System enhancement (WP4) and design of new registries (WP5)
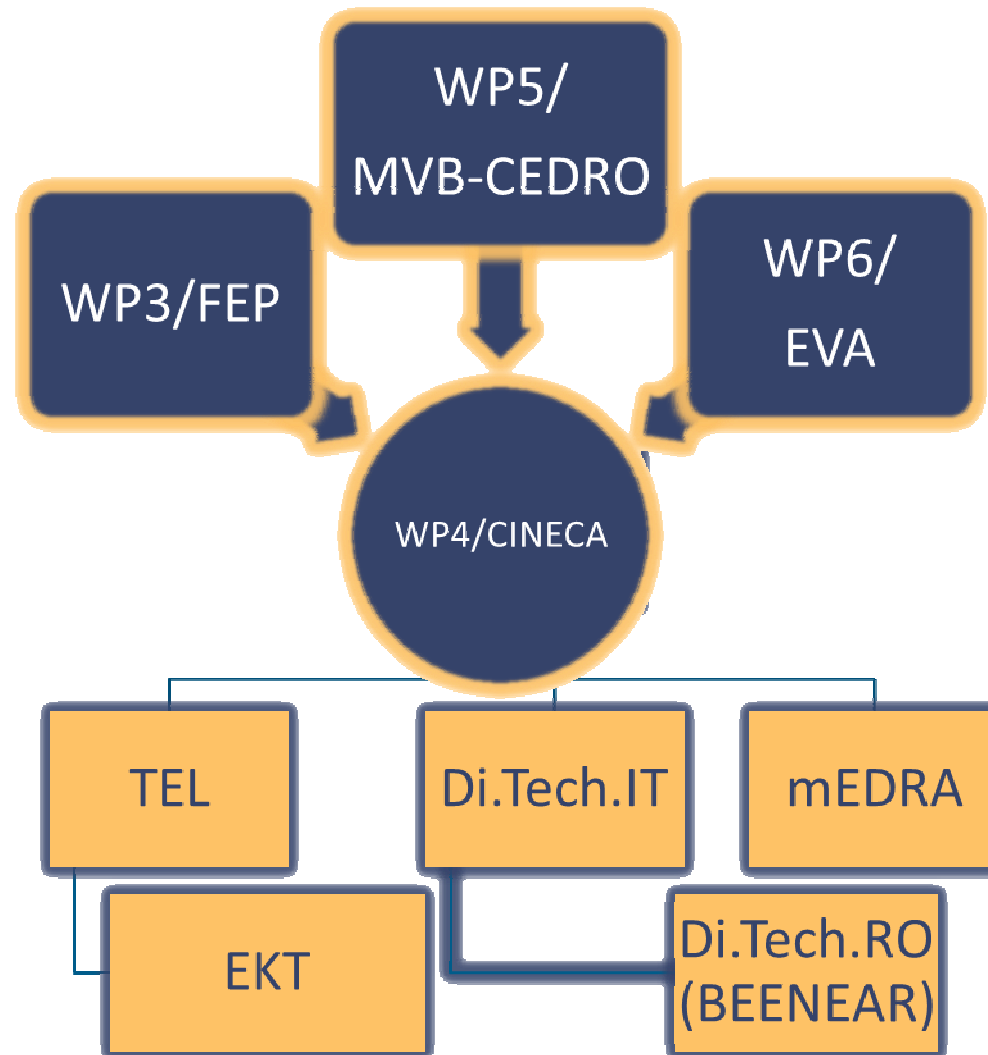
Brussels, March 20th

WP4: management strategy

WP4: Expected results, outputs, dependences

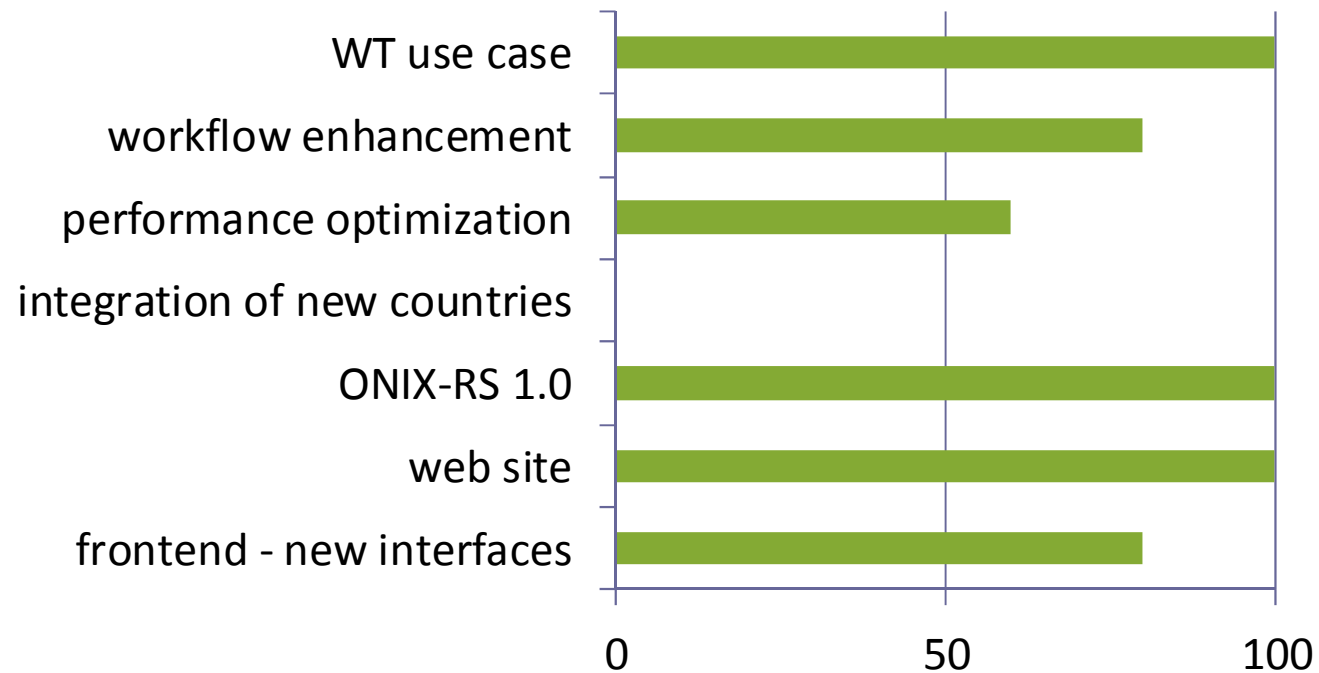| System Enhancement | • System maintenance
• Use Cases (Wellcome Trust)
• Enhancements | CINECA-TEL-mEDRA
D4.1 - M13 | ARROW |
New countries inclusion	• Integration of new data providers in new countries (roadmap for inclusion)	CINECA-TEL-Ditech D4.4 – M30	WP3
Set up BIP service	• minimum set of core functionalities, according to the technical specifications (MVB)	CINECA-mEDRA-Ditech EKT D4.3 – M18	WP3 WP5
Set up RRO service	• minimum set of core functionalities, according to the technical specs (CEDRO)	CINECA-mEDRA-Ditech D4.4 – M20...M23	WP3 WP5
Integration of visual material	• Enhance metadata schema • Review the Arrow business logic	CINECA-TEL-Ditech D4.4 – M30	WP6
Registry of Orphan Work	• Design and implementation of new tools for rightholders claiming of rights in collaboration with UIBK.	CINECA-mEDRA D4.2 – M15	UIBK

