

City-Zen Open linked data for more Visible Historical and Cultural Heritage

Prof. Basma Makhoulf Shabou

University of Applied Sciences of Western
Switzerland

HES-SO GENEVA

Prof. Maria Sokhn

University of Applied Sciences of Western
Switzerland

HES-SO Valais Wallis

Archival preservation and assessment: Relevant Scientific & professional Initiatives
ALA-ICA 2017 – Mexico City

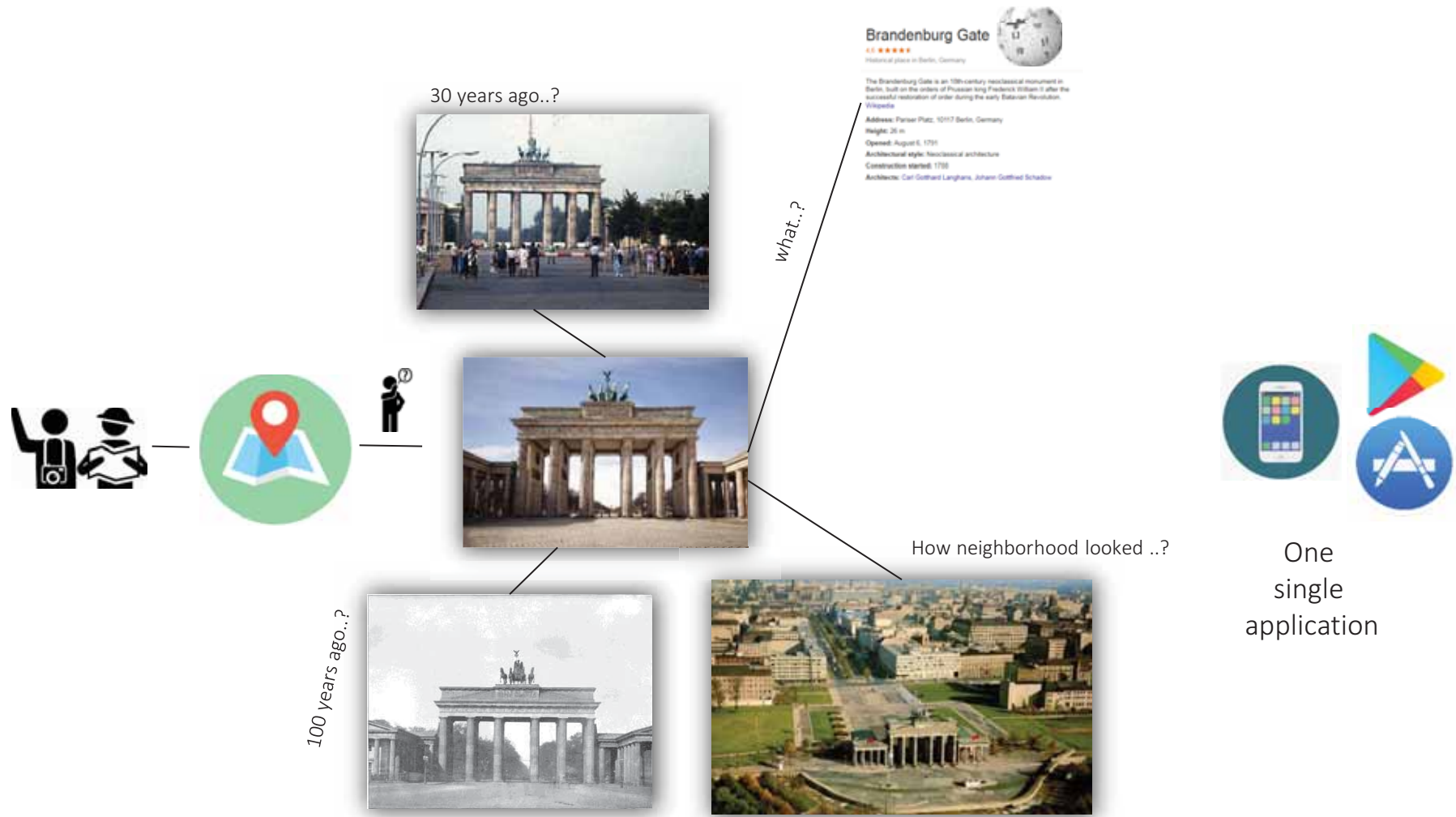
Outline

- Context
- Motivation
- Challenges
- City-Zen system
- Conclusion & Future work

Context

- Lead
 - University of Applied Sciences (HES-SO Valais - Geneva)
- Partners
 - HES-SO IS-NET
 - Hasler Foundation
 - State archives of Valais Switzerland
 - Digital Switzerland

Motivation



Motivation

- Information systems
 - Mostly rely on up-to-date content on attractive places
 - Additionally use multimedia digital archives
 - Historical data: photos, films, documents
- Interesting to explore evolution of cultural interests over time



