View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Deep Blue at the University of Michigan

Reconstructing the development of legacy database platforms

Alexandria J. Rayburn, Andrea K. Thomer
University of Michigan School of Information
arayburn@umich.edu, athomer@umich.edu

INSTITUTE of Museum and Library SERVICES

SCHOOL OF INFORMATION
UNIVERSITY OF MICHIGAN

Database curation to support future collection migration

Through the "Migrating Research Data Collections" project, we are working to better understand and support data collection migration and change over time. We are developing case studies of database migration at a number of memory institutions, first focusing on Natural History Museums (NHMs; Thomer, Rayburn and Tyler, 2020; Thomer, Weber and Twidale 2018).

In developing our case studies, we have found that there are surprisingly few published histories of database platforms and software – particularly for domain- or museum-specific platforms. Additionally, documentation of the function and format of legacy systems can be quite hard to find. This has motivated our efforts to reconstruct the development history of the data systems described by our study participants in interviews.

The timeline presented on this poster has been developed through review of academic publications (if any) and other documentation of these data systems. Much of this has been found through digital archival research (e.g. the Internet Archive's Wayback Machine).

The full dataset, with references, underlying this timeline is available at **bit.ly/36vaEoM**

We are using this resource to contextualize the evolution of the data systems in each of our case studies, and we further hope this work will be of interest to others studying infrastructure development and change.

Trends and observations

Surprisingly, the oldest database mentioned by our participants was not a commercial platform, but rather, was TAXIR, an information system developed to share taxonomic descriptions. TAXIR was developed in the 1960s – a full two decades before the launch of any of the other systems. This underscores our previous observation that NHMs were quite early adopters – and indeed, innovators – in database technology and information systems.

Most of the NHM-specific databases were initially funded by the National Science Foundation. Additionally, several grew out of the same research groups and grants; for instance, both Muse and TAXIR were developed by researchers at the University of Michigan, one of whom went on to build Specify at the University of Kansas. We hope to build on this observation to further explore the evolution of the biodiversity informatics community as a subset of the NHM community in future work.

In general, the commercial databases were less well documented than the NHM-specific databases. Where we found documentation for commercial systems in the form of technical specifications and archived websites, the NHM-specific platforms are typically documented through grant proposals, user manuals, and academic publications. In both cases, there is a clear need for more systematic archival of documents describing these platforms' evolution and technical specifications.

Few of the defunct platforms we surveyed reached their "end of life" in a particularly intentional way; rather, they simply went out of business with little warning. Memory institutions must take heed of this legacy, and be sure to plan for their collections management platforms' demise, as well as adoption.

TAXIR

1966 Idea for a boolean taxonomic database described in, *A General Method of Taxonomic Description for a Computed Similarity Measure*, by developers of TAXIR

1969 "The Theory of the TAXIR Accessioner" published by George Estabrook

1975 University of Michigan Herbarium collection databased in TAXIR

1980-1990 Personal computer systems began to replace mainframe operating systems, making programs like TAXIR obsolete

1988 University of Michigan Herbarium collection migrated out of TAXIR

dBase

1980 dBASE hits consumer market, released by Ashton-Tate

1982 dBASE II is launched

1986 dBASE III+ is launched

1989 *Ashton-Tate stock* falls by 43% in response to dBASE IV

1988 dBASE IV is launched

1990 Ashton-Tate sues FoxPro for copyright infringement

1991 Ashton Tate sold to Borland for $439 million

2017 dBASE PLUS 11 released

Visual FoxPro

1984 Initial design similar to dBASE II. Developed due to gaps in their market.

1985 Fox Software releases Multi-User FoxBase.

1987 Ashton-Tate tries to buy FoxPro.

1990 Ashton- Tate sues FoxPro for copyright infringement

1992 Microsoft merged with Fox Software

1995 FoxPro changes from a procedural language to an object-oriented language

2004 Visual FoxPro 9 is released, the last version of FoxPro.

2007 Company disbands

2015 Last iteration of FoxPro is supported

2020-2025 Planned obsolescence

FileMaker

1985 First version released, originally called Nutshell, and owned by Nashoba

1987 Nashoba declined a Microsoft offer to buy Filemaker

1988 Apple purchases Nashoba under their child company Claris International. Was renamed Filemaker

1992 Filemaker versions released for both Macs and Windows

1998 Apple dissolved Claris, which became FileMaker Inc.

2000 Debut of FileMaker Mobile

2011 Continued popularity U.S. public schools, universities and government agencies

2019 Can exchange data with SQL databases. Has functions for JSON and API requests.

Access

Mid 1980's Project Omega was developed and a precursor to Microsoft Access. Not as efficient as dBASE, so not very successful

1992 First version released. First major relational database program for Windows.

1995 Microsoft Access 7.0 became part of the Microsoft Office Professional Suite

1997 Microsoft Access increased the maximum database size to 2GB from 1GB

2007 MS Access rolls out a new feature allowing it to connect to SharePoint

2018 Access 2019 is scheduled to be released as part of Office 2019 suite.

MySQL

1994 MySQL was created by a Swedish company, MySQL AB

1995 First code for My SQL is released

1996 Version 3.19 is released

1998 Windows version released for Windows 95

2008 Sun Microsystms acquired MySQL AB for $1 billion

2010 Oracle acquired Sun Microsystems

2018 Version 8.0 Released. Supported until April 2026

KE Software

Late 1970's The University of Melbourne, the National Museum of Victoria and the Australian Museum, work to develop an electronic system for museums

1984 Titan Database is released from the previous mentioned work. This is the precursor to KE Software

1996 Website was created. Titan had evolved into KE Textpress. Their main database is KE Textpress

1998 KE Software website includes information on why KE Textpