Creative Engagement with Creative Works: a New Paradigm for Collaboration

Jones, Steven E.

sjones1@luc.edu
Loyola University Center for Textual Studies and Digital Humanities

Shillingsburg, Peter

Loyola University Center for Textual Studies and Digital Humanities

Thiruvathukal, George K.

Loyola University Center for Textual Studies and Digital Humanities

---

Funded in part by a Digital Start-up Grant from the National Endowment for the Humanities, **Creative Engagement with Creative Works** is a project to build a new online environment (e-Carrel) and integrated tools with the aim of improving understanding of creative processes across various humanities disciplines and genres. Studies of the production and reception of literary, historical, musical, and philosophical works are all built on primary materials that are textual in the broad sense-- documentary, material objects: manuscripts, newspapers, periodicals, pamphlets, books, and images. But current solutions to digitizing and providing access to these materials are structurally flawed and lead away from the often-stated goal of extensive collaboration. Digital transcriptions relying on XML or other inline markup can often prevent or limit collaboration on the files themselves, can (paradoxically) threaten a project's integrity, and can lead to early maintenance problems and the ultimate abandonment of projects beyond the lifetimes of their initial creators.

We propose using standoff markup instead, indexing these files to inviolate core files, thus affecting the way humanities disciplines interact with primary materials--moving away from proprietary, look-but-don't-touch window-case projects toward secure and enduring projects that are open to ongoing annotation and re-markup, and which thus encourage widely collaborative knowledge sites.

Electronic projects that restrict their construction and enhancement to a select few persons as a way to protect the intellectual integrity of their resources inadvertently introduce a significant threat to the durability of text files, whose hard-won accuracy is the result of expert attention. The creators' anxiety to protect the results of their time-consuming and labor-intensive in-line-tagging leads to a proprietary attitude that's detrimental to scholarly collaboration, because everyone knows that the text files are vulnerable to inadvertent change every time they are reopened for further tagging. Protective restriction also restricts truly collaborative work, the life of digital humanities projects in the world.

These restrictions can be loosened or eliminated by a fundamentally different approach to collaboration, durability, and maintenance, pioneered in scientific fields following principles of:

- modular component structure,

- connectivity,

- extensibility,

- distribution and aggregation systems,

- stand-off enhancement mechanisms, and

- methods for identifying and crediting researchers with their individual contributions to composite research projects.

These trends in software development allow for decentralized alterations. But even free/ open source (FOSS) software projects are not without problems. The past was dominated largely by projects where the code was kept under tight wraps, using version control systems such as CVS (Concurrent Versions System) and Subversion. Only a handful of FOSS projects succeed with this model, notably the Firefox project, but many FOSS projects are already migrating to more distributed approaches (the Python language, the Linux kernel, etc.). Much like archives, software development projects in the FOSS community tend to self-organize and establish their own governance. While they make their code available (as required by the FOSS licensing schemes) they also tend to fall into disuse, making it hard for new developers to come along and take the project to new levels.

Distributed version control systems record a project's entire history and state, which can be copied freely, allowing derived works to take place. When a copy is made, however, the entire history is kept intact, allowing new contributors to either make their own changes or to push their changes back to the original maintainer. The push/pull model is so sophisticated that anyone who makes changes can get recognition for their work, because their specific changes to the code are encoded in the derived history.

Our Creative Engagement with Creative Works and e-Carrel environment incorporates these principles and adds the functions of stand-off files for markup and annotation, along with a dynamic authentication mechanism. Experimentation has shown the considerable promise of such functionality. (See the Just InTime Markup system rototyped at the University of New South Wales at ADFA by our Senior Consultant, Paul Eggert and his team working on Australian literature; and see Desmond Schmidt's MVD or Multi Version Document system, the architecture and some code of which we are incorporating). These are the building blocks of our project's coherent vision for archiving creative works for creative collaboration, preservation, and dynamic interactive access, realized in the form of tools and programming frameworks.

We establish an image file and a base text for each significant version of a work. Text data for all versions are compressed in a composite, inviolable CorTex file, which anchors all stand-off contributions. Each participant's contribution is credited and protected from work by other contributors. Endusers choose a historical text (or other object) plus desired types of enhancements from a menu dynamically aggregated from distributed sources. The e-Carrel processes and presents perspectives of texts and enhancements for viewing, printing, or export in commonly used formats. Individual projects using the system can vet and certify parts of a project. And--most important--this system allows for the storage, retrieval, and coordination of different, even conflicting editorial or critical approaches to the same literary work. In this way, it opens the horizon of any given project's ongoing reception.

By giving scholars and students a vested interest in a growing integrated collaborative project, CECW ensures preservation and access to textual research projects and their superstructure of critical analysis at the same time that it promotes collaboration beyond the project initiators' participation and goals. Project viability follows community ownership and becomes a widely distributed responsibility. The system ensures long-term maintenance and growth through collective ownership, distributed storage, and the principle of LOCKSS (Lots of copies keep stuff safe). Our strong definition of LOCKSS, which distributes multiple accurate copies, and in an ongoing process verifies them by way of the persistent CorText data and checksum system, is significantly different from the "soft" version of LOCKSS that just sends "copies" into the world in whatever state and trusts to the hive mind for endurance and integrity.

CECW makes use of RDF capabilities and is XML aware but is by design markup agnostic. It allows for the importation of materials from other systems and prepares perspectives of the E-Carrel materials for export in various forms compatible with other systems (such as PDF and in-line coded XML).

This poster session will provide demos of portions of the software in development, in particular, the standoff markup tool used within the E-Carrel environment. A case-study CoreText will be loaded for the purposes of demonstration.

References

Eggert, Paul (2009). 'The Book, the e-Text, and the Worksite'. *Text Editing, Print and the Digital World*. Ashgate.

Loyola University Center for Textual Studies and Digital Humanities. `http://www.ctsdh.luc.edu/`.

Schmidt, Desmond (2009). 'A Data Structure for Representing Multi-Version Texts Online'. *Int. J. Human-Computer Studies*. **67**: 497-514.

Shillingsburg, Peter (2006). *From Gutenberg to Google*. Cambridge: Cambridge UP.