



LiWA - Living Web Archives

Project Presentation
at the Information Day
December 17, 2007
Luxembourg

Claudia Niederée
L3S Research Center

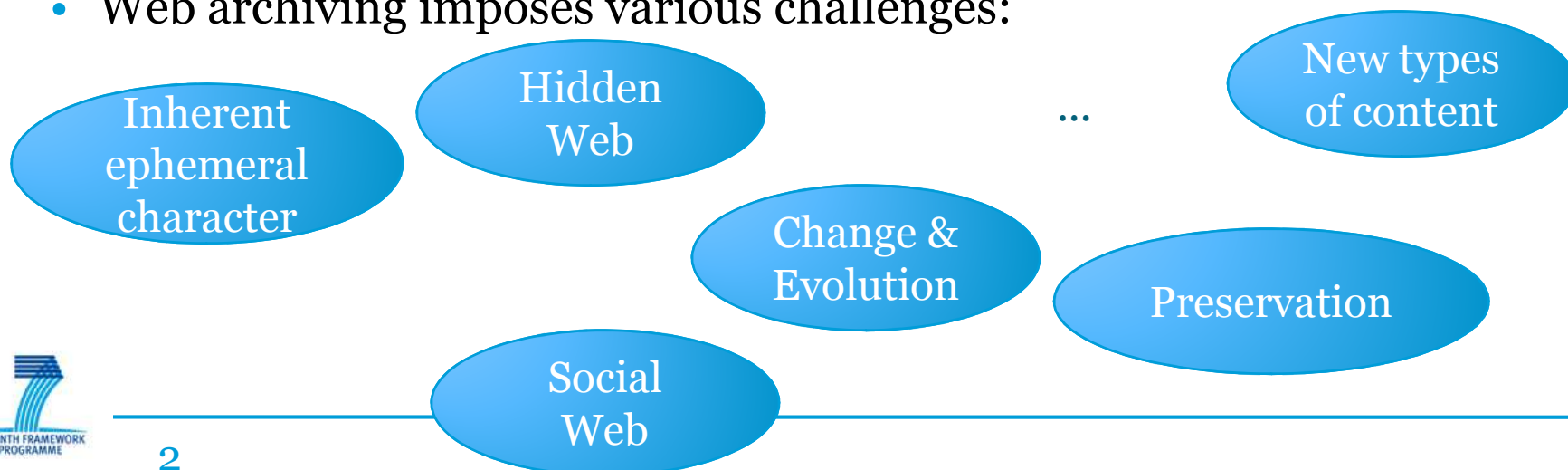




Motivation

- Role of Web:
 - providing information and services for seemingly all domains
 - reflecting all types of events, opinions, and developments within society, science, politics, environment, business, etc.
 - giving room for the articulation for a multitude of stakeholders
- Archiving this quickly changing multifaceted information space has become a relevant issue for cultural heritage

- Web archiving imposes various challenges:

