

MTA SZTAKI

Department
of Distributed
Systems

**The evaluation computer:
a model for structuring evaluation activities**

László Kovács
András Micsik

Our goals

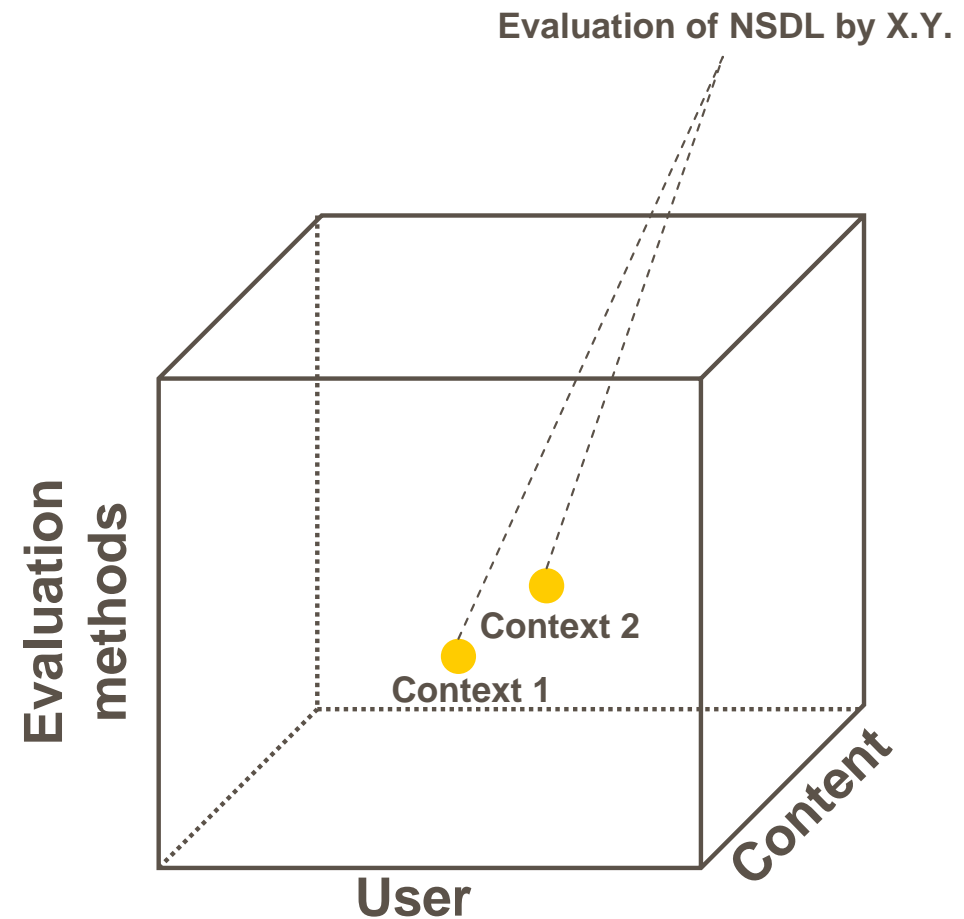
- Analysis of relation between evaluations
 - Are they comparable or not
 - How many common parts they have
 - „Distance” between evaluations
- Building visualisations of multiple evaluations
 - Provide overview of activities
 - Locate white spots

