

Toward the Interoperability of Language Resources

July 13-15, in conjunction with the 2007 LSA Summer Institute



A Development Perspective on Interoperable Tools

Shane Stephens

Commonwealth Scientific and Industrial Research Organisation

I am currently a developer on the Annodex Platform, a cross-platform approach towards multimedia publication, annotation and sharing. In this position statement, I discuss the approaches which the Annodex team have taken towards ensuring the Annodex Platform's interoperability with other media, annotation and web-based tools and platforms.

The Annodex Platform allows: alignment of textual annotations with time points in multimedia documents; streaming access to arbitrary time points within multimedia documents; web-based modification of annotations; and integration of annotated multimedia playback into web documents.

This makes the platform suitable for inclusion within a wide range of applications, including research websites with shared annotation of video and audio; reference websites with permanent URIs for important features within annotated multimedia; and language teaching resources that include annotated multimedia.

The following key points explain the approaches that the team has taken towards ensuring interoperability:

- **Open standards and open formats:** The Annodex Platform uses and enhances Ogg, an open-standard multimedia format. This ensures that a range of tools and players are available for Annodex media. The team has also released a number of open standards to ensure that Annodex formats are understandable by other teams wishing to interoperate.
- **Format extension:** When a new standard is required, the Annodex team has ensured that this standard extends and interoperates with existing standards. For example, Annodex documents are valid Ogg documents, Annodex annotations are valid XML files, and Annodex time-point URIS are valid HTML URLs. This allows existing tools to be used with the platform without requiring an awareness of the new formats.
- **Modular approach:** The platform consists of a number of interoperable components. By considering how internal components should interchange data, we have helped ensure that external data formats are sane. Furthermore, developers are free to pick



and choose components according to their requirements, and develop additional components as needed.

- **Integration into Firefox:** With numerous plugins and extensions, Firefox is already a platform for interoperability. Integrating into the DOM and providing access through Javascript ensures that a web-based UI incorporating annotated multimedia can quickly and effectively be generated.
- **Small libraries; many bindings:** most of the capabilities of the platform are factored out into a number of small libraries, with bindings for several languages, including Python, Perl, and PHP. If developers wish to include Annodex capabilities without using the prepared components, then these libraries can be imported instead. For example, libannodex is currently used in our python-based CmmWiki application to serve Annodex media, and liboggplay can be integrated into standalone applications to provide annotated multimedia capabilities.

