



Quality Interoperability

The DL.org Quality Working Group

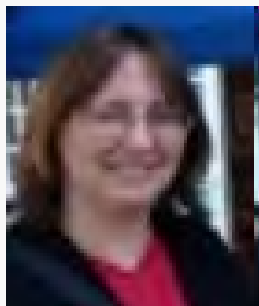
Sarah Higgins, Digital Curation Centre, University of Edinburgh, Quality WG Testimonial

Giuseppina Vullo, HATII, University of Glasgow, Quality WG Coordinator

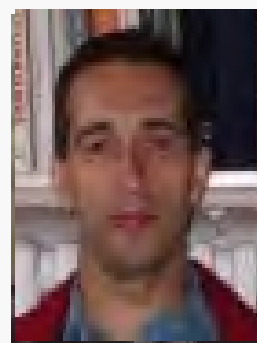
The Quality WG



Genevieve Clavel



Sarah Higgins



Sarantos Kapidakis



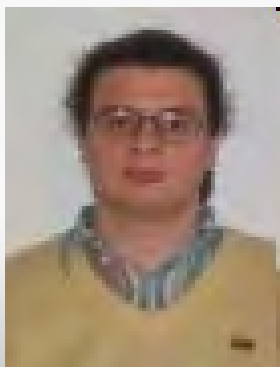
Giuseppina Vullo



Dirk Roorda



Seamus Ross



Nicola Ferro



Tefko Saracevic



Wolfram Horstmann

The Quality WG key features

- Start date: March 2009 (M4) - End date: July 2010 (M20)
Operational start date: May 2009 (M6)
- 7 external experts, 2 internal project staff members
- 8 projects/initiatives/institutions represented: AlfaLab (DANS), DRAMBORA, DRIVER, EUROPEANA, PLANETS, SHAMAN, TEL Plus, University of Edinburg DL
- WG Scientific Leader: Nicola Ferro, University of Padua
- WG Testimonial: Sarah Higgins, Digital Curation Centre, University of Edinburgh
- WG Coordinator: Giuseppina Vullo, HATII, University of Glasgow

The Quality WG Mission

- Identification of main **interoperability** issues and select the most urgent from the Quality perspective
- To discuss the **possible approaches** to identified issues of Quality Interoperability
- Elaboration of **effective patterns** and identification of **best practices**
- Refinement & enhancement of main **DL concepts** w.r.t. Quality

The Quality WG Scope

- adopts the **DELOS Reference Model** as its conceptual framework and the definition of *Quality*
- elaborates **interoperability requirements** between Digital Libraries with respect to the *Quality Parameter*
- works on the **Quality Core Model**, which will be applicable to a broad range of DLs, investigating the definitions, the relationships and the examples of selected Quality parameters (*Generic Quality Parameter, Content Quality Parameter, Policy Quality Parameter*)
- investigates the *Quality measurement* considering the Digital Library as an **organisation** that covers the existing levels of *Digital Library, Digital Library System* and *Digital Library Management System*
- identifies and collects **best practices** in view of setting effective guidelines to promote quality interoperability
- provides suggestions for **a standard vocabulary** for quality interoperability
- given the high significance of **Policy** within the Quality Core Model, collaborates closely with the Policy Working Group

What is Quality?



Quality is something which makes the difference

Quality Interoperability key-issues

- Definition of Quality
- Different approaches to Quality: quality of content, quality of services, quality of policies
- Quality Interoperability, i.e. how different DLs can share a common Quality framework

The Quality WG preliminary findings

- **organisational approach to the DELOS Digital Library Reference Model:** the Quality WG recommends considering an additional level that is termed “Organisation”, wrapping the existing Levels of Digital Library, Digital Library System and Digital Library Management System. There is an organisation beyond a Digital Library that defines the policy of the overall system, in which a Digital Library is operating
- **interdisciplinary research:** relevant studies on Digital Libraries quality are taking place within LIS, computer science, HCI
- **development of a Quality Core Model:** in order to broaden the applicability of the Quality framework within the DELOS Digital Library Reference Model, the Quality Working Group identified some selected aspects to define and develop objectives and criteria for their evaluation

The Quality Core Model

Based on the RM Quality concept map, “*the **Quality Core Model** is thought to be most characteristic for DLs and shall help to identify best practices. This simplified pattern should help DLs to interoperate in the quality domain*”
(From: Testimonial of the Quality Working Group, Tirrenia, 3 July 2009)

Phase 1.