Structure of DL Evaluations

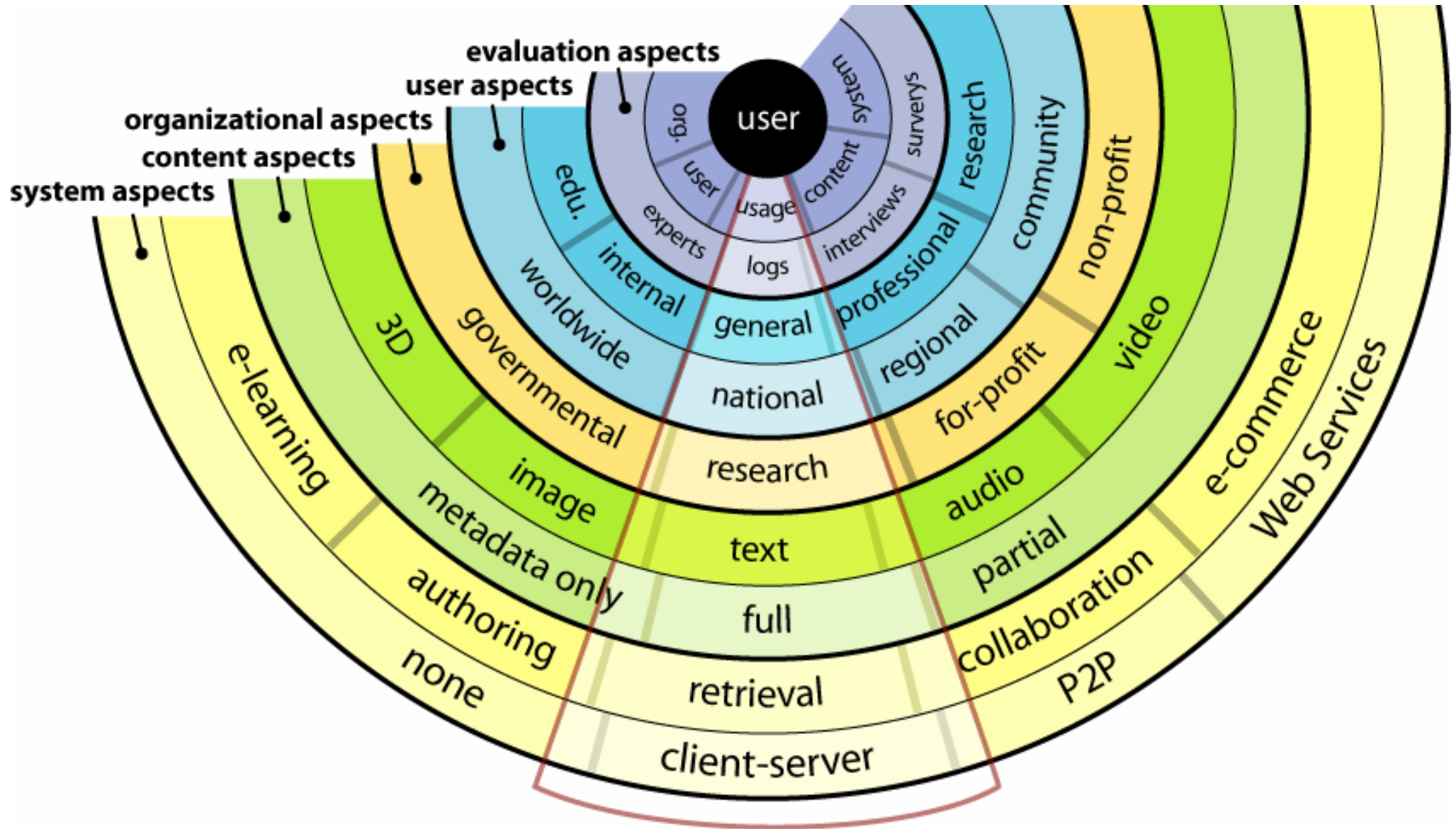
- Typical elements of evaluations
 - Defining the context of evaluation
 - e.g. image collection for research
 - General metrical outcomes
 - e.g. avg. response time, availability
 - Context-sensitive metrical outcomes
 - e.g. avg. size of images
 - Analysis (textual)
 - e.g. summary of user feedbacks

Classification of evaluations

- Evaluation atoms
- Visualisation of evaluation contexts
- N-dimensional space
- Top-down approach:
 - Defines metrics
 - ...research to be done



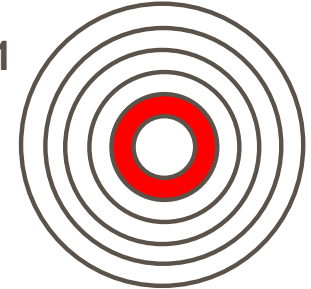
The “evaluation computer”



Rings of the evaluation computer

EVALUATION aspect

Ring 1



Target:

Usage
User
Content
System
Organization

Type:

Summative
Formative
Iterative
Comparative

Measure:

Usability
Satisfaction
Performance
Economic value
Educational value
...

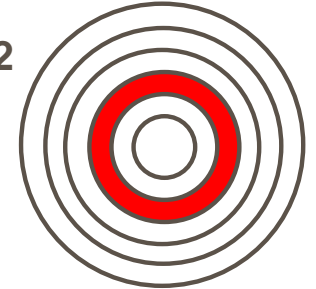
Method:

Surveys
Focus groups
Log analysis
...

Rings of the evaluation computer

USERS aspect

Ring 2



Type:

Internal
 General
 Education
 Professional
 Research
 Children
 Senior
 ...

Number:

0 – 100
 100 – 1000
 1000 – 10000
 10000 – 100000
 100000 –
 ...

Distribution:

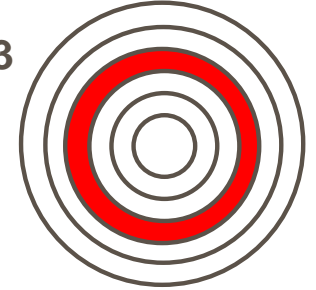
Community
 Regional
 National
 Worldwide
 ...

...

Rings of the evaluation computer

ORGANIZATIONAL aspect

Ring 3



Location:

Single place
Distributed

...

Profit:

For-profit
Non-profit
Self-organizing
Governmental
Research

Services:

Trad. Library
Digital Content Mgmt.
Unified interfaces
Mediating Content

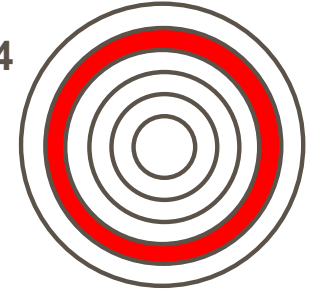
...

...

Rings of the evaluation computer

CONTENT aspect

Ring 4

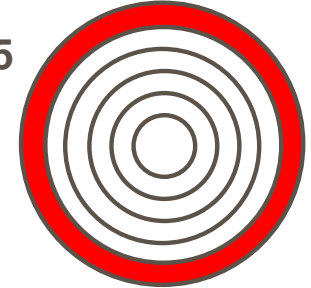


Type:	Number of:	Extent:	
Text	0 – 1000	Full	
Image	1000 – 100000	Partial	
Audio	100000 – 1000000	Metadata only	...
Video	1000000 –	...	
3D			
...			

Rings of the evaluation computer

SYSTEM aspect

Ring 5

**Communication:**

None
Client-server
P2P
WebServices
...

Multilinguality:

None
UI only
Cross-language

Functionality:

Retrieval
Authoring
Collaboration
E-commerce
E-learning
...

Repository:

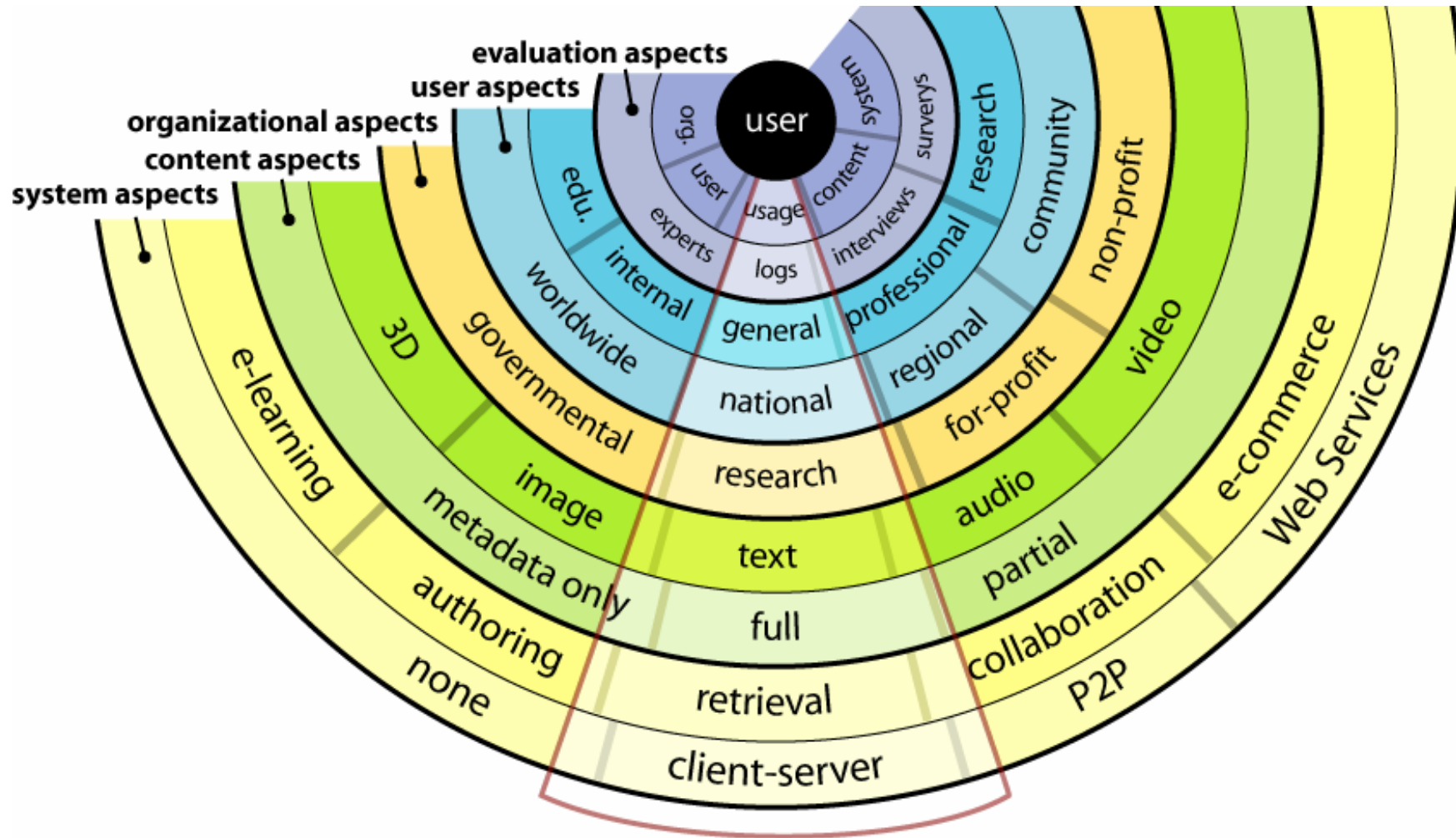
DBMS
Object store
XML DB
Other IR
Multimedia
Proprietary
...

...

Evaluations

- Evaluate a single service
 - Justify, improve, plan, etc.
 - Evaluate iteratively
- Evaluate multiple services or scenarios
 - Comparison, ranking, etc.
- Experimental evaluation
 - Try new evaluation methods or new evaluation targets

Defining evaluation context

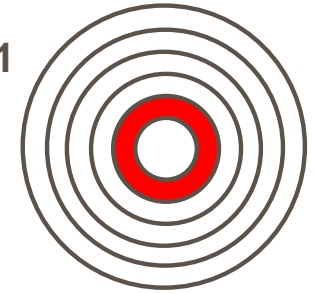


Steve Jones, Sally Jo Cunningham, Rodger McNab:
An Analysis of Usage of a Digital Library,
ECDL 1998

Comparison of evaluations

EVALUATION aspects

Ring 1



Target:

Usage
User
Content
System
Organization

Type:

Summative
Formative
Iterative
Comparative

Measure:

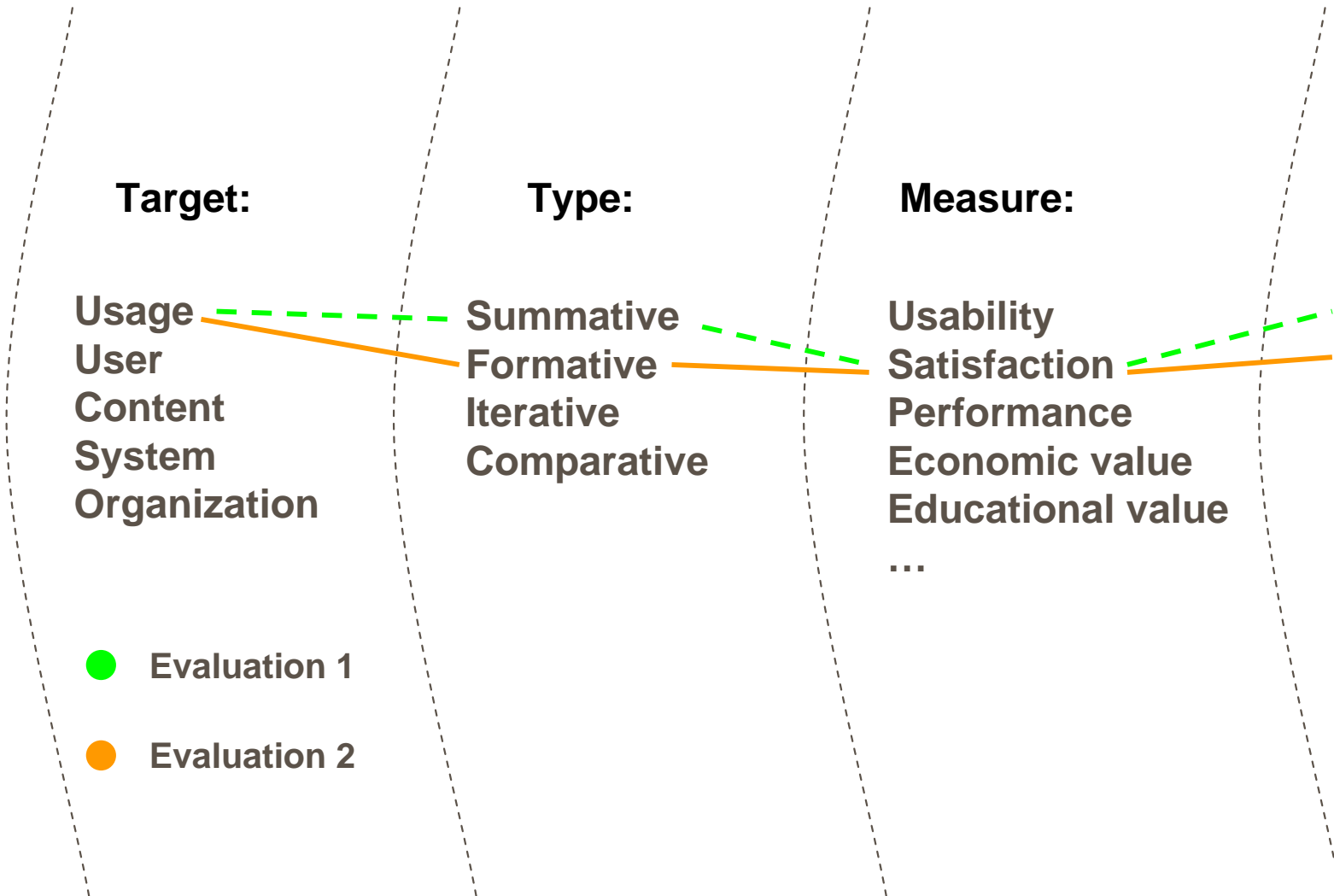
Usability
Satisfaction
Performance
Economic value
Educational value
...

Method:

Surveys
Focus groups
Log analysis
...