WP4: Activities done M1-M10

T4.1 subtasks	Activities	Date of completion
Frontend - New interfaces	• Diligent Search Results for different User Profiles • Diligent Search Partial Results for Basic Users • Data Entry Interface Enhancements • Other modifications	06/2011
Wellcome Trust use case	• Schema upgrade to 0.27 • No match management • Automatic validation	08/2011
Update of ARROW messages schema to ONIX-RS 1.0	• Schema validation • Updating of data model • Updating of service components involved	02/2012
Workflow enhancements	• No match management • Automatic validation • Workflow configurability	02/2012
Performance optimizations	• Load test • Improvement of responsiveness in 'Monitoring Area' • Better error management with retry mechanism • Other system enhancements issues	In Progress

T4.1: Current status

Work done %

WT use case	100
workflow enhancement	80
performance optimization	60
integration of new countries	0
ONIX-RS 1.0	100
web site	100
frontend - new interfaces	80

Integration of New Countries

- Delivered the **Arrow Integration Specifications to BIP and RRO** data providers

- Organised several **video conference call** with the following Data Providers:
  - Greece: EKEBI (BIP)
  - Italy : IE (BIP)
  - Belgium: Boek….
  - Netherland: Scheduled for end of March with Central Book House

  Goal:
  - provide more clarifications for the requirements,
  - analyse the work that should be done at data provider level
  - highlight and address possible difficulties at technical level
  - define a scheduling for integration in Arrow

WP4: first five months planning

Action	Output (also including final deliverables)	Partner and person in charge	Date of completion
Workflow configurability	D4.1 Interim report about enhancements	CINECA – mEDRA	30/04/2012
D4.1 Interim report about enhancements	D4.1 Interim report about enhancements	CINECA	30/04/2012
Inclusion of new countries	D4.1 Interim report about enhancements	CINECA – mEDRA – DI.TECH - AIE	To be defined with each country
Inclusion of new countries in TEL's catalogue	D4.1 Interim report about enhancements and D4.4 Arrow Plus production system	KB	30/06/2012
Implementation in TEL system of improvements identified in Validation Report from project ARROW	D4.1 Interim report about enhancements	KB	30/06/2012
Integration of the TEL prototype from project ARROW into the TEL production systems	D4.1 Interim report about enhancements	KB	30/04/2012
Task 3.1: Implementation of BIP Registry	D4.2 BIP registry	CINECA – mEDRA – DI.TECH – BEE - EKT AIE – MVB – Editeur	30/09/2012
Task 3.2: Implementation of RRO Registry	D4.3 RRO registry	CINECA – mEDRA – DI.TECH – BEE - EKT AIE – MVB – Editeur	28/02/2013

WP4: Enhancement and Inclusion at TEL level

Action	Output (also including final deliverables)	Partner and person in charge	Date of completion
Inclusion of new countries in TEL's catalogue	D4.1 Interim report about enhancements and D4.4 Arrow Plus production system	KB – Nuno Freire	30/06/2012
Implementation in TEL system of improvements identified in Validation Report from project ARROW	D4.1 Interim report about enhancements	KB – Nuno Freire	30/06/2012
Integration of the TEL prototype from project ARROW into the TEL production systems	D4.1 Interim report about enhancements	KB – Nuno Freire	30/04/2012

Integration of the TEL prototype from project ARROW into the TEL production systems

- The prototype was an independent system completely operated by the Technical Team
  - The new system is integrated with the internal TEL systems and is operated by the Collections Team, responsible for all data ingestion in TEL
  - Also involves training of the Collections Team for getting the national bibliographies in the data formats required for ARROW, with links to VIAF when possible
- Hosting of the VIAF data set in the TEL authority file
- Clustering and Matching performance and scalability investigation, and improvement.