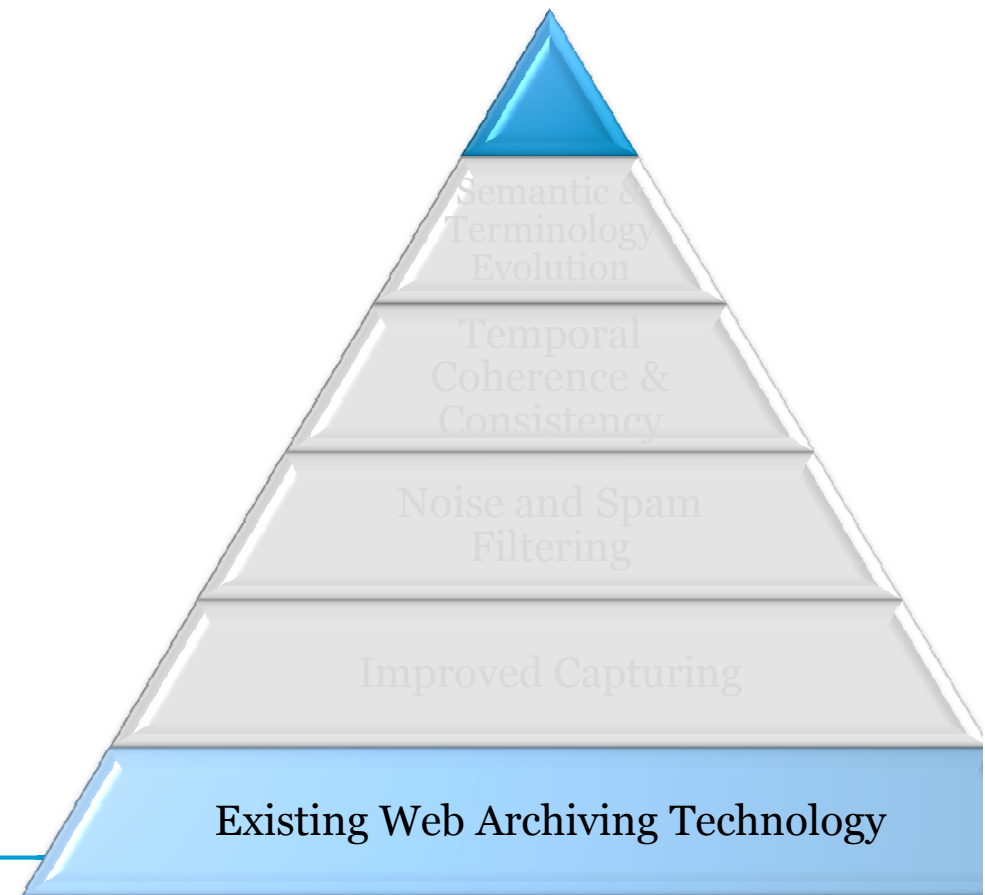
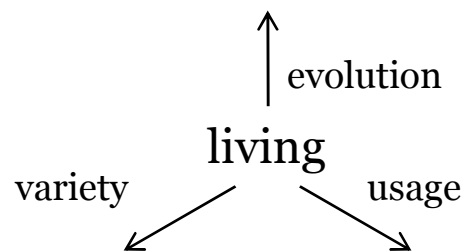




LiWA Goal

- Next generation Web Archiving technology for:
 - high Quality Web Archives
 - long-term Archive usability

→ From Web page storage to “Living Web Archives”