● Evaluation 1

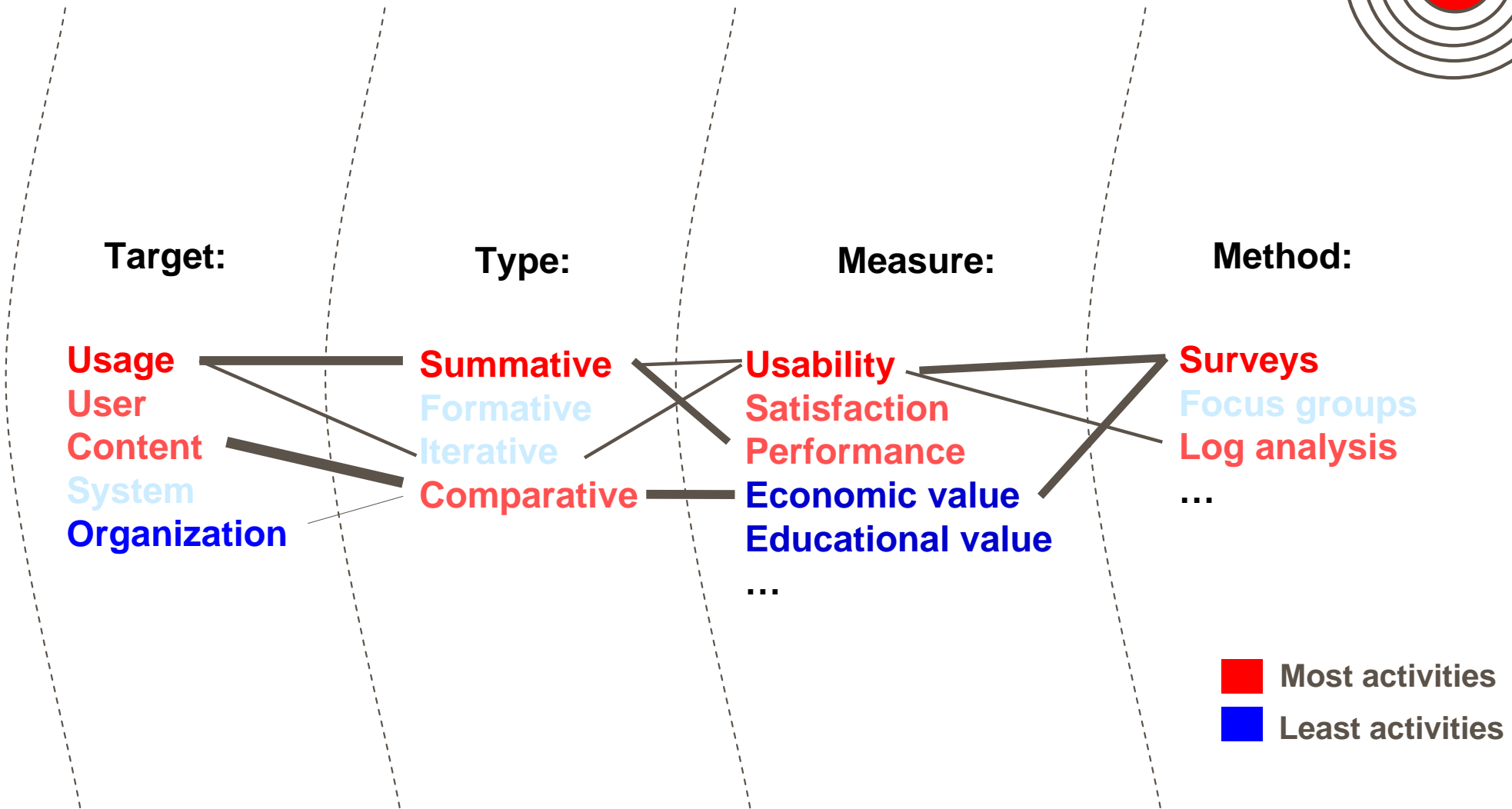
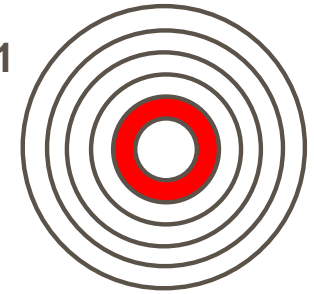
● Evaluation 2