Challenges (1/3) - Multimedia retrieval

- Support different queries

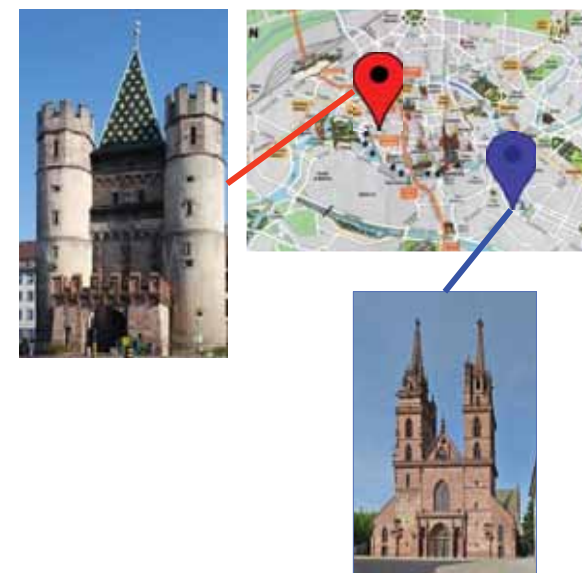
Simple location queries



Temporal queries



Spatial queries



Multimedia similarity queries

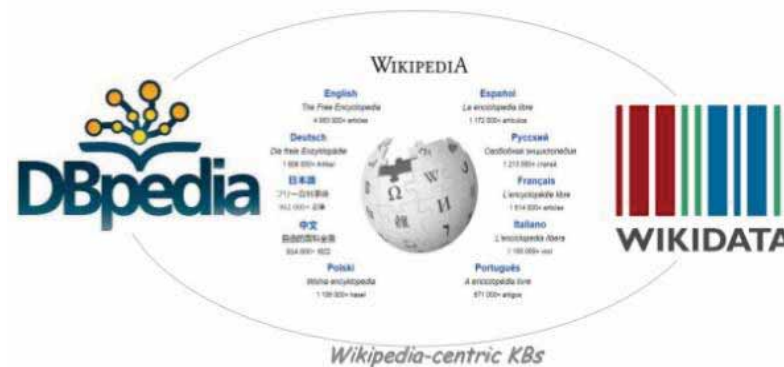


Radar mode



Challenges (2/3) - Entity recognition and Linking

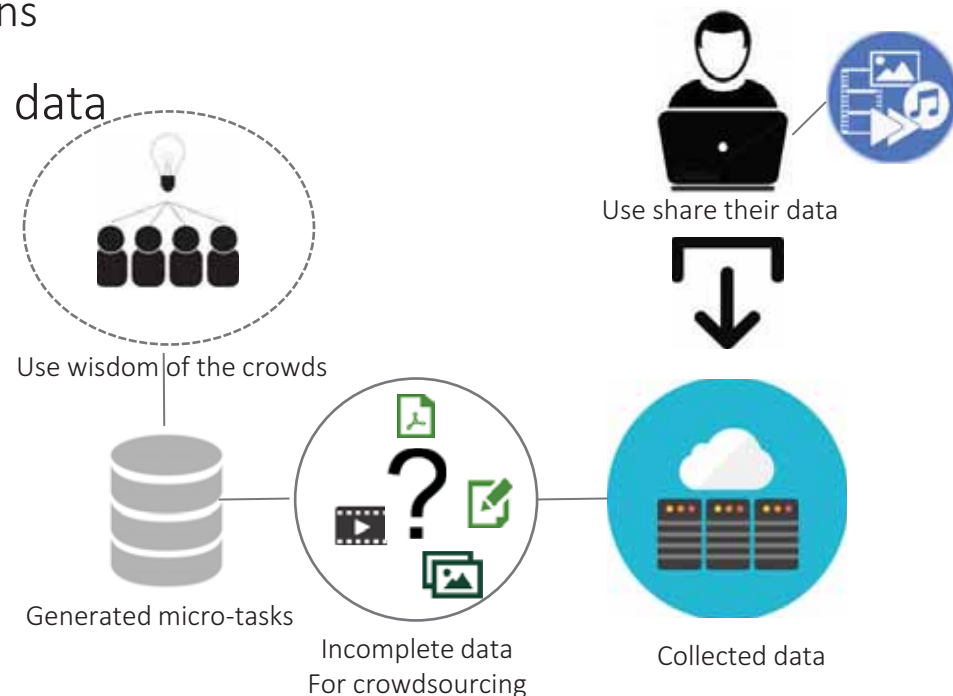
- Different sources and formats
 - Schema and data integration
- Entity recognition
- Linking with external data sources: e.g. DBpedia



