

# A Metadata Collection Technique for Documents in WinFS

*Stijn Dekeyser*

Department of Maths & Computing, University of Southern Queensland (USQ), Australia

**Aim:** To propose a relatively preliminary yet sufficiently general GUI-based technique to collect rich metadata for documents, to be used in a file system wrapper such as Microsoft's forthcoming WinFS.

**Methods:** With the recent beta release of Microsoft's WinFS file system wrapper, the traditionally disjunct research areas of databases, document computing, and file systems are merging. In such environments, it is relatively clear how to query the file system for metadata about documents. However, it is not yet clear how rich metadata can be collected for documents in an intuitive manner suitable for novice users. We present a GUI-based technique for capturing both *formative* and *contextual* metadata. The former relates to metadata that can be set automatically when a document is created (such as the shutter-speed setting for a digital camera image), while the latter represents the much richer metadata that can only be determined through user interaction (e.g. the persons appearing in the image). The technique, described in more detail in [2] entails three parts:

- **Class Hierarchy.** We propose to employ a hierarchical inheritance-based class diagram that complements WinFS's built-in technically-oriented classes. The hierarchy is built in three phases: first, we propose that the first few levels in the hierarchy contain built-in classes that represent tangible concepts (e.g. **Person**, **Product** etc.) likely to be of general use. Secondly, we propose that businesses create classes in the next few levels of the hierarchy representing concepts in the company's specific branch of work (e.g. **Insurance**), and make those classes available in staff's computers. Finally, the lowest levels of the hierarchy should contain classes created by individual users (e.g. **Flooding**) to increase their productivity. For all newly created classes, their *properties* should reflect formative metadata while their *relationships* to other classes should reflect contextual metadata. We further propose straightforward, intuitive GUI actions to create new classes and their properties and relationships.

- **Capturing formative metadata.** A newly created class will be represented by an icon that acts as a folder for objects of that class. Creating a new document thus is simply adding an object to the folder; formative metadata is automatically provided by the class' constructor method (e.g. setting an image's metadata can be done through reading its EXIF information when copying it from the camera).

- **Capturing contextual metadata.** In WinFS all users can create relational views over documents and their metadata. We propose to create Class Folders which are updatable views [1] usually over just one table (class), although more complex ones are possible (see [2]). Such views are homogeneous (although taking inheritance into account) and in effect become virtual classes, in which objects and their formative metadata can be created as before. As our proposal ties contextual metadata to relationships between classes, our approach to collecting it upon saving is to supply the user with an Explorer-like window that contains several tabs. Because the document belongs to a certain class, we know which relationships to other classes are associated to it. Thus, we create one tab in the save-as dialog window per relationship. Inside those tabs, all member documents of the target class are displayed, together with all Folder views defined on that target class. The user is then asked to, *per tab* (i.e., per relationship of which the document's class is a source), select the documents and/or Folders to which the document is associated through that specific relationship. This selection procedure is akin to the current "Save As" dialog (although may take longer for the user to perform depending on the number of relationships) and serves to visually capture the contextual metadata rather than requiring its input through keyboard entry.

**Results:** This paper presents the concepts involved in our metadata collection technique rather than a working prototype. Hence we do not yet have but are planning to obtain results measuring usability (e.g. through ISO 9241 metrics).

**Conclusions:** We have proposed a generic technique to capture rich metadata for documents in the context of WinFS. The technique is based on sound relational theory and extends current "File Save As" GUI dialogs to assist users in capturing complex metadata for user-defined types in a reasonable, intuitive manner. We plan to implement a proof-of-concept system to test usability and effectiveness.

## References

- [1] U. Dayal and P. Bernstein. The updatability of relational views. In *VLDB'78*, pages 368–377, 1978.
- [2] S. Dekeyser. Metadata collection for documents in WinFS. Technical Report SC-MC-0524, USQ.

Proceedings of the 10th Australasian Document Computing Symposium, Sydney, Australia, December 12, 2005. Copyright for this abstract remains with the authors.