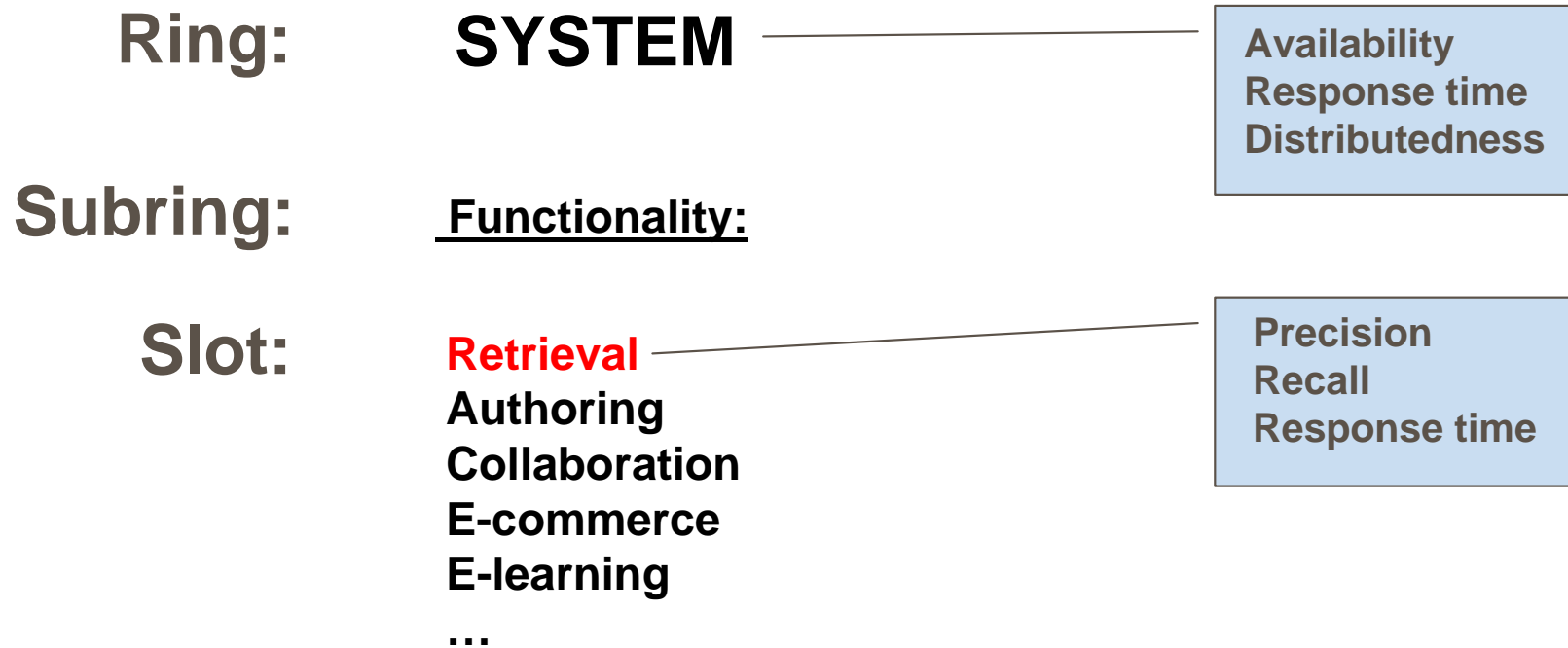
Overview of evaluation activities

EVALUATION aspects

Ring 1



Generation of evaluation parameters



Parameter *or* criterion?

Systematic planning of evaluations

**Same
evaluation
method**

**Similar user
groups**

**Similar
content**

**Same
functionality**

**Different
repository
technologies:**

**DBMS
Object store
XML DB
Other IR
Multimedia
Proprietary**

...

Summary

- The “evaluation computer”
 - Calculates atomic evaluations
 - Defines distance between evaluation atoms
 - Provides overview of DL evaluation scientific field
 - Helps to identify new research areas
 - User-centric approach
 - Helps to visualize the activities and research done in the field
 - Generates parameters for evaluation