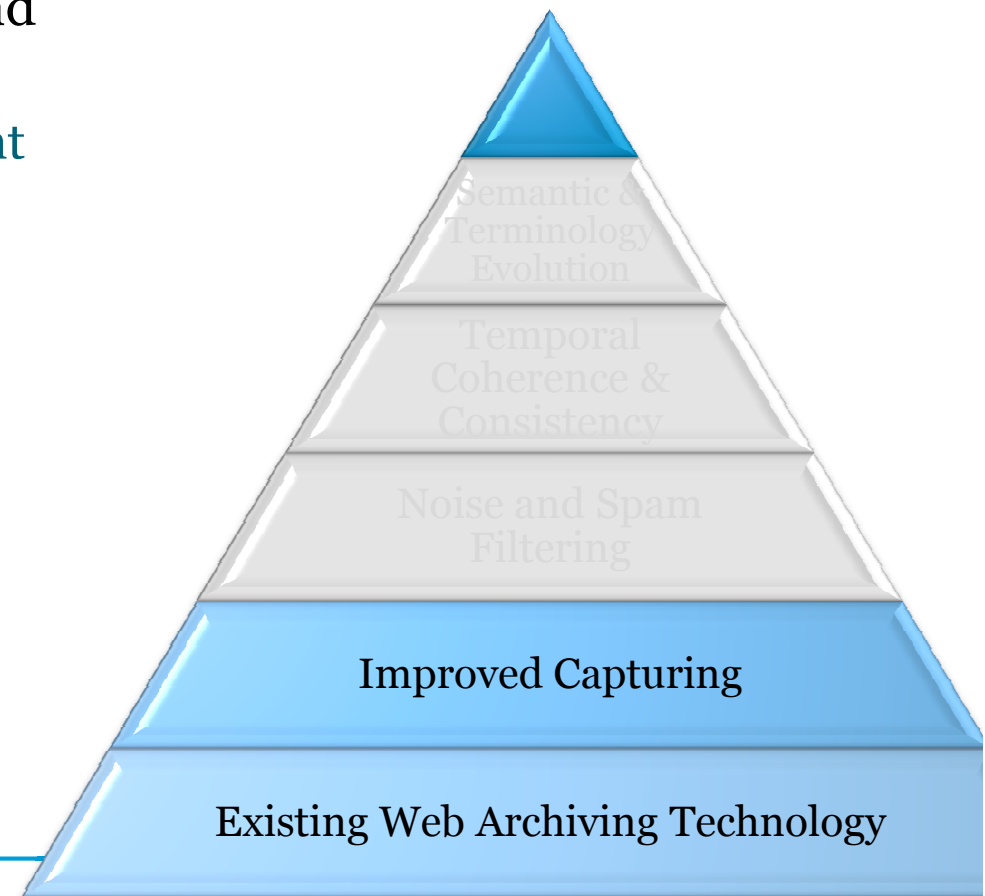


LiWA Objectives: Archive Fidelity

Next generation Web Archiving methods and tools:

- enhancing Archive Fidelity and authenticity by
 - capturing all types of content
 - capturing of Hidden Web
 - detecting traps

“Living Web Archives“



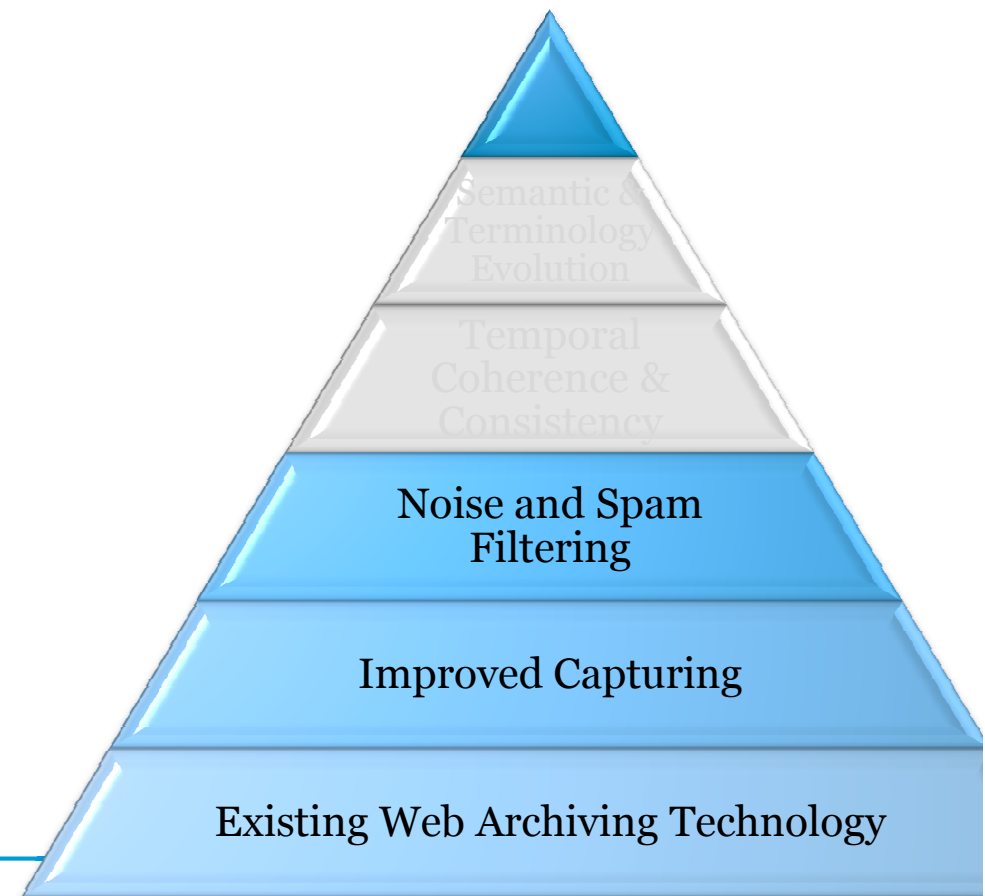


LiWA Objectives: Archive Fidelity

Next generation Web Archiving methods and tools:

- enhance **Archive Fidelity** and authenticity
 - capture all types of content
 - detect traps
 - filtering Web spam
 - filtering noise

“Living Web Archives“



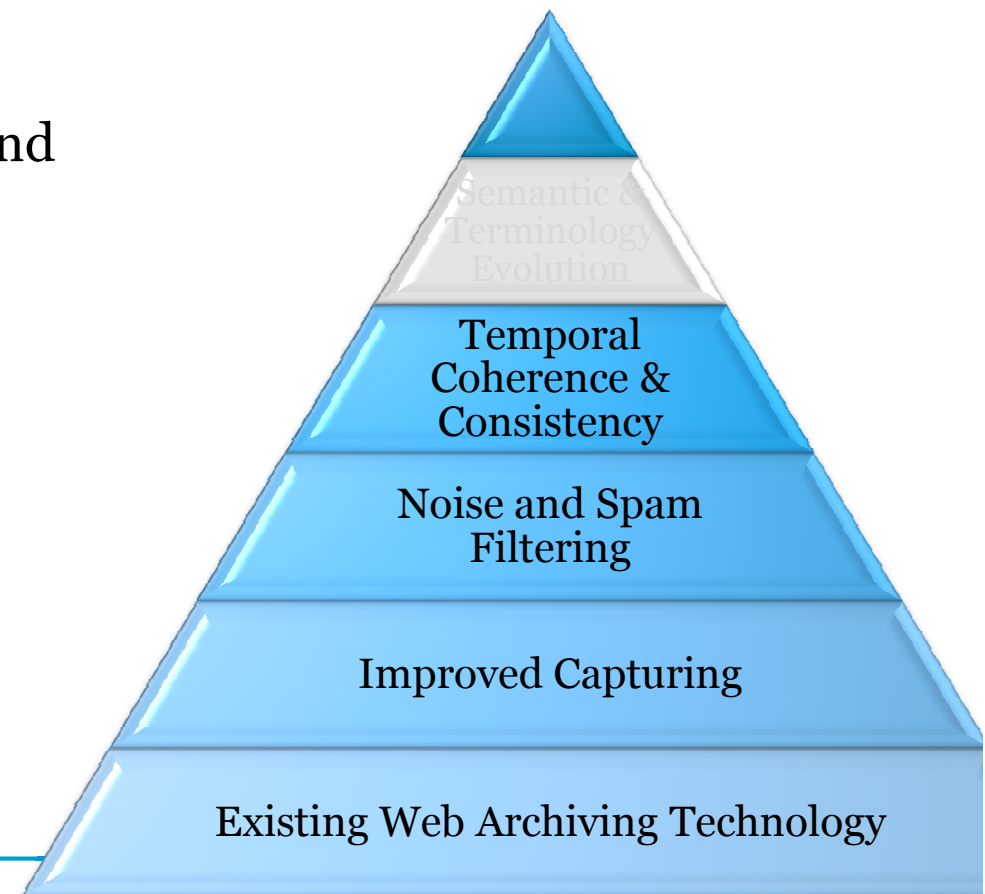


LiWA Objectives: Archive Coherence

Next generation Web Archiving methods and tools:

- enhance Archive Fidelity and authenticity
- improve **Archive Coherence and Integrity**
 - deal with issues of temporal Web construction
 - identify, analyse and repair temporal gaps
 - consistent Web archive federation