Analysis of Quality within the DELOS RM

Selection of core quality parameters

Investigation of the selected parameters

(definitions, position within the RM, user scenarios, key issues)

DELOS RM enhancement

The Quality Core Model

Based on the RM Quality concept map, “*the **Quality Core Model** is thought to be most characteristic for DLs and shall help to identify best practices. This simplified pattern should help DLs to interoperate in the quality domain*” (From: Testimonial of the Quality Working Group, Tirrenia, 3 July 2009)

Phase 2.

Proposing (identifying) **best practices**

Understanding (testing) **feasibility**

Suggesting (setting) a **Quality Interoperability framework**

Definition of Interoperability

The Quality Working Group (QWG) adopts two **definitions of interoperability**:

- IEEE – “the ability of two or more systems or components to exchange information and to use the information that has been exchanged”
- ISO/IEC 2382-2001- “the capability to communicate, execute programs or transfer data among various functional units in a manner that requires minimal knowledge of the unique characteristics of those units”

Definition of Quality

The QWG adopts the ISO 9000:2005 **definition of quality**:

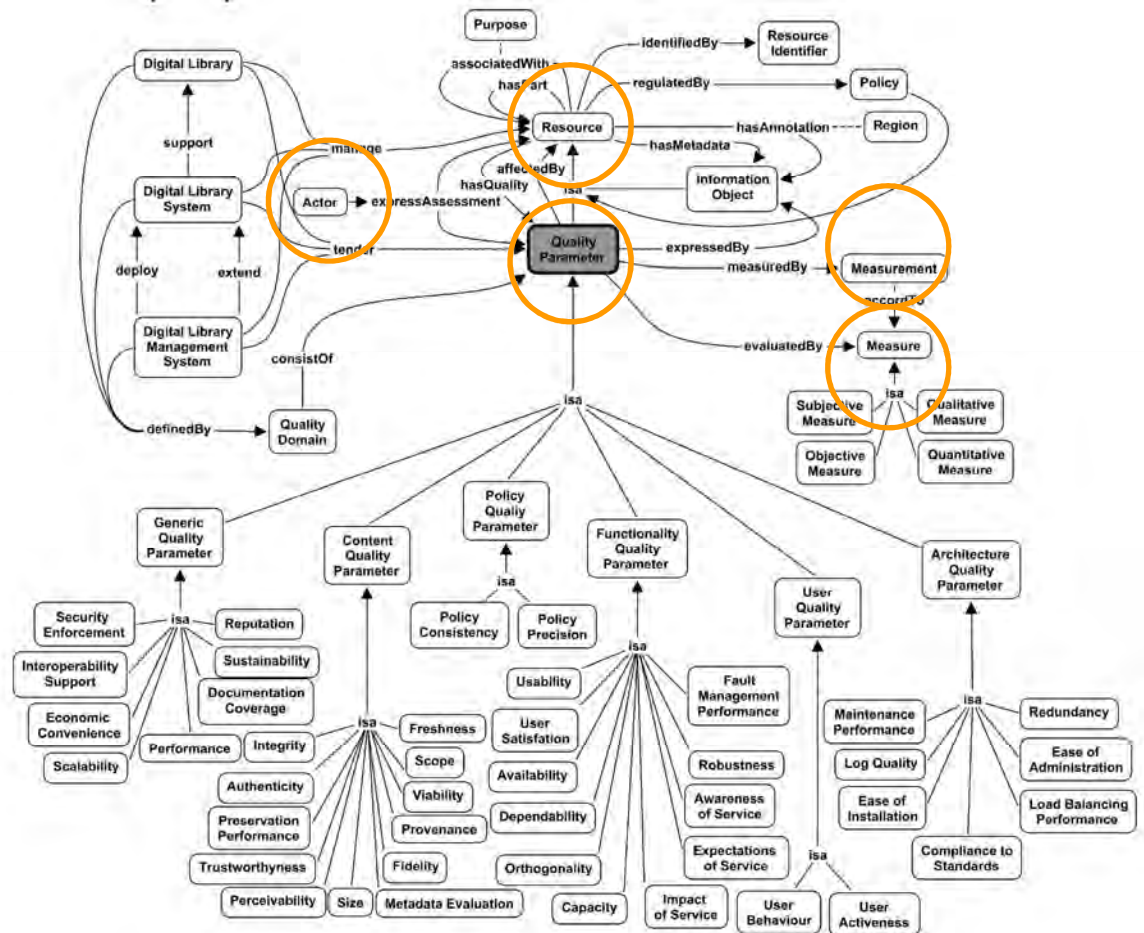
“the degree to which a set of inherent characteristics fulfils requirements”

Where:

Inherent characteristics =
Resource + DRM Quality
Parameter

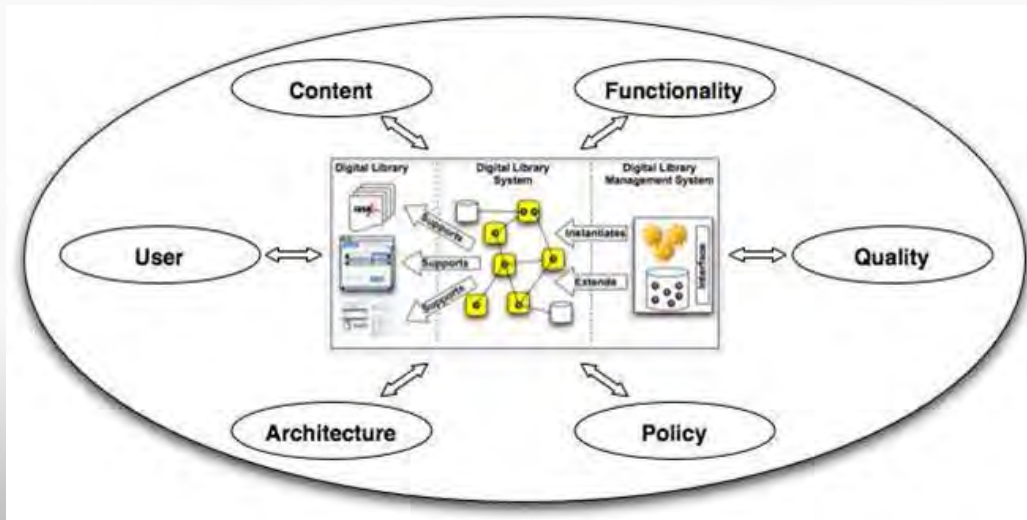
Degree of fulfilment =
Measure + Measurement

Requirements = assessment
expressed by Actor



Delos Digital Library Reference Model

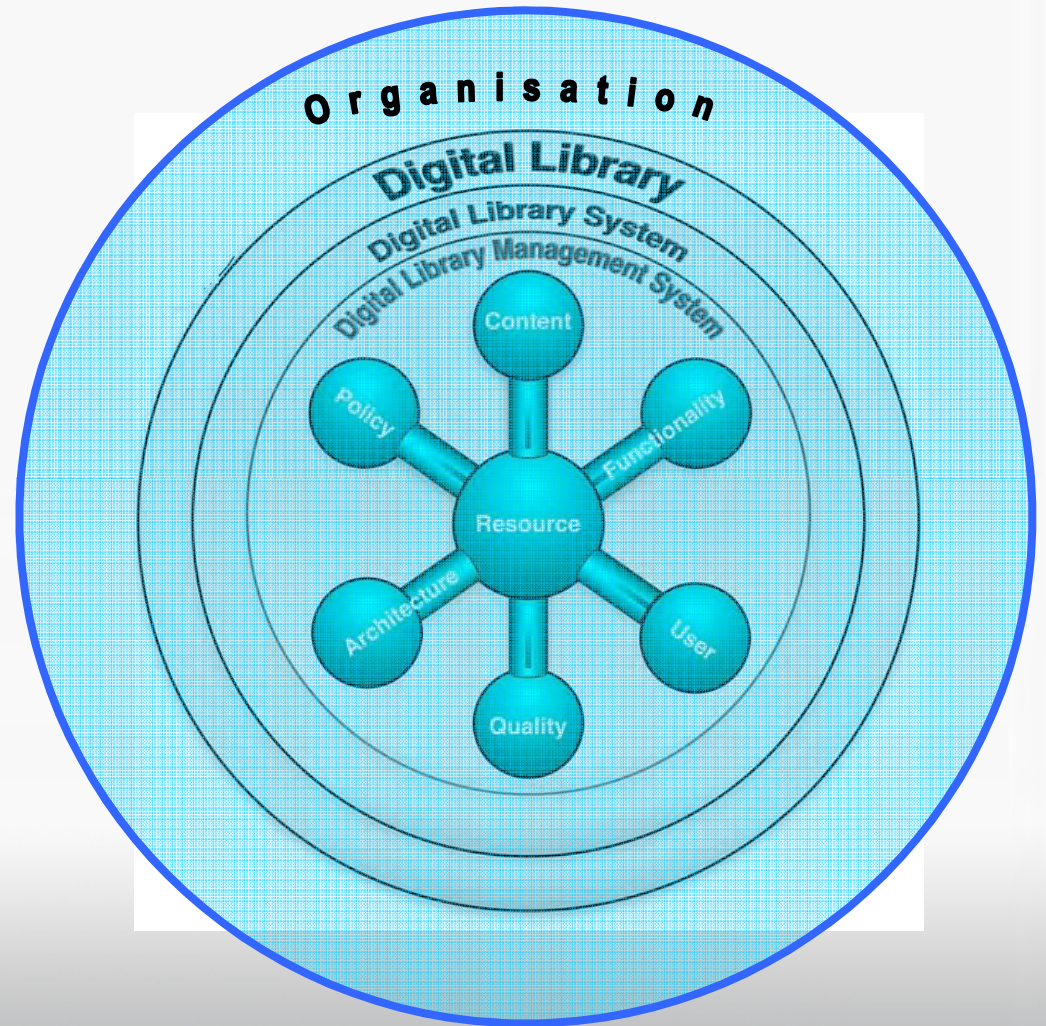
The QWG adopts the
Delos Digital Library Reference Model (DRM)
as a conceptual framework for elaborating the
“Quality of a Digital Library that supports interoperability”
(not “quality of interoperability between Digital Libraries”)