Implementation in TEL system of improvements identified in Validation Report from project ARROW

□ Two particular issues will be addressed until June:

- ▫ Extract from MARC records, the contributors within the natural language text in *statements of responsibility*

- ▫ VIAF integration: search for a contributor in VIAF when the bibliographic record does not contain its identifier in the national authority file

Implementation in TEL system of improvements identified in Validation Report from project ARROW

- Extract from MARC records, the contributors within the natural language text in *statements of responsibility*

  - The first author is almost always present in the structured data fields of MARC, but sometimes contributors are only present in the *statement of responsability*

  - As a consequence, primary work and secondary work clusters may be imprecise
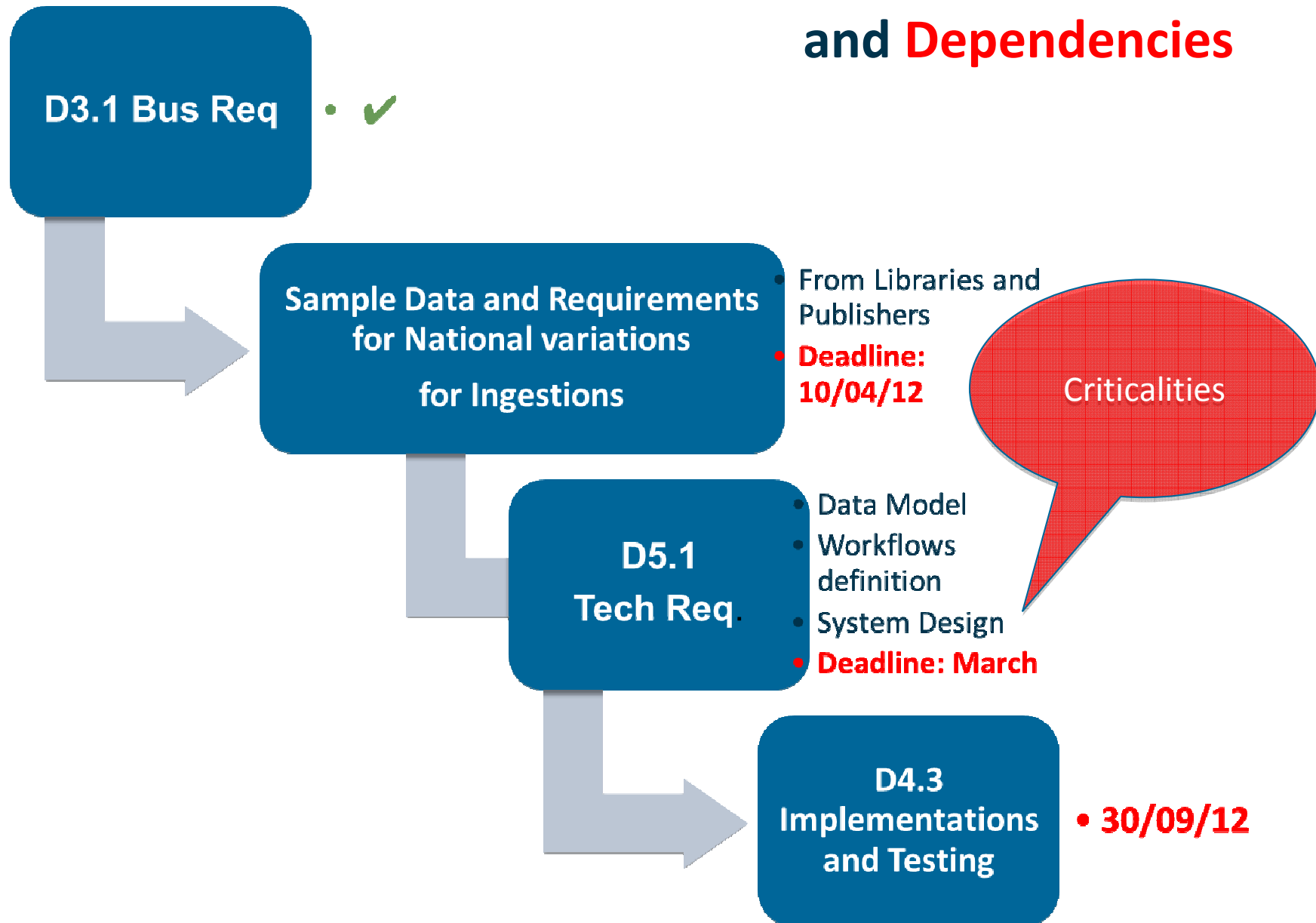
Implementation in TEL system of improvements identified in Validation Report from project ARROW