Challenges (3/3) – Crowdsourcing & Quality control

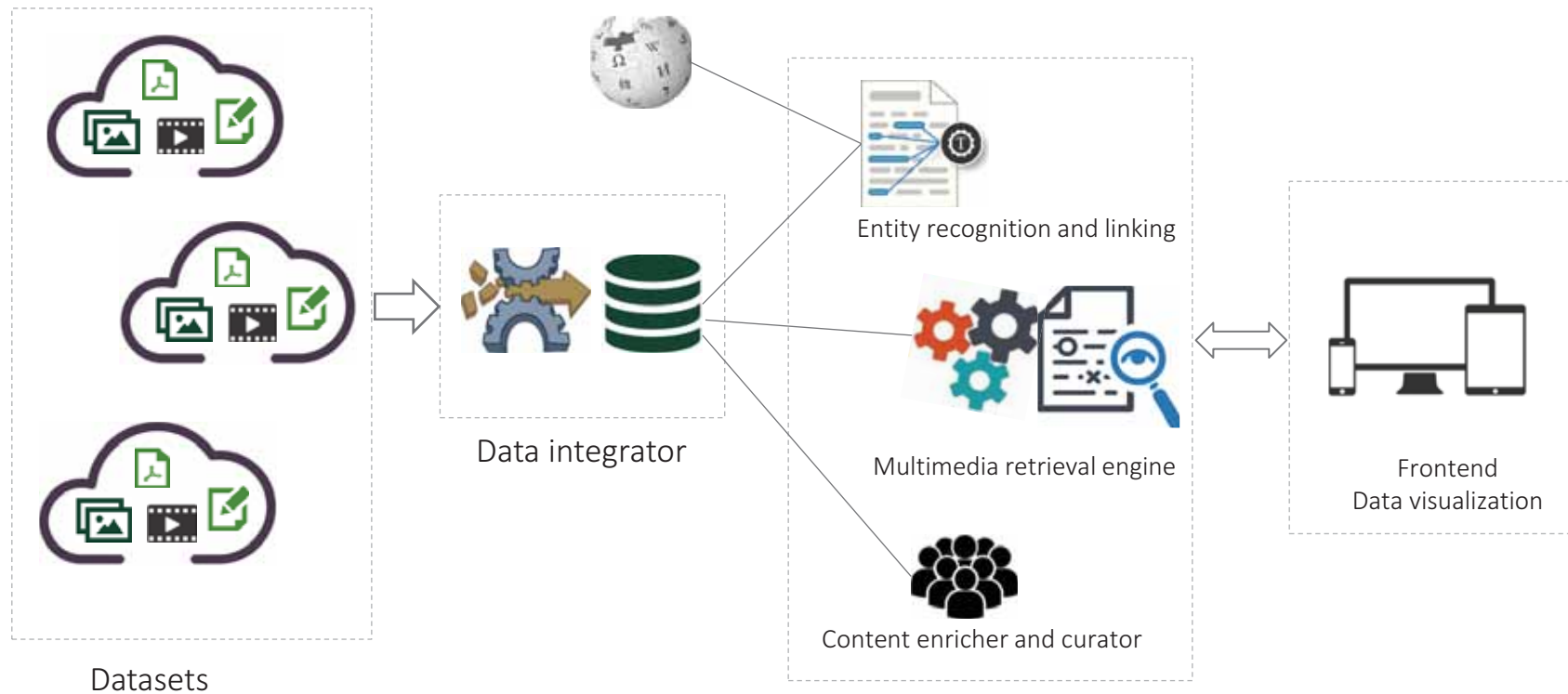
- Enable provision of user content data
- **Motivate** participation
 - Incentives
 - *intrinsic* or *extrinsic* motivations
- Control the **quality** of crowdsourced data

altruism
reputation
freedom reciprocity
self-expression
competition
community
autonomy
fun



City-Zen high level system design

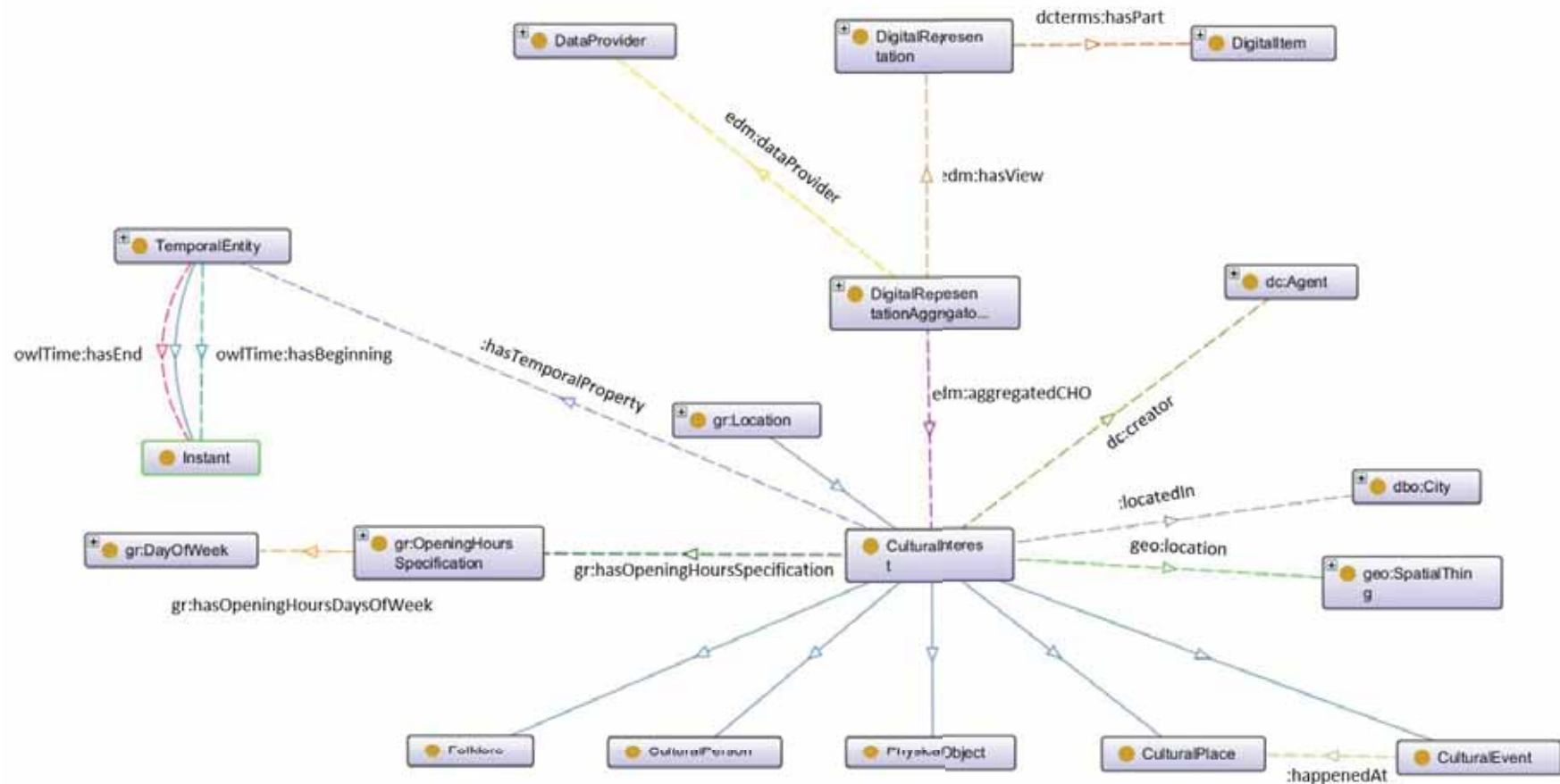
Provide integrated interactive access to heterogeneous contents