Organisational Issues

But recognises that a Digital Library may operate within an **Organisation** which defines over-arching **policies** (not necessarily specific to Digital Libraries) which affect interoperability *eg*:

- Subject community
- University



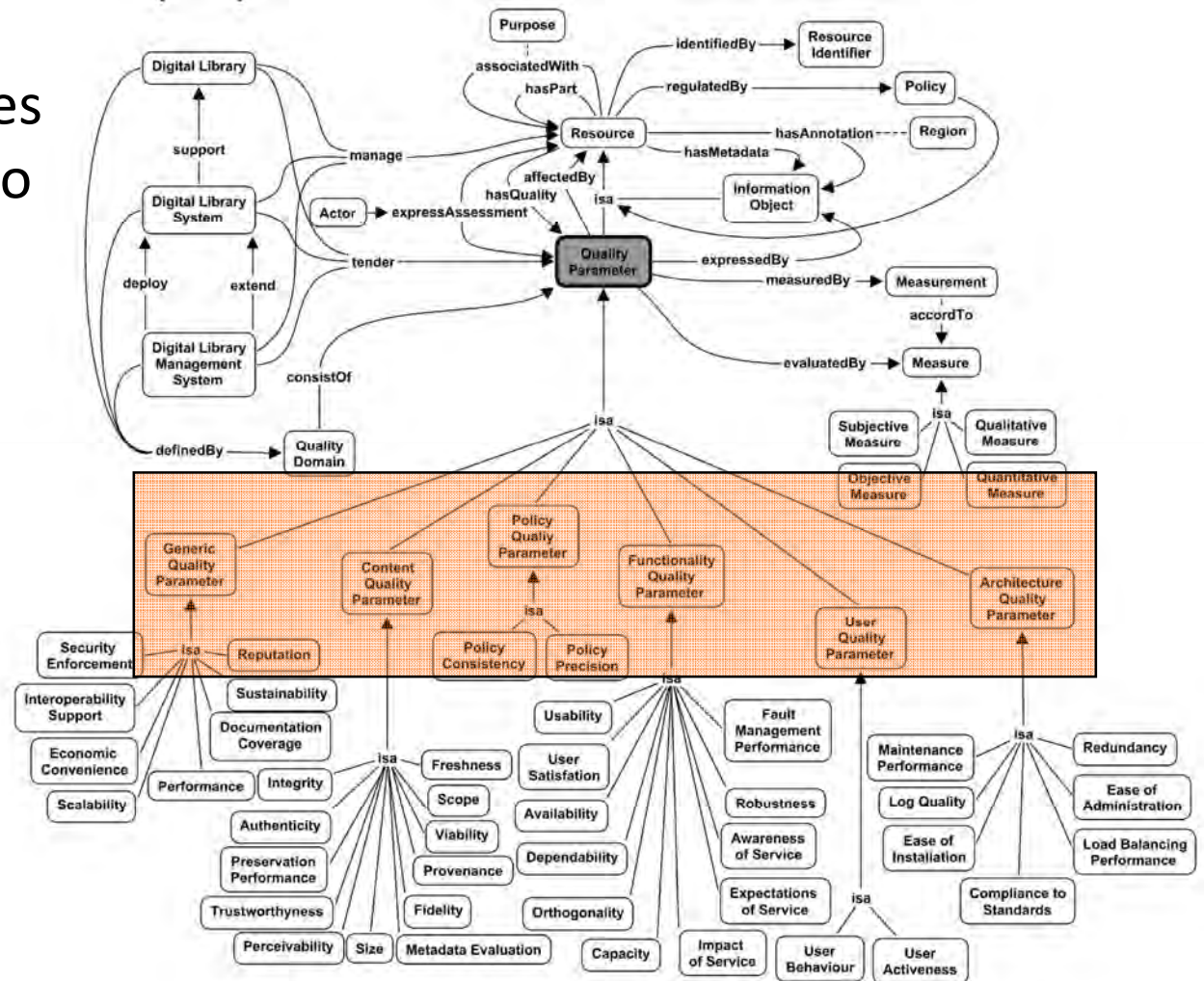
Quality Measures

- The core business of a Digital Library is identified as collection management
- To support interoperability a Digital Library requires an acceptable **Quality Measure** and needs to pass a quality threshold

Quality Measures

All parameter classes may be applicable to a given **Quality Measure**:

- Generic
- Content
- Policy
- Functionality
- User
- Architecture



Quality Core Model

The QWG identified an **Application Profile** of the DRM Quality Parameter which are essential to:

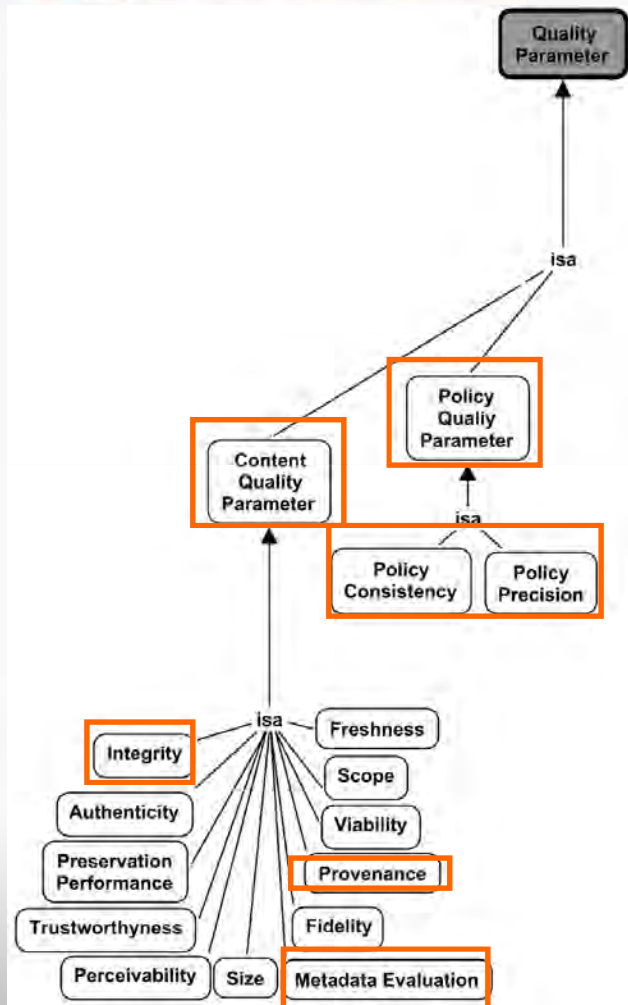
- the nature of a Digital Library
- interoperability across Digital Libraries
- most characterises the parameters needed for a **Digital Library Interoperability Quality Measure**

The identified **Application Profile** has been called **Quality Core Model**

Quality Core Model

Most important for a DL Interoperability Quality Measure are:

- Policy
 - Policy consistency
 - Policy Precision
- Content
 - Integrity
 - Provenance
 - Metadata



