Automated XML Schema Representations for Sensor-based Information Processing Systems

By: Luz Acaba, MS Student

Advisor:

Prof. Domingo Rodriguez

WALSAIP

Automated and Information Processing Laboratory

University of Puerto Rico at Mayaguez (UPRM)

March 14, 2007

Problem Formulation

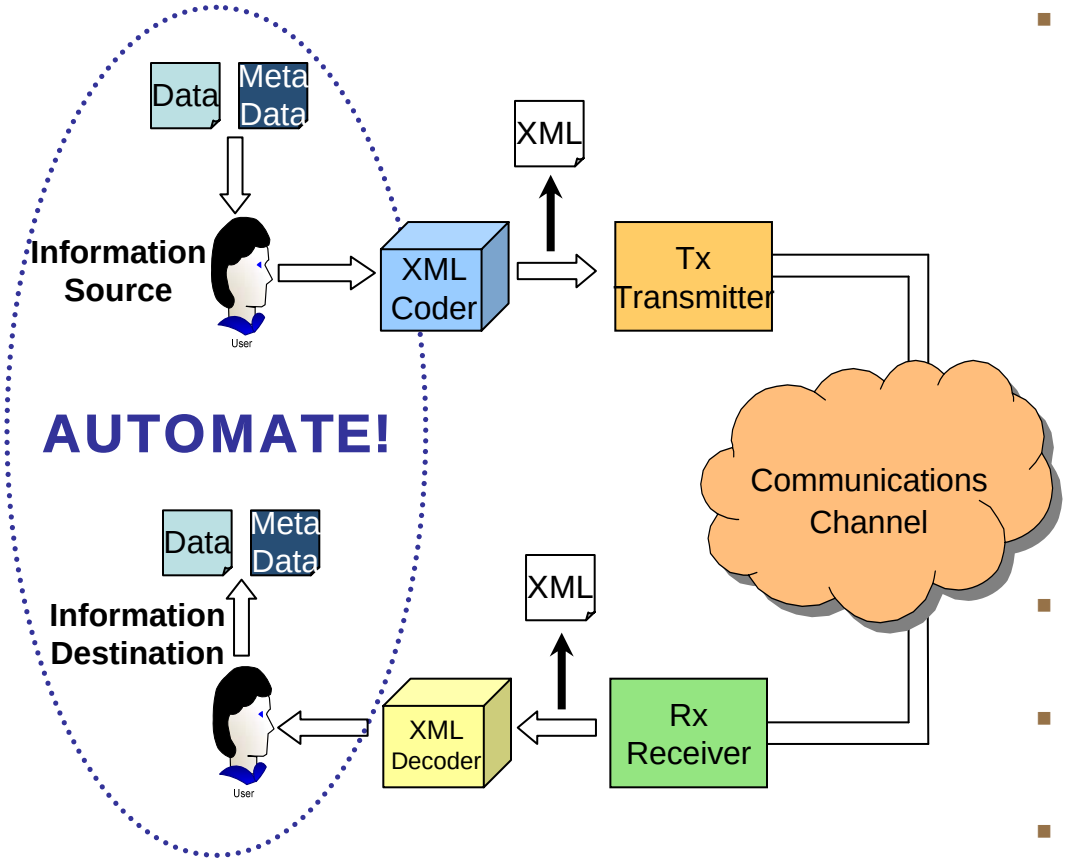
Raw Data + Metadata + Processing

Decision Making!

- **Raw Data** – Defined as all readings collected directly from sensors
- **Metadata** – data that describes raw data. Metadata is crucial providing researchers concrete data about the real conditions in which data was collected. Metadata is a determinant of how the environment influence the measurement case of abnormal findings.

