

A Conceptual Data Model and Schema for Curating Collections of Video Game Development Artifacts

Statement of Need: The Game Research Group (GAMER Group) at the University of Washington Information School (UW iSchool) – with the Video Game History Foundation (VGHF)¹ as the partner organization – requests \$249, 996 for a two-year research project to create a framework for describing, organizing, and preserving artifacts related to the development of digital games. Today more than ever, digital games play a widespread role in society. The Entertainment Software Association reports² that more than 150 million Americans play video games and that the industry employs more than 220,000 Americans. Over 350 US colleges and universities offer degrees or coursework to prepare game industry professionals, and games are used in education, science, and engineering as learning and literacy tools. The American Library Association (ALA) supports and promotes gaming in libraries and a growing number of university libraries circulate video games for the purposes of academic inquiry. Reflecting these trends, an increasing number of memory institutions—including libraries, museums, and archives—are collecting and providing access to video games as part of our cultural heritage. For instance, the National Videogame Archive in Great Britain and the Strong National Museum of Play focus on collecting and preserving hardware, original software, and other artifacts such as design documents, marketing materials, and company-specific archival items.

Recently, the GAMER Group started collaborating with the VGHF to catalog, classify, and preserve the VGHF’s collection of artifacts surrounding the digital game development process. Unlike most collections that focus on final, released games, this collection includes pre-release versions of games, marketing and sales materials, and internal documentation, which can provide unique insight into a game’s history. Many of these assets were generated prior to a game’s completion, often revealing the game as a work in progress, and the cultural and societal contexts in which the game was created. This development-related information is valuable to students and scholars studying games, in terms of technical development, business processes, and their social and historical context, akin to giving the public access to the sketchbooks, notebooks, and unpublished drafts of renowned artists, authors, and musicians. These collections preserve the often inaccessible historical contexts of one of the most important global media forms developed since the 1970s, one to which the US makes significant contributions.

Currently, there is limited peer-reviewed research regarding the preservation of these artifacts despite their importance. Furthermore, existing standards related to bibliographic descriptions such as Functional Requirements for Bibliographic Records (FRBR) and Resource Description and Access (RDA) are established with an emphasis on published materials rather than artifacts resulting from a creation process, and thus have limited applicability for representing and describing such collections.

Project Goal and Research Questions: This research project aims to build a conceptual data model and metadata schema for representing artifacts related to the development of video games to complement ongoing research on video game metadata. Our primary research questions are:

RQ1. Which information entities, attributes, roles, and relationships are relevant and important to describe the artifacts created during the development of video games and interactive media?

RQ2. Which information entities, attributes, roles, and relationships are perceived as valuable by various user groups/stakeholders (including librarians, museum curators, game researchers, students, and users of library and museums) interested in game development artifacts?

RQ3. What issues and challenges emerge as we apply our model and schema to describe existing collections?

The research findings will improve our theoretical understanding of the entities and their relationships within the domain of game development. We will ground the research and field test our findings by developing metadata records for the collection owned by the VGHF and museums/libraries/archives with similar materials in their collections.

Prior Work and Collaborations: The PI, Jin Ha Lee, will lead this work, leveraging her research expertise in metadata for music, multimedia, and interactive media; evaluation of IR systems; and user studies. She has

¹ <https://gamehistory.org/> <https://gamehistory.org/>

² <http://www.theesa.com/about-esa/industry-facts/>

extensive experience in representing metadata for game, music, and other multimedia objects based on data collected from qualitative and quantitative user studies. Since 2011, the GAMER Group has been working with Pacific Northwest not-for-profit organizations— including the Seattle Interactive Media Museum (SIMM), Living Computers: Museum + Labs (LCM+L), and the UW Libraries—and collaborating with scholars at universities around the globe to develop, test, and publish recommendations for metadata related to digital games, resulting in the Video Game Metadata Schema (VGMS). The VGMS has been made freely available through the GAMER Group website³ and the Open Metadata Registry⁴, and some related CVs have been given Library of Congress source codes. This open work has benefitted our partners and other public and private institutions and organizations (including LCM+L and Michigan State University) who have adopted it. The PI will work closely with the VGHF staff, whose extensive domain knowledge will be critical in creating the conceptual model and metadata schema. As an emergent organization with an extensive uncatalogued collection of development artifacts, much of which exists on aging, volatile storage media, the VGHF provides an ideal environment for creating and testing any emerging model and schema. Furthermore, the VGHF is well situated to work with museums with extensive video game collection and has recently started discussions for collaboration with museums such as the Strong National Museum of Play.

Proposed Work Plan: The first year will primarily focus on cooperative development of the model and schema, combining top-down (consulting domain experts) and bottom-up (analyzing existing data structures and records of game artifacts) modeling approaches. The second year will involve testing the model and schema with applicable user groups (including museum curators, librarians, archivists, game researchers, students, and library and museum patrons) by employing interviews and participatory design sessions. In addition to budgeted resources, the effort will include volunteer graduate and undergraduate researchers using the methods, processes, and lessons that made VGMS development successful.

Contributions: The main scholarly contribution of our work is a theorized understanding of how to best represent the entities and relationships in the domain of video games and interactive media development. This research will build upon the VGMS and extend other standards such as the FRBR or CIDOC Conceptual Reference Model. We also anticipate advancements in the fields of library and information sciences, and understanding the requirements for archiving and preserving digital materials. Many other burgeoning forms of electronic, born-digital, and interactive media (e.g., e-books, computer software, digital images, and smartphone applications) can benefit from research on non-book-focused metadata. Inquiry into the differences in metadata needs for physical and digital artifacts speaks to larger questions about the transition from physical to digital materials and the implications of that transition for libraries, archives, and museums. Our work may also be applicable to other media objects, such as film and animation, which are also created from a complex development process where many artifacts are produced. At UW, iSchool graduate students will have the opportunity to work with the project's frameworks through its annual Metadata for Interactive Media course.

Beyond academic research, our conceptual data model and metadata schema will enable catalogers to describe video games and related materials more accurately and thoroughly, improving the quality of metadata shared among organizations and used to locate materials. Standardized and thorough descriptions of the artifacts related to video game development can offer improved access to museum curators, archivists, and librarians who acquire, catalog, and provide reference services for these materials; users of these collections; and commercial organizations in the video game industry trying to organize and preserve their own development assets. As an increasing number of organizations adopt this new open standard, it will support interoperability among institutions, facilitating new consortia and resource sharing opportunities, as the VGMS has. As the proposed project is in partnership with the VGHF, there is already a guarantee of adoption by that institution.

Estimated Budget: The proposed research project requests \$249, 996. This request includes .85 FTE of one month of the PI's time and funding of one doctoral student for two academic years. It also requests \$4,000 for travel each year, for presentations of work in progress and consultations with organizations with extensive game-related collections.

³ http://gamer.ischool.uw.edu/official_release/

⁴ <http://metadataregistry.org/schema/show/id/132.html>