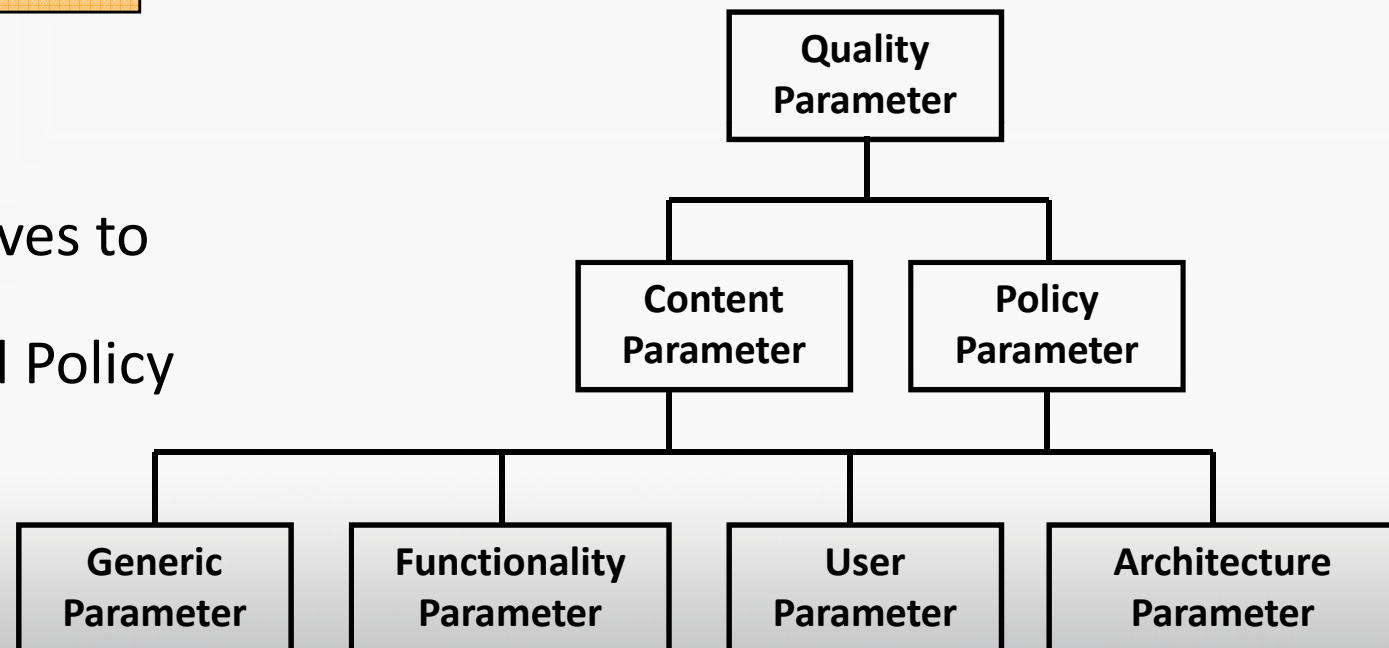
Quality Core Model

- Functionality
- Architecture
- User

- Depend on the content for quality
