Archiving Information Workflows

Marie Vans [1] **and Steven Simske** [2]**;** [1]**HP, Inc; Fort Collins, CO/U.S.A.** [1]**;** [2]**Walter Scott, Jr. College of Engineering Systems and Mechanical Engineering; Colorado State University, Fort Collins, CO/USA** [2]

Abstract

Understanding the context that accompanies content is an important aspect of archiving documents. In a twist on that idea, we present a new approach to archiving information workflows that includes context. Using language translation as an example of context-shifting relevancy for different tasks in an object workflow, we show how device options can determine the language presented to the user. This context-shifting behavior is important information that should be archived along with the content in a workflow.

Introduction

Archiving information associated with workflows needs not only the content but also the context in which that information is used[1]. We need context for indexing, classification, clustering, and other workflow tasks. Context may include such things as language of origin and the language of the people querying the object. For example, it is known that "the use of multiple languages could result in some systematic or potential bias because the use of multiple languages requires more than just translation"; that "the use of multiple languages is further complicated by other contextual factors including the characteristics and roles of translators"; and that "the use of multiple languages is further complicated by other contextual factors including the characteristics and roles of translators."[2] The results of this study found that keeping cultural and conceptual equivalence was one approach to that was useful when creating electronic translators. In this work, we introduce a third level of archiving information which allows changes to "options on context" by adding the decisions made to change the context.

Context-Shifting Relevancy for Language-Related Tasks

Language is one example of context that may change during an information workflow. Currently, translation services are clunky and require several steps when an object's associated information is not presented in the needed language. However, just because the information is originally presented in one language does not mean that it can't be presented and understood in a different language. Context-shifting relevancy for tasks dependent on language understanding is one approach for addressing this. This means that regardless of the original language associated with an object, whomever is querying the object will access the information in the language they understand. For example, a label associated with an object in a workflow may have several information-carrying components including barcodes, websites for related information, and even appropriate standards related to the product. Figure 1 is a high-level description of this idea: An object can be examined as a whole and information on the

object itself can be queried. A context shift occurs once a single object item is selected for additional information processing. In this paper, we propose context-shifting relevancy for tasks dependent on language understanding
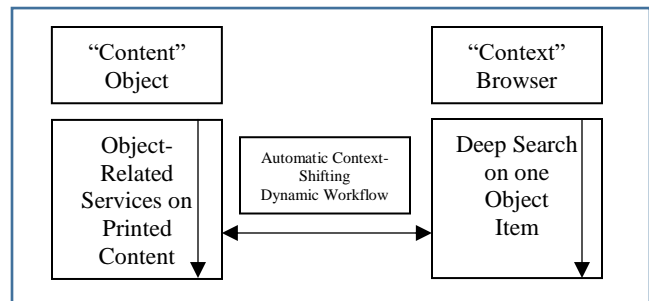

Figure 1: Two ways of delivering information based on the context: a) Object-Related services for all of the printed content and b) Browser-Related services for a single item on the object.

Equally important is the ability to capture this context shift for archiving purposes. For example, the language printed on package labels can be considered a "setting", among other settings, on the package. Normal package interactions such as inventory or distribution activities, using standard or public equipment, such as barcode readers, occur in the printed language. Context shifting occurs when personal devices, such as cellular phones or tablets, are used to query those same marks on the package; however, in this case, settings on the personal device are used to determine the language and the information presented to the user. Archiving the context information cans create benefits for litigation and auditing purposes as well as improving the user experience, among other advantages.

As figures 2-4 show, information may appear in one language, for example, English, but other language speakers may need to interact with the package.
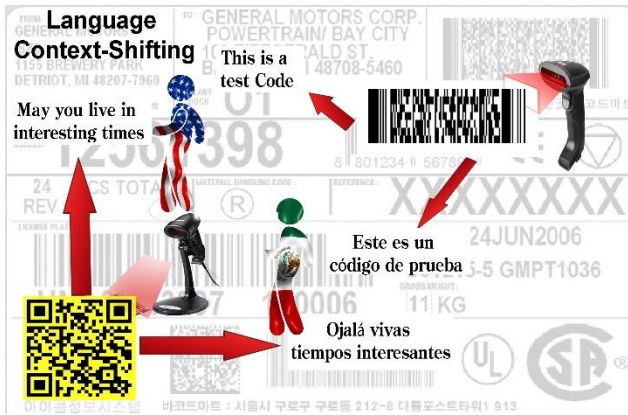
Figure 2. Example application of Language Context-Shifting: use of standard, off-the-shelf barcodes. Depending on the user's language setting, barcodes return the contents of the barcode in the language needed by the user, regardless of the actual data encoded.

Figure 3: Another example of the Language Context-Shifting: the mailing address changes depending on the location of the user. Address and URL are read by the context-shifting device (in this case a cell phone) but interpreted according to user settings. For the Spanish-speaker located in Mexico, the URL acquired is not only in Spanish, but is the localized website for users in Mexico.

Preferences on the mobile device (e.g. GPS is, preferred language is French) may result in a different, third behavior. Jacques from Quebec City gets the French website in both the locations (U.S., Mexico) depicted.

Figure 4 demonstrates the scenario in which content is not available to address the contextual needs. In this instance, we pull in real-time natural language processing, including translation.

Note also, that there is a logical link between these two fields of text data filtering and transduction. Namely, the accuracy

Figure 4: Example – Context Shifting on symbols. The left side shows an example of a symbol for which no information is available in a language other than English. For non-English speakers, a translation service is needed between the AR (Augmented Reality) device and the information associatedwith the symbol. The right side shows that the information associated with the symbol is directly accessible in the user's language.

of the language translation approach can be directly gauged by comparing the search results on the un-translated and subsequently translated corpora. If the translation is accurate, then the documents should respond very similarly to un-translated and translated queries against the corpora. This type of functional testing of un-translated and translated corpora also warrants further, quantitative investigation.

Conclusions & Future Work

This approach provides a seamless experience for object handling in object-centered workflows regardless of the language understood. In addition, a switch from specific content whilst interrogating the package to broader, more general context can be done once we pull a term from the package and use it for search/data-augment, etc. It also allows multiple language speakers to interrogate a single object and retrieve information in their native language.

References

[1] Vajcner, M. (2008). *The importance of context for digitized archival collections*. Retrieved from: https://quod.lib.umich.edu/j/jahc/3310410.0011.102/--importance-of-context-for-digitized-archival-collections?rgn=main;view=fulltext, accessed November 16, 2017.

[2] Im, E. O., Lee, S. J., Hu, Y., Cheng, C. Y., Iikura, A., Inohara, A., ... & Chee, W. (2017). *The use of multiple languages in a technology-based intervention study: A discussion paper. Applied Nursing Research.*

Author Biography

Marie Vans is currently a Senior Research Scientist with Hewlett-Packard Labs in Fort Collins, Colorado. She has a Ph.D. in computer Science from Colorado State University. She also received a Masters of Library and Information Science the Department of Information from San José State University in 2016, where she is focused on technologies for distance education.