Dataset

- Uses metadata
 - 2 datasets (*DigitalValais* and *MediathequeValais*)
 - Three languages (French, German and Italian)
 - Title and description

Dataset - Model



Dataset - Model

Europeana

Digital
Concept

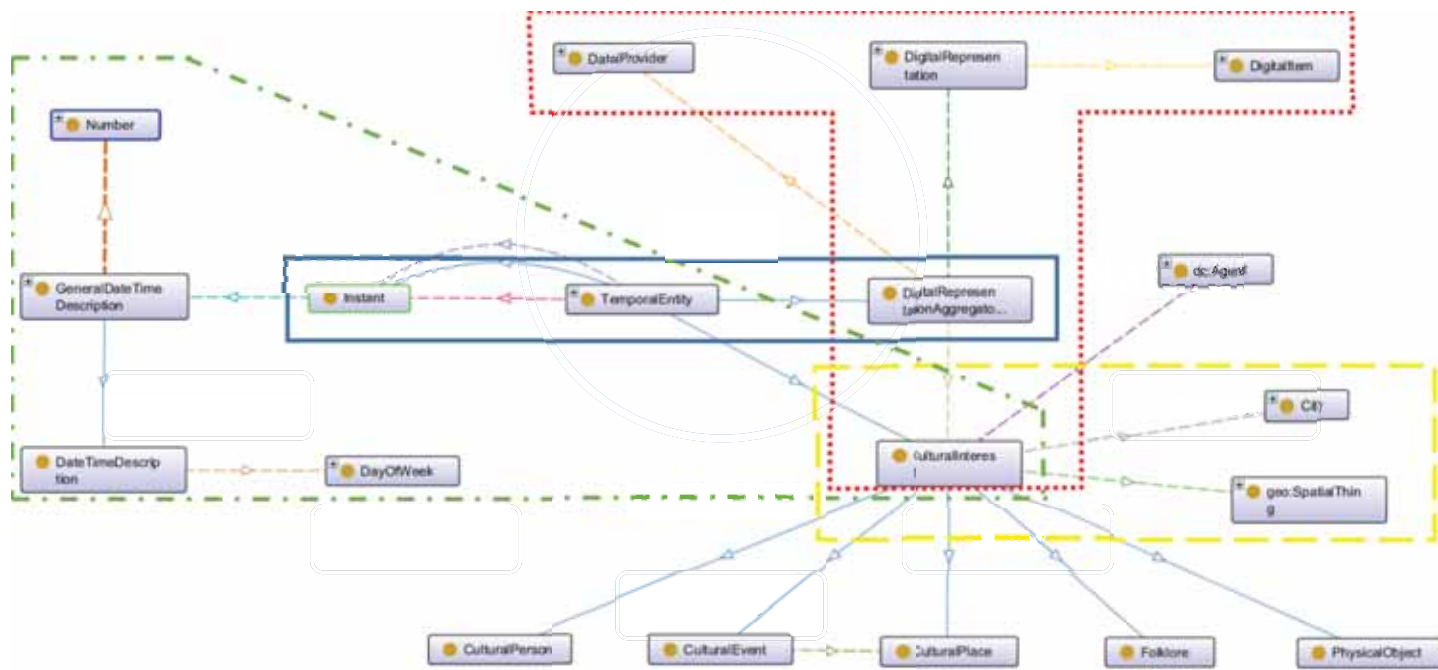
OWL Time

Timing
Concept

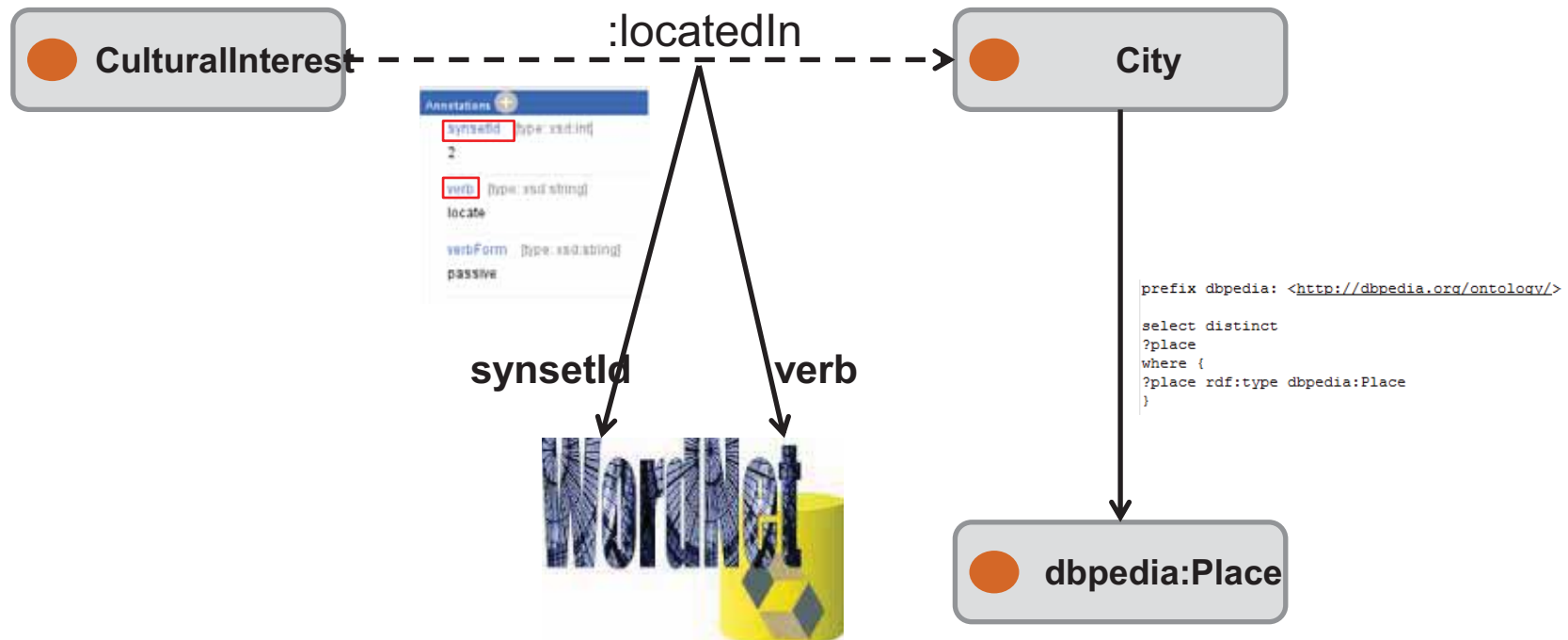
Planning
Concept

Geonames

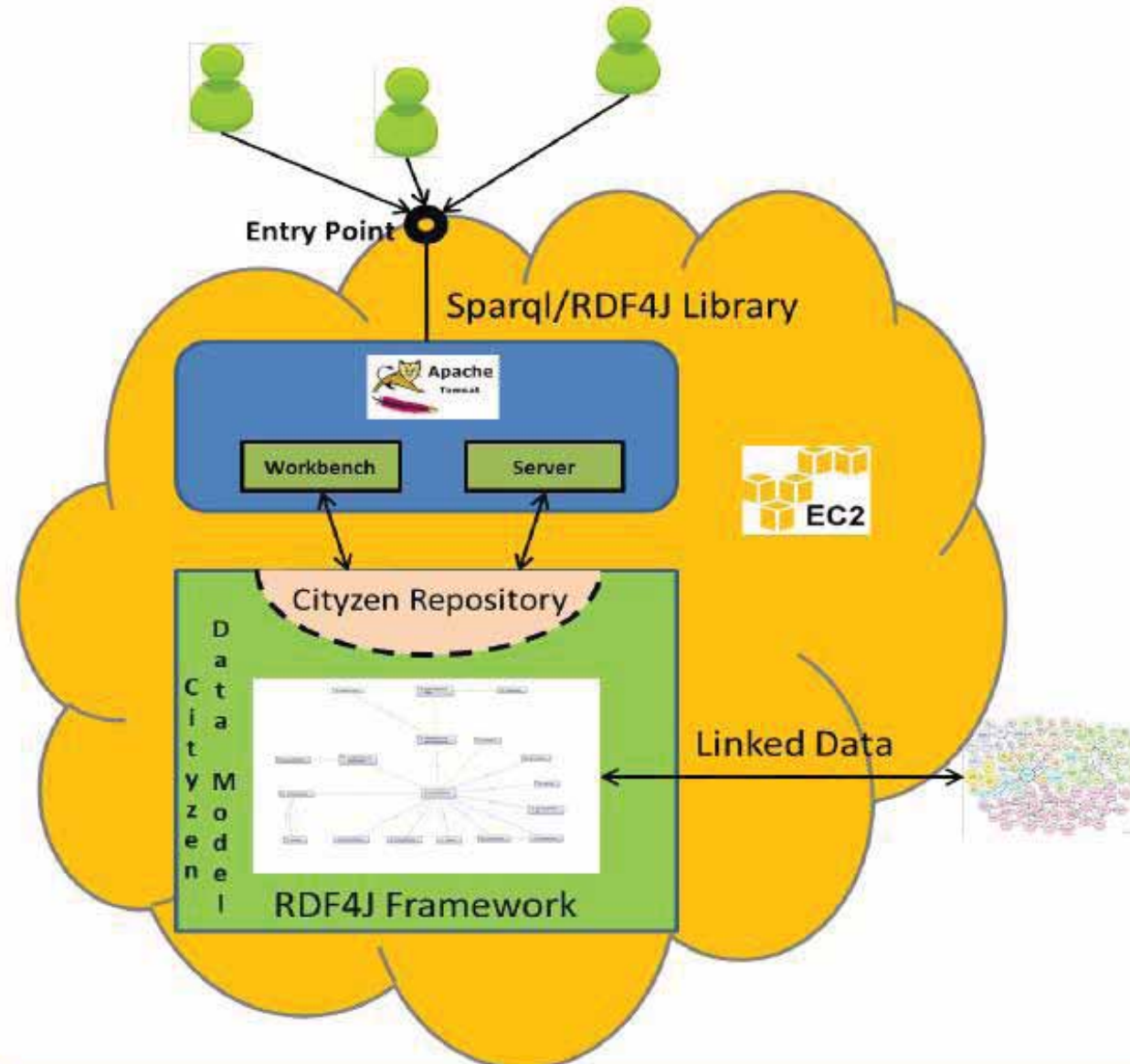
Spatial
Concept



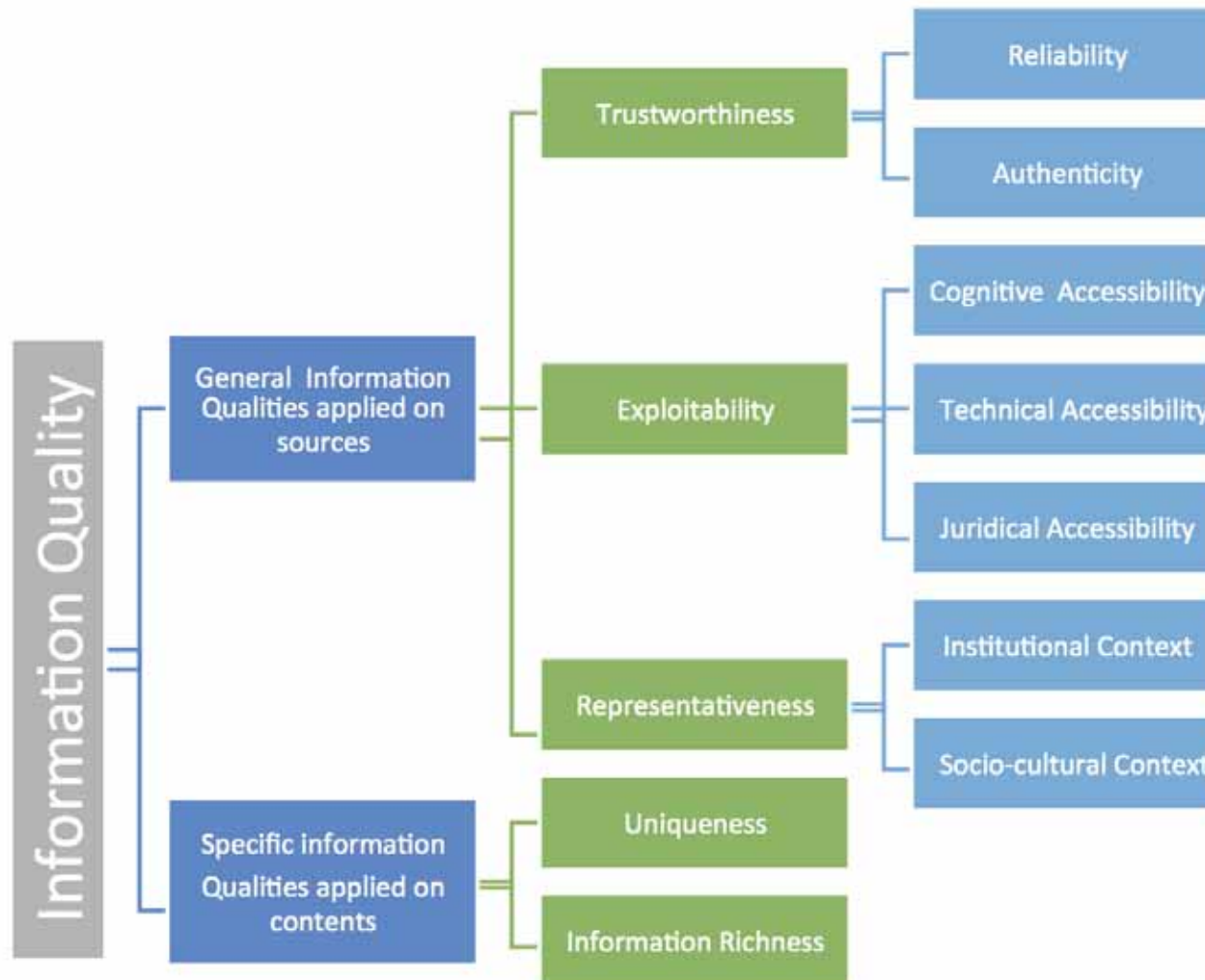
Dataset - Linking



Dateset – triple store design



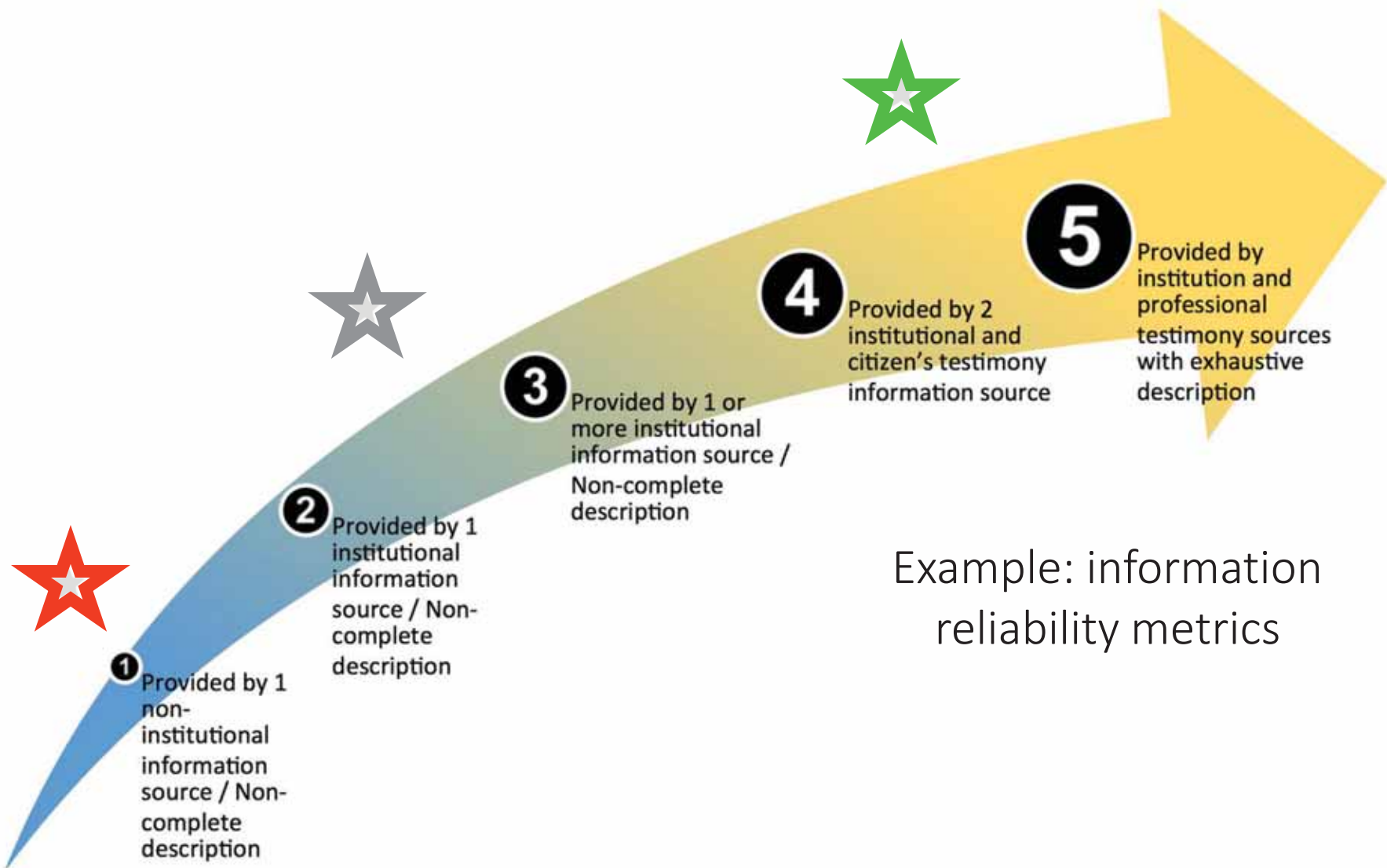
Dataset – Quality control



Score	Information Quality Levels				
	*	**	***	****	*****
Criteria					
General Information Qualities applied on sources					
Trustworthiness					
Reliability	Provided by 1 non-institutional information source Non-complete description	Provided by 1 institutional information source Non-complete description	Provided by 1 or more institutional information source Non-complete description	Provided by 2 institutional and citizen's testimony information source	Provided by institution and professional testimony sources with exhaustive description