- Can be adapted with quality interoperable content

- Generic

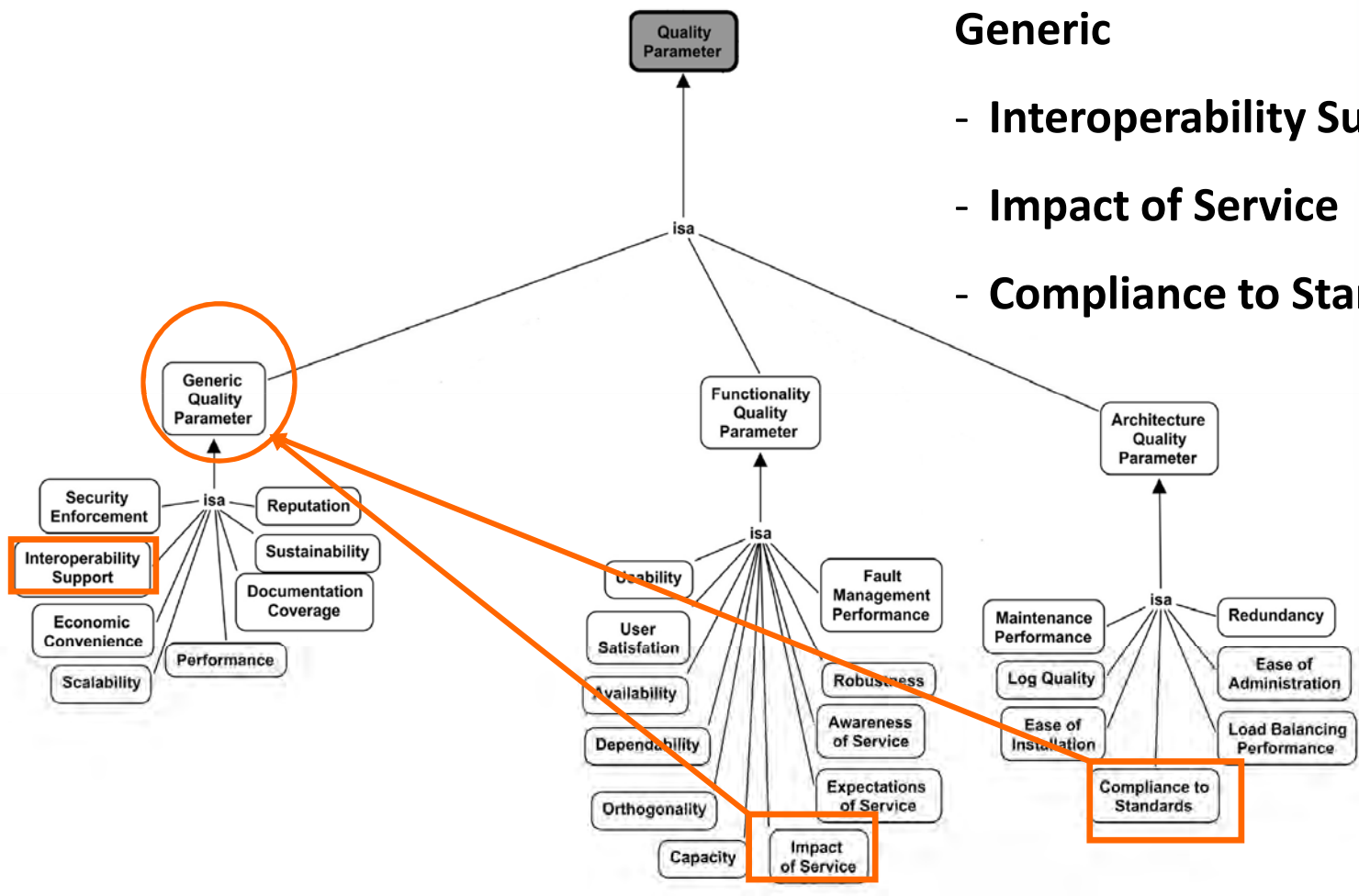
....are all slaves to
Content and Policy