“Living Web Archives“



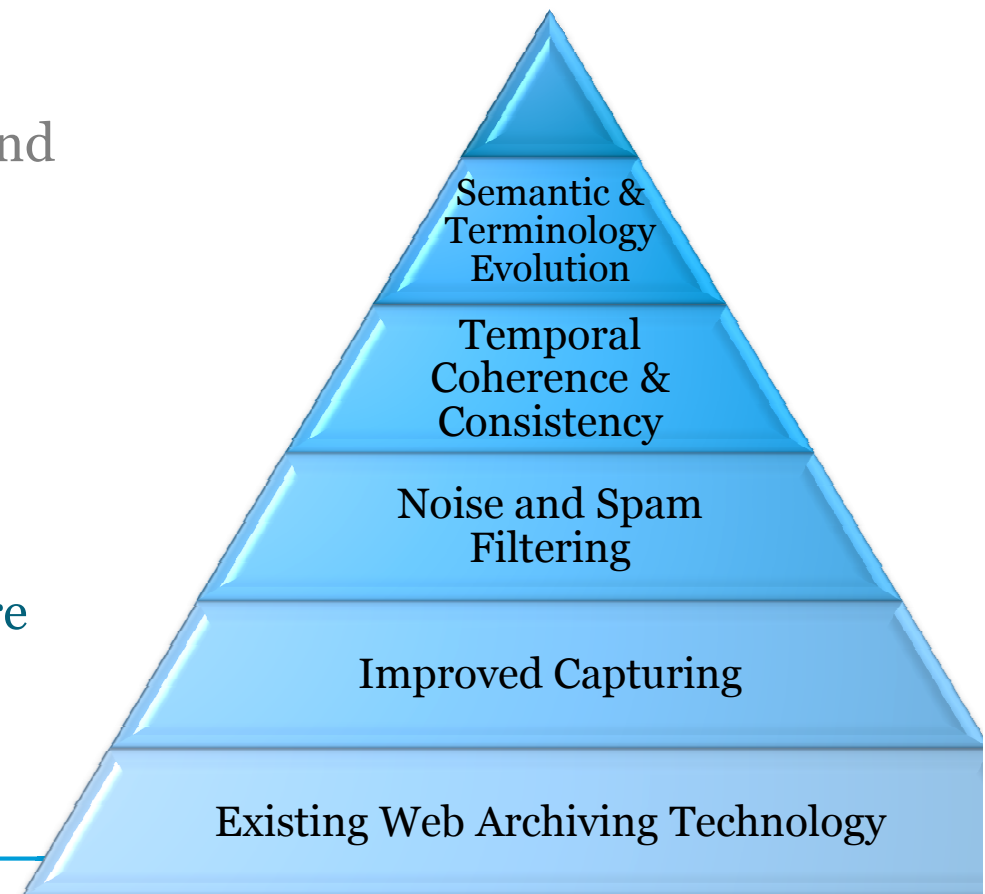


LiWA Objectives: Archive Interpretability

Next generation Web Archiving methods and tools:

- enhance Archive Fidelity and authenticity
- improve Archive Coherence and Integrity
- facilitate (long-term) Archive Interpretability
 - dealing with terminology evolution
 - handling semantic evolution
 - preparing for evolution aware access support

“Living Web Archives“





Outcome & Impact

- Outcome
 - set of innovative services for capturing, preserving, and enriching Web content
 - validation of LiWA technology by two exemplary applications
 - integrated reference open source solution for web archiving aligned with work of International Internet Preservation Consortium (IIPC)
- Impact
 - technology and best-practices for all stakeholders involved in Web archiving
 - contribution to Europe wide coordination of Web archiving technology
 - high-quality LiWA based Web archives as an European Asset



Thanks!

LiWA Team:

L3S Research Center

European Archive

Max-Planck-Institut für Informatik

Hungarian Academy of Sciences

Stichting Nederlands Instituut voor Beeld en Geluid

Hanzo Archives Limited

National Library of the Czech Republic

Moravian Library



see also: www.liwa-project.eu