Authenticity	Provided by source non-identified	Provided by Citizen's authentic testimony	Provided by Authentic professional testimony	Provided by multi-source information	Provided by authentic multi-source information
Exploitability					
Cognitive Accessibility	French	French/German/Italian	French/German/Italian/English/Spanish	Multi-language (7 languages)	Multi-language (more than languages)
Technical Accessibility	Easy to understand for the French tourist or citizen	Easy to understand for the French/German/Italian tourist or citizen	Easy to understand for the tourist or citizen	All tourists and citizens can understand easily	All tourist citizen understand easily
Juridical Accessibility	Information is accessible but non-consultable physically	Information is accessible but non-consultable physically	Information is accessible but non-consultable physically	Information is Open access and can be consulted with a request to the institution with an identifier (download if check-in as user)	Information Open access can be consulted with a request to the institution with an identifier (download if check-in as user)
Representativeness					
Institutional Context	Recommended by Office of Tourism but information not up to date (of last year)	Recommended by Office of Tourism but information not up to date (of last year)	Recommended by Office of Tourism but information not up to date (of last year)	Recommended by Office of Tourism and information is up to date	Recommended by Office of Tourism and information is up to date
Socio-cultural Context	Information is not scarce, in a context summary (abstract)	Information is not scarce, in a context summary (abstract)	Information is not scarce, in a context summary (abstract)	Information scarce by the exhaustiveness of the context	Information scarce by the exhaustiveness of the context

Score	Specific information Qualities applied on contents				
	*	**	***	****	*****
Criteria					
Specific information Qualities applied on contents					
Uniqueness	-Density of information, lacks of precision, - Abundance of format, media, information typology, subject, period / context	-Density of information, lacks of precision, - Abundance of format, media, information typology, subject, period / context	-Selection of information, lacks of precision, -A selection of format, media, information typology, subject, period / context	-Exclusive information, precise and indexed - Scarcity of format, media information typology subject, period/context	-Exclusive information, precise and indexed - Scarcity of format, media information typology subject, period/context
Information Richness	The Information will be searched by personal words, there are no feedback available, the information will be limited audio, with an Impersonal source (ex. Tourism Office)	The information will be searched by a list of keywords, the feedback will be slow but available, the information will be limited in Visual, with an Impersonal source (ex. Tourism Office)	The information will be searched by a list of Keywords, the Feedback will be slow, the information will be diffused in Visual or in Audio, with an Impersonal source	The information will be searched by a Numerical Index, the Feedback will be Fast, the information will be diffused in Visual or in Audio, with a Personal and an Impersonal source (ex. Citizen's testimony and Tourism Office)	The information will be searched by a Numerical Index, the Feedback will be Immediate, the information will be diffused with a developed Visual and Audio, with a Personal and an Impersonal source (ex. Citizen's testimony and Tourism Office)