- Extract from MARC records, the contributors within the natural language text in *statements of responsibility*

  - Example *statements of responsability*:

    - *translated by Alexander Pope, with notes and introduction by the Rev. Theodore Alois Buckley ... and Flaxman's designs.*

    - *Ėrnst Teodor Amadej Gofman. Chudožnik Boris Tržemeckij. [Per. I. Tatarinovoj]*

    - *Ernst Theodor Amadeus Hoffmann. Mit Holzstichen von Andreas Brylka*

    - *Vicente Aleixandre ; estudio previo, selección y notas de Leopoldo de Luis*
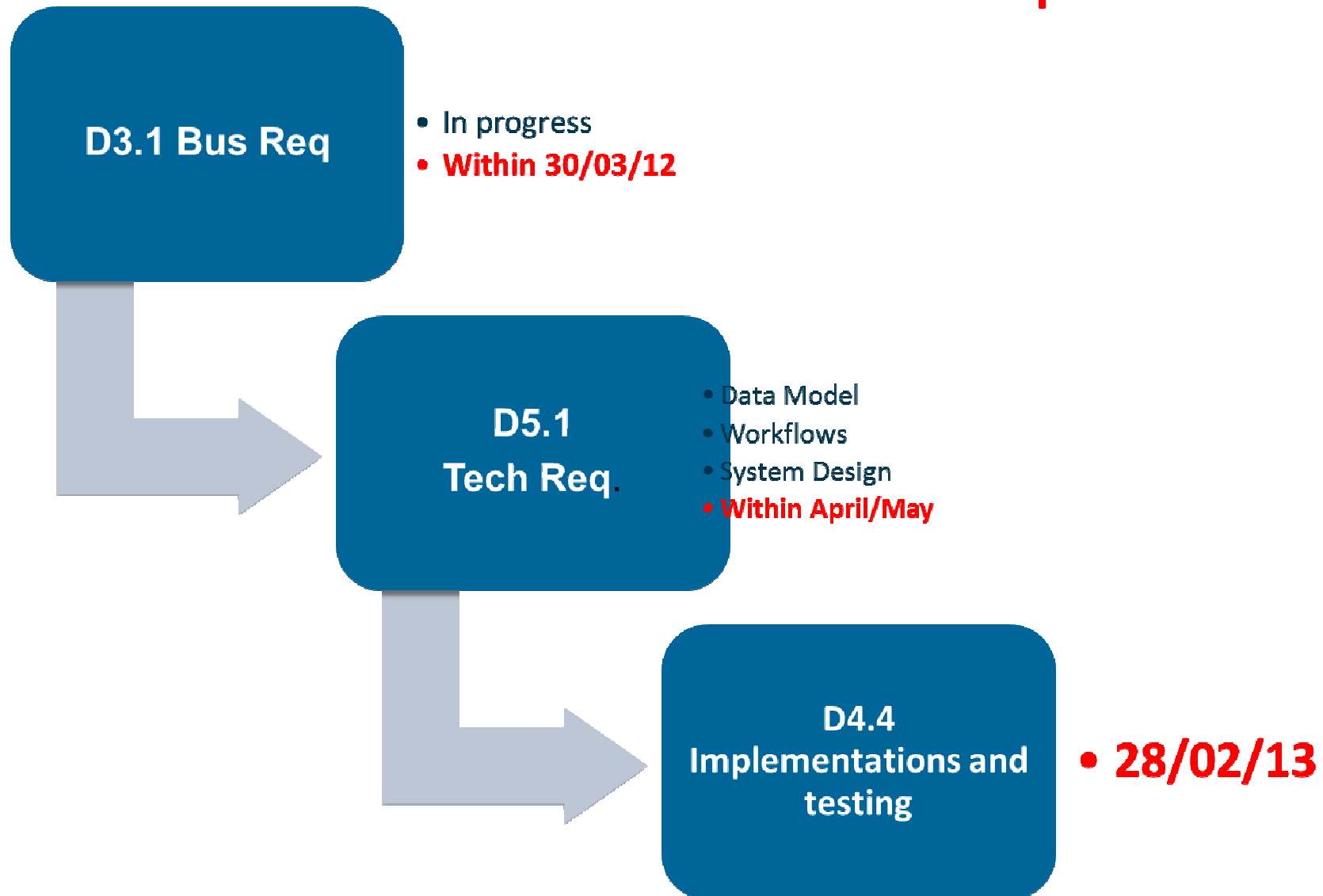
Implementation in TEL system of improvements identified in Validation Report from project ARROW

