

<NaML />

A Markup Language for Narrative Knowledge Annotations, Mining and Management

Jens Fendler
j.fendler@polytechnic.edu.na

Department of Software Engineering
Polytechnic of Namibia

Narrative Knowledge, as the name implies, is not naturally bound to any media other than speech. In order to preserve such knowledge and to make it accessible through computer systems, a formal and standardised basis needs to be identified, and to be integrated with other information sources.

While Audio and Video recordings of stories (Bidwell et al., 2011) are tempting due to the simplicity of their creation, such attempts have not yet resulted in significant progress regarding formalisation, querying, or further IT-based processing. Though tools such as Anvil (Kipp, 2011) can be used to manually add meta data and formal aspects to the recordings at a later stage, such recordings remain a hardly exploitable - and thus limited - source of story information for IT systems.

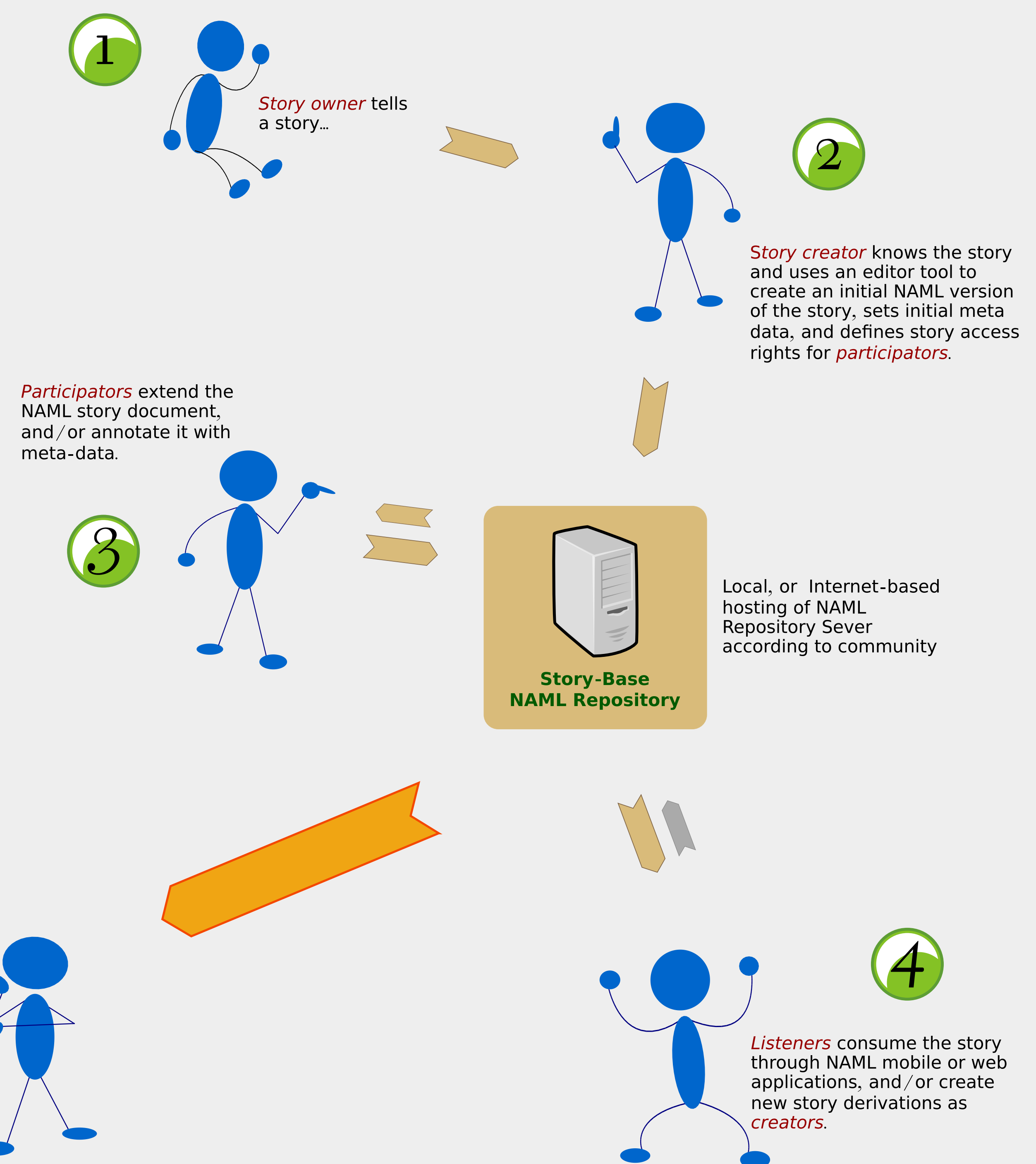
The author argues that narrative knowledge should be primarily stored in a format as closely related to the verbal content as possible, while providing sufficient means of formalisation and automated processing capabilities. For this purpose, the Narrative Annotation Markup Language (NaML), a custom XML-based language (XML, 2006) for story representation and annotation is being proposed.

Initially capturing narrative stories in their written form in a standardised format, NaML documents can later be extended with meta data and reference links to related stories or additional external information to form a comprehensive and fully navigatable story repository. Readily available XML technologies such as XPath (Clark and DeRose, 1999) or XSLT (Clark, 1999) can be used to query and export selected parts for the interested individual (listener), as well as allowing flexible story-mining towards a deeper understanding of narrative information for research communities. Access control mechanisms are built-in from the beginning in order to give story owners full control over read- and write access to their stories.