Dataset – Quality control





1965



Assemblée générale



Assemblée générale

10:41

1979

brisolée des boulan

Le 4 novembre 1979, les membres du Vieux-Québec ont célébré la traditionnelle brisolée annuelle. Pour ce jour, le protocole de l'assemblée est simple : les dames se réunissent dans une salle de réception pour partager un moment de détente et de convivialité. Il s'agit d'un moment très apprécié par tous les participants. On se retrouve tous ensemble pour partager un moment de détente et de convivialité. (Note ajoutée)

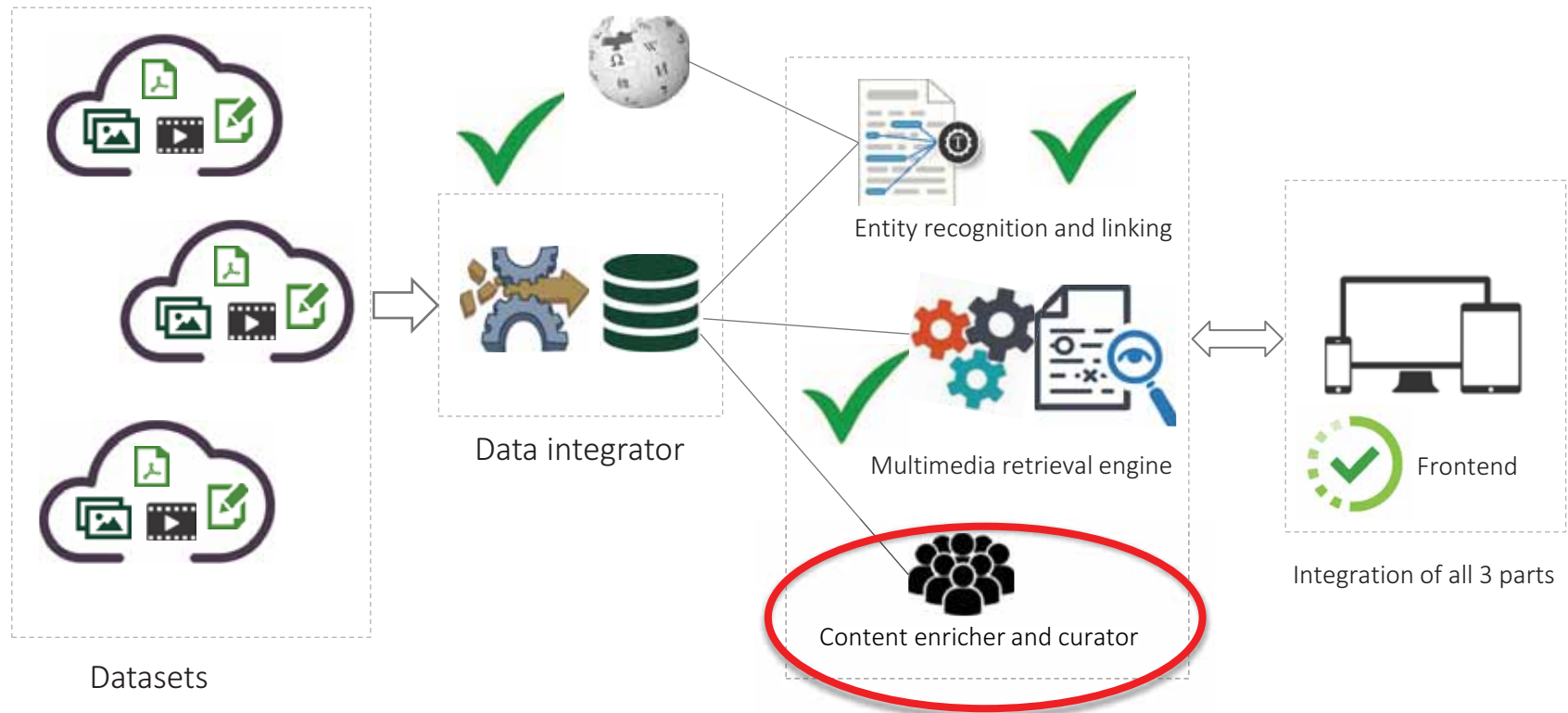


Un moment de récréation dans une vie très active : la brisolée annuelle !

Section féminine

10:41

Conclusion



Evaluation on the full integration

Future developpments...

(1) Crowdsourcing

- Issues with the existing data
 - Part of the data is *not complete*
 - *Missing* information: GPS, subject, location
 - *Conflicting* data from different sources
 - e.g., two different dates for the same event



Who?



When?



Where?

- Initiative for collecting new digital data from individuals
 - Many individuals have data about the region they live



(2) Project extension

- Enable a more performing quality assessment focusing on the metadata use
- Enabling a multilingual functionality: over the structured and non structured extracted data
- **Historical and Cultural Heritage exploration based on geolocalisation**
(Radar mode service)
- Relevance assessment of City-Zen when applied in pilot cities

QUESTIONS ?

Prof. Basma Makhoulf Shabou

basma.makhoulf-shabou@hesge.ch

Prof. Maria Sokhn

maria.sokhn@hevs.ch

Thank you!