- Matching contributors of works with VIAF records of persons
  - Matching in VIAF is already available for VIAF participating libraries, which can export the contributors' ids (France, Germany and Spain)
  - Enhancements are necessary when:
    - The bibliographic record does not contain its identifier in the national authority file
    - Or the library does not participate in VIAF
  - Involves:
    - Searching in the VIAF data set for the contributor's name
    - Confirming/disambiguating matches by using the birth and death dates, known titles, and co-authors.

RRO Registries Tasks and **Dependencies**

D3.1 Bus Req

- In progress
- **Within 30/03/12**

D5.1 Tech Req.

- Data Model
- Workflows
- System Design
- **Within April/May**

D4.4 Implementations and testing

- **28/02/13**

WP5 Design of new registries

Brussels 20 March 2012

Current activities

- The BiP Business requirements have been finalised by WP3. They are the basis for the system design of the BiP database.

- On March 5 Cineca and MVB had a workshop that was meant to evaluate the BiP Business requirements with regard to feasibility.

- A first draft of the technical specification has been discussed. Cineca and MVB will work very closely together on that document in order to refine the specification.
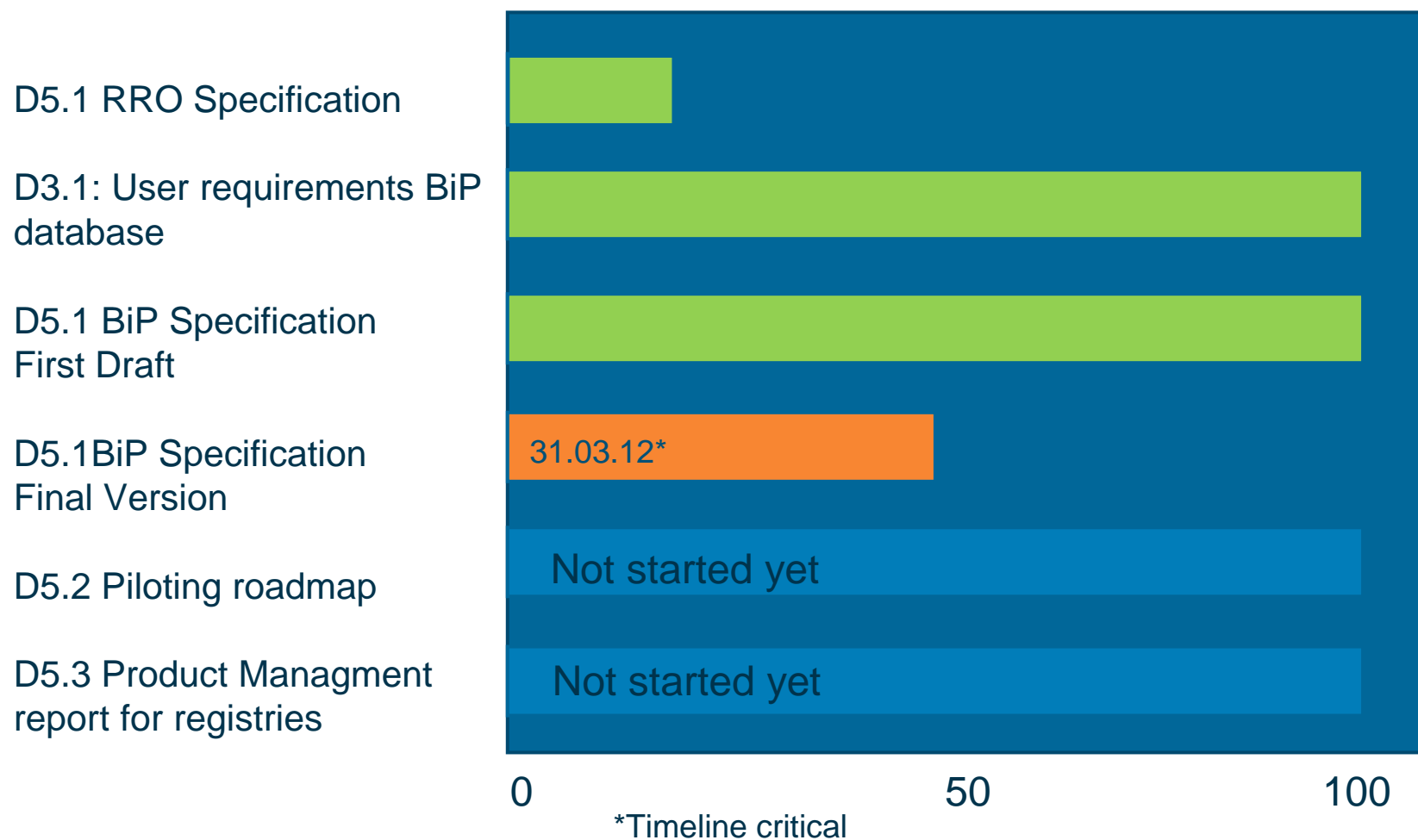
Current activities