NaML does not intend to replace existing knowledge management systems (e.g. systems based on audio and video recordings), but aims to provide an extensible foundation to consolidate the various narrative knowledge elements and information sources in a standardised and accessible way.

Background

Process



NaML Example

```
<?xml version="1.0" encoding="utf8"?>
<!-- Example NaML Document -->
<story xmlns="http://www.polytechnic.edu.na/naml/0.1" xml:lang="en" id="s://na/38">
  <!-- Story Metadata (references, history, restrictions, etc.) -->
  <title>The Beginnings of NaML</title>
  <!-- This story is based on another story (see URI) -->
  <basedOn id="s://na/17" />
  <!-- History of when the story was told to whom -->
  <history>
    <told on="2011-05-25T10:14:39Z"
          by="p://na/1124"
          to="g://na/19"
          occasion="o://na/43992" />
  </history>
  <!-- access restrictions -->
  <access default="allow">
    <include id="g://na/19" />
    <exclude id="p://na/1339" />
  </access>
  <!-- story content -->
  <content>
    <when id="t0" type="fuzzy" scope="months">Early in 2003</when>,
    when <who type="group">some IT people</who> were sitting together
    over a strong cup of <concept ref="c://coffee">coffee</concept> in
    an office at the <place ref="p://na/Polytechnic">Polytechnic</place>,
    they were thinking of how <concept ref="c://knowledge/traditional">
    traditional (and other) knowledge</concept> available so far only
    in stories of the <who ref="g://na/1">Namibian people</who> could
    be <act ref="management">persisted, managed and hopefully further
    utilised</act> to the benefit of generations yet to come.
    <when relativeTo="t0" delta="after" scope="years">Some years later</when>
    after plenty of long discussions with <who ref="p://na/1124">Heike
    </who>, <who ref="p://na/389">Gereon</who>, <who ref="p://na/975">
    Shilumbe</who>, and <who type="fuzzy" scope="many">many others</who>,
    <who ref="p://na/1088">Jens</who> had the <concept ref="c://idea">
    idea</concept> of developing a custom <concept ref="c://language">
    markup language</concept> representing <concept ref="c://story">
    stories</concept>. Based on <concept ref="c://XML">XML</concept>,
    it would be possible to use readily available tools and technologies
    for the management part, and to easily integrate <concept ref="c://story">
    story</concept> content in a variety of <concept ref="c://computer">
    computer-based systems</concept>.
  </content>
</story>
```

2011
2012
2013
2014

NaML is still in its early stages and not yet a working tool. We are currently busy with the identification of last features and extension points, followed by the development of a full XML Schema specification.

The next step will then be the development of a reference implementation of an NaML editor, allowing the research community and selected user groups to capture and annotate an initial set of stories for a testing repository. Existing narrative knowledge will be analysed and possibly imported into the repository.

Once the repository has been sufficiently populated to represent not only the most trivial cases, a web front-end will be developed for story listeners to receive and re-publish stories, and for researchers to run repository queries. It is also envisaged to complement the web application with a mobile version for further field experiments with the indigenous community.

Once the system has proven to meet its primary objectives, we are looking at the integration of additional - probably web-based - knowledge sources, to increase the overall story relevance and to provide additional context. Finally we are going to integrate existing narrative knowledge.

Future Work

Pros & Cons

Advantages

- Based on XML Standard (Tool availability)
- Flexible and easily extensible
- Story Versioning Support (XML Diff)
- Simple Integration and Transformation (XSL)
- Common format for capturing and processing
- Allows evolutionary story mark-up/annotations

Disadvantages

- Not yet implemented, Proof-of-Concept pending
- No native tool support (to be developed)
- Possible loss of story aspects unless explicitly marked-up or attached
- Performance impacts for large repositories

finally {}

Acknowledgements

Thanks to all colleagues and students in the School of Information Technology at the Polytechnic of Namibia. Especially to Heike Winschiers-Theophilus, Gereon Koch-Kapure and Shilumbe Chivuno-Kuria for their contributions to this line of research and a lot of field work. Special thanks also go to Matthias Rehm for his encouragement and lots of work yet to come.

References

- Bidwell, N.J., et al., *Pushing personhood into place: Situating media in rural knowledge in Africa*. International Journal of Human Computer Studies (2011), doi: 10.1016/j.hcs.2011.02.002
- Kipp, M., *Anvil. The Video Annotation Research Tool*. Online at <http://www.anvil-software.de/>. Last accessed 26 May 2011.
- Tim Bray et al. (Eds), *Extensible Markup Language (XML) 1.1 (Second Edition)*. w3c. Online at <http://www.w3.org/TR/2006/REC-xml11-20060816/>. Last accessed 26 May 2011.
- James Clark and Steve DeRose (Eds.), *XML Path Language (XPath). Version 1.0*. W3C Recommendation 16 November 1999. Online at <http://www.w3.org/TR/xpath/>. Last Accessed 26 May.
- James Clark (Ed), *XSL Transformations (XSLT) Version 1.0*. W3C Recommendation 16 November 1999. Online at <http://http://www.w3.org/TR/xslt>. Last accessed 26 May 2011.

For more information please contact me:

j.fendler@polytechnic.edu.na, jf@teamskill1.de, jens@iktc2011.org,

126 Department of Software Engineering, Polytechnic of Namibia, Private Bag 13388, 13 Storch Street, Windhoek, Namibia

