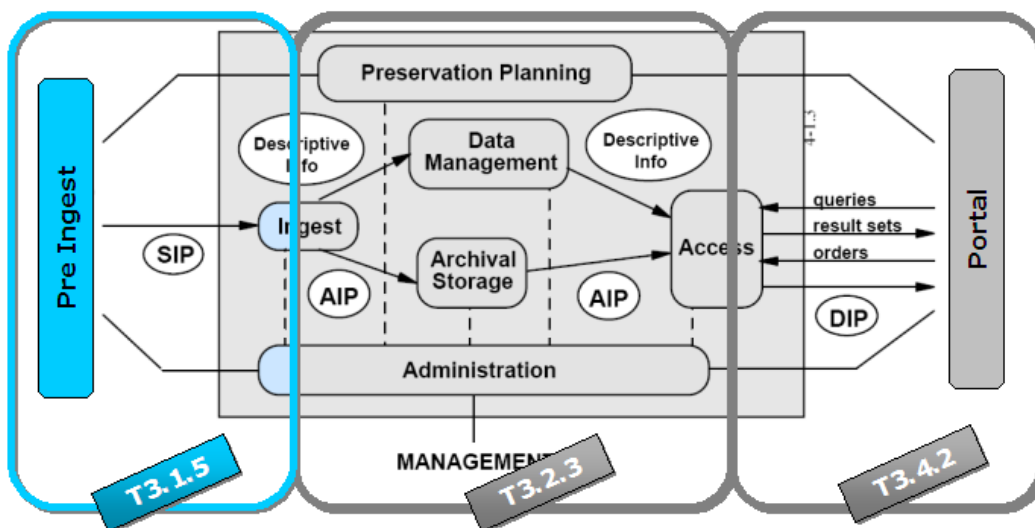


Figure 3-1 OAIS components covered by this sub task

As part of the interface definition, the payload needs to be defined. The payload must use existing standards and will therefore use MARC as descriptive metadata and METS as container for the SIP.

During the creation of the SIP creation toolset and the interface definition, scanning tools and procedures need to be created. These correlate with T3.1.3 and T3.1.4. The methodology is outlined in Figure 3-2.



Figure 3-2 Methodology outline

Task Brief 3.1.5

4 Description of Work

This chapter describes the activities done by each partner to achieve this sub task.

4.1 *TBW by task leader*

4.2 *Defining the SIP*

Participants: AIT, Sp2000, Mobot

Related tasks: T3.1.3, T3.1.4, T3.2.3

Deliverables: D3.5, D3.6, D3.7, D3.9

AIT defines a SIP format based on the standards MARC and METS. Integration of FRBR and relation of serials will provide input for T3.2.3 and the data model. The SIP will be walked through with Mobot and Sp2000 to find technical flaws. The requirements will be taken from T2.1.3 and T3.1.4.

4.3 *Build one SIP creation toolset*

Participants: AIT, Sp2000, Mobot

Related Tasks: T3.2.3

Deliverables: D3.5, D3.6, D3.7, D3.9

With the help of Mobot and Sp2000, AIT will build one SIP creators that will provide the following features:

- Provide BHL fields on item and page level
- Link images in the file and structMap of METS
- User Management
- Handle the ingest process

4.4 *Adding ontology to the SIP creation toolset*

Participants: AIT

Related tasks: T3.4.1, T3.2.3

Deliverables: D3.5, D3.6, D3.7, D3.9

AIT will add ontology functionality, required for the portal, to the toolset. This covers the integration of web services into the interfaces as well as placing the data in the SIP.

4.5 *TBW by task leader*