- As it is very important to get a clear and detailed picture on the system design the process of refining the technical specification will take some time. Cineca and MVB have agreed to focus on singular technical aspects and work them out successively.

- The same procedure will be applied regarding the RRO technical requirements. Cedro will work out the technical specification for the RRO system supported by Cineca.

Status of deliverables

Work done (percentage)

D5.1 RRO Specification

D3.1: User requirements BiP database

D5.1 BiP Specification First Draft

D5.1BiP Specification Final Version — 31.03.12*

D5.2 Piloting roadmap — Not started yet

D5.3 Product Managment report for registries — Not started yet

0   50   100

*Timeline critical

The BiP system

Publishers

Import title data

- Onix-Import
- Excel-Import
- Manual Import
- Other formats

Booksellers
Libraries
Online Shops

Access title data

- Onix-Export
- Excel-Export
- Browsing the Web-Catalogue

WEB-Frontends

Data Entry
(Title Service)

Data Display
(Title Catalogue)

Data Input

Data Access

Backend

- Quality Control
- Reporting and statistics
- User and Rights Management
- Customised exports

Feedback and update on price and availability information

Customised reports
Serviceline
Login and account administration

ARROW

Next Steps and open issues

- Refine the technical document for the BiP system with a focus on data entry (title service)
  - Use cases
  - Role definition (who uses the title service for what purpose)
  - Creation of flow charts

- Define the data model of the BiP system

Next Steps and open issues

□ Get a clearer picture on national variations

- What data sources and formats used for initial loading will be provided (e.g. MARC21, MARCXML, CSV, ONIX)?

- Different cardinalities

- Us of different code values

- Test data files of participating countries

- What percentage of the records in the initial loading will include ISBN?

- How many new titles are published each year?

- What is your estimate of the number of titles currently "in print"?

Next Steps and open issues

□ Get a clearer picture on national variations

  ◻ Data policies and cataloguing rules

  Some open issues might need some clarification

  ◻ Quality control (duplicate check) regarding the fact that not all titles are ISBN related.

FURTHER INFORMATION

Pattock, Natascha (MVB)
Vogelbacher, Michael (MVB)
n.pattock@mvb-online.de
m.vogelbacher@mvb-online.de

Scipione, Gabriella (Cineca)
g.scipione@cineca.it

http://www.arrow-net.eu