- **There is a need for proper characterization of binding/coupling relationships between data and metadata files to improve information content analysis.**
- **Data should be interoperable across heterogeneous users with different data architectures, storage systems, and platforms.**
- **Lack of support for dynamic metadata management**
- **Systems need to incorporate information from "human sensors".**

Proposed Solution

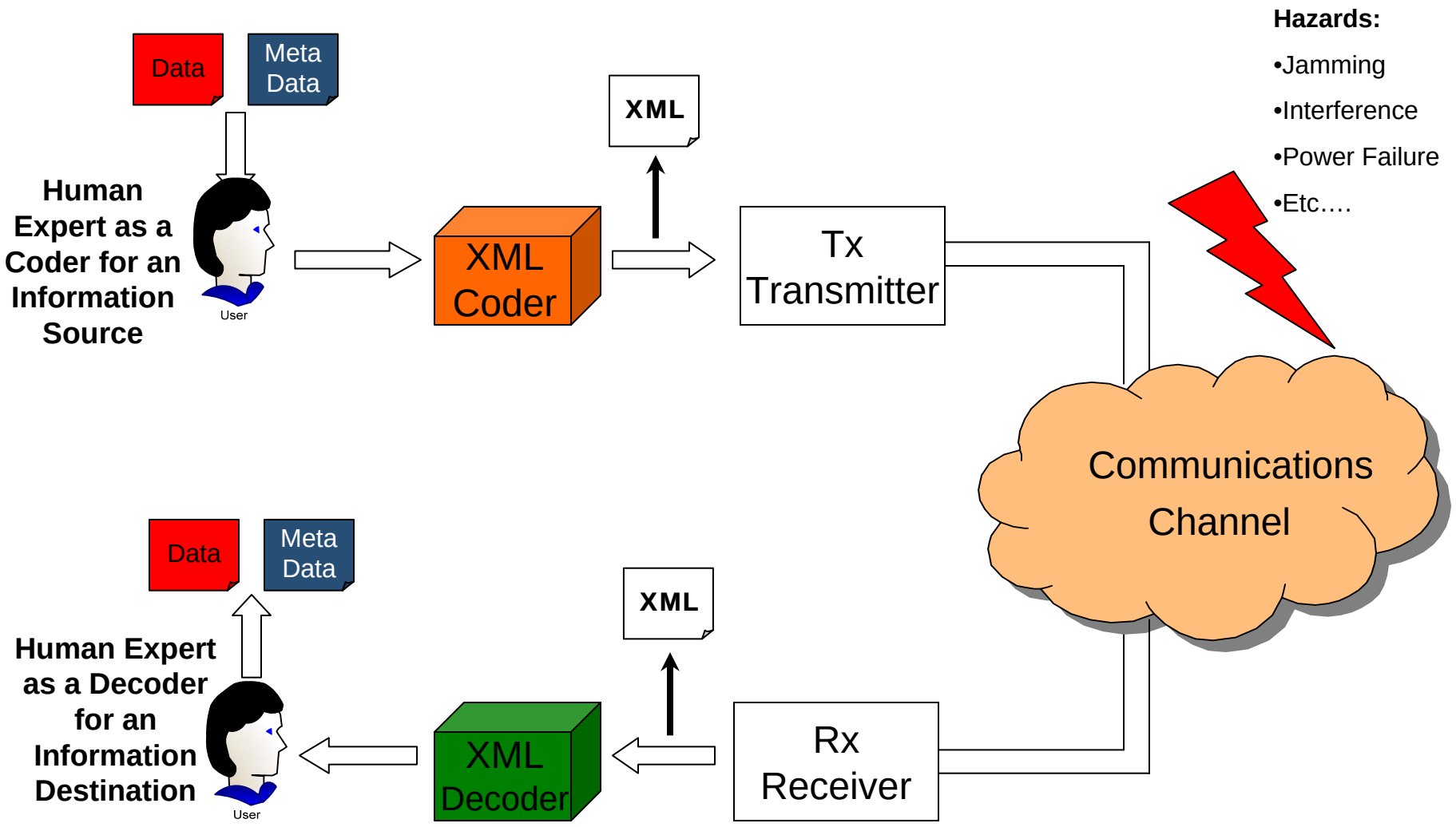


Shannon's Theory and XML Processing

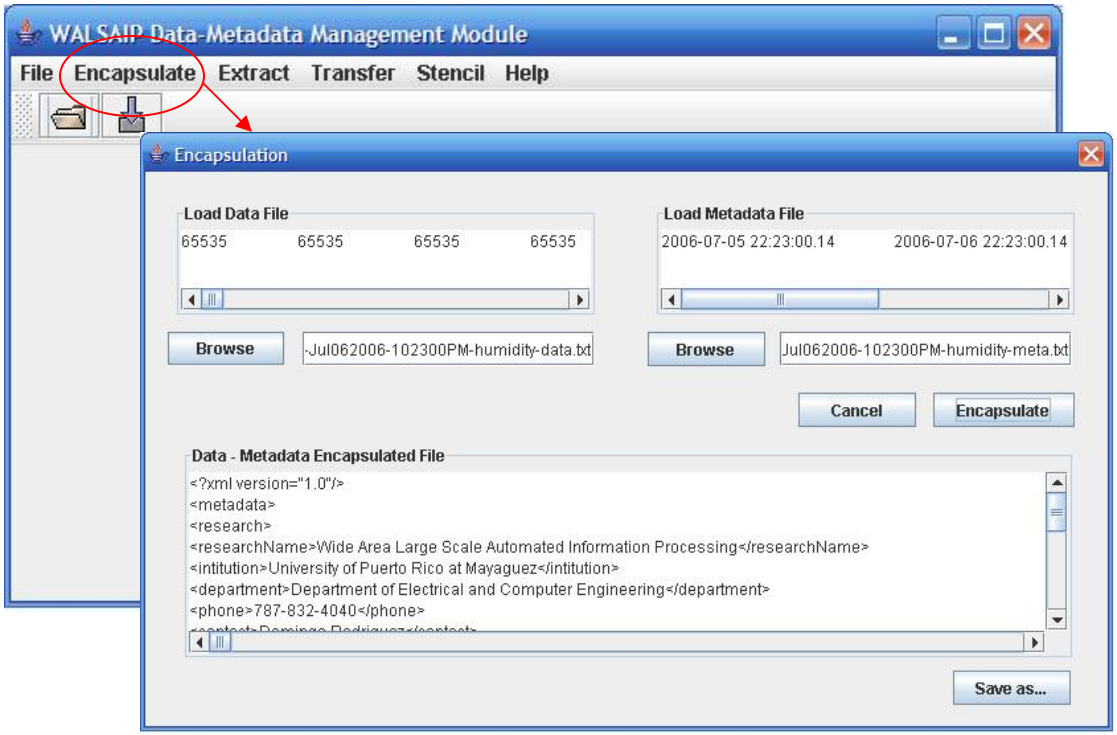
- Information theoretic measures are used to study how the extensible markup language (XML) may serve as a means for integrating symbols and meaning (semiotics and semantics parts), from metadata, with signals and structure (syntactic part) from sensor based raw signal-data.

  User may develop "stencils" in order to customize "XML tags" during encapsulation.
- Proposed solution contemplates *dynamic metadata management (DMM)*.
- Data and metadata can be enhanced with user observations.
- User can edit the metadata by annotating additional comments and parameters.

XML Source/Channel Coding



XIR: XML Information Representation



- XIR – XML Information Representation tool is a java based application for the coupling/binding representation of data and metadata entities associated with physical sensors pertaining to environmental surveillance monitoring (ESM) applications.

- Tools for Application Implementation
  - Java
  - FTP – File Transfer Protocol will be used to transfer encapsulated files.
  - XML – eXtensible Markup Language is a general purpose markup language capable of describing many different sets of data. Provides a text-based means to describe and apply a tree-based structure to information.