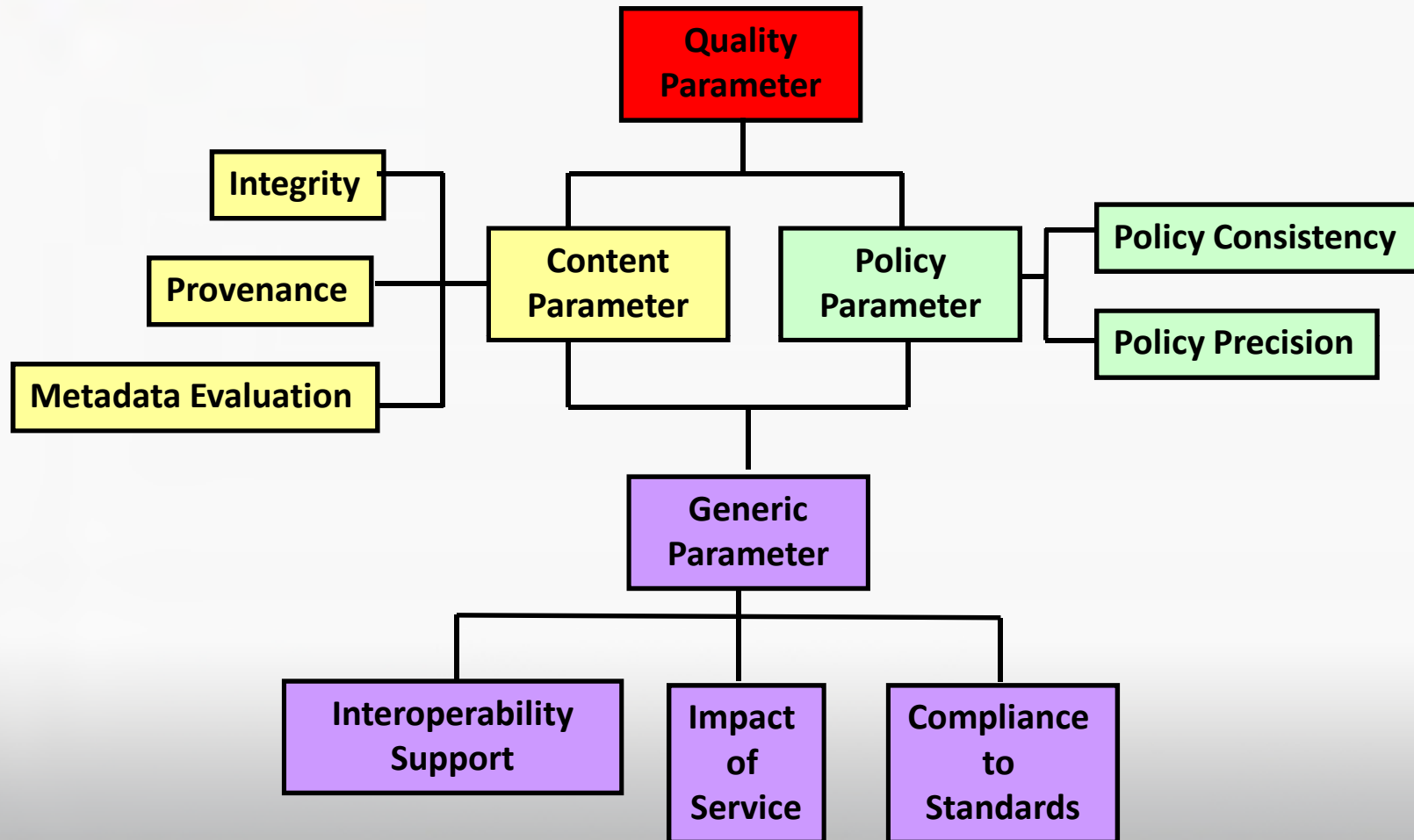
Quality Core Model

Generic

- Interoperability Support
- Impact of Service
- Compliance to Standards



Quality Core Model





Generic Parameter: Interoperability Support

Approaches to interoperability:

- Define generic interchange protocols – OAI-PMH
- Set up research infrastructures which define a framework for participants – CESSDA, CLARIN, DARIAH

Possible parameters:

- OAI-PMH compliance
- Use of persistent identifiers
- Metadata specifications
- Authorisation and authentication procedures
- Licences
- Continuity of service

Related to:

- Compliance to standards

Generic Parameter: Compliance to Standards

- Quality interoperability depends on the extent a DL adheres to a set of pre-determined rules or codes, which include:
 - Data / content standards
 - Metadata standards
 - Web interface standards
 - Data sharing protocols
- Which framework to adopt depend on the community or discipline involved
- Establish a measurable standards compliance agreement
- Related to:
 - Interoperability support
 - Sustainability



Generic Parameter: Impact of Service

Impact of service can be measured by:

- Increase of user knowledge
- Improvement in DL practical skills over time



Content Parameter: Integrity

DL's Information Objects:

- Sufficient breadth, depth, scope to achieve aims
- Completeness, accuracy

Related to:

- Metadata integrity
- Policy consistency
- Regular content update
- Accurate format migrations

Content Parameter: Integrity

User scenario

Collection of journal articles:

- Does the final version of each article appear in DL?
- Are all the pages and figures available?
- Does the scanning quality mean that all pages are clear?
- Has OCR scanning been proof-read and corrected
- For merged collections:
 - Is there only one entry in the catalogue
 - Have all entries copied correctly
 - Does the collection only contain what is expected?

Content Parameter: Provenance

- Tracking origins and history of the Information Object to know if it is fit for purpose:
 - Transformations? Cleaning? Rescaling? Modelling? Mergers?
 - Authorship, IPR, integrity and authenticity
- Issues for quality provenance information:
 - metadata standards for tracking provenance?
 - How to capture
 - What to capture
- Related to: Metadata, Annotation, Preservation Policy

Content Parameter: Metadata Evaluation

- Metadata evaluation should measure the **support for digital items** against the Content Quality Parameters.
- Metadata evaluation should look the support in all classes of metadata:
 - Descriptive, Technical, Administrative, Use, Preservation
- Evaluation of metadata for:
 - Use of structure standards
 - Use of content standards
 - Metadata creation
- Related to: Content Quality Parameter , Policy Quality Parameter, Compliance to Standards, Interoperability Support, Scalability, Sustainability

Content Parameter: Provenance

User scenario

A bioinformatics DL, which supports the analysis of gene expression and analysis, requires tools to be applied to the raw data in a defined workflow.

Are the following maintained?

- Results of workflow
- Intermediate steps of the workflow
- Configuration of tools and algorithms

Policy Parameter: Policy

- **Policy consistency** - free of contradictions eg consistency across *Digital Rights Policy* and *Digital Rights Management Policy*
- **Policy Precision** – policy detailed and defined enough to constrain behaviours, deal with consequences and enforce:
 - Envisage aspects of governance
 - Sufficient knowledge of technology – architecture and software

Summing Up

- **Quality interoperability** firstly means the possibility for DLs to share a common qualitative framework
- The QWG is focusing on a **set of quality parameters** considered essential within this framework
- **New research** and **best practices** are needed



THANK YOU

<http://www.dlorg.eu/>

workinggroups.wiki.dlorg.eu/index.php/Quality_Working_Group

quality@dlorg.eu; g.vullo@hatii.arts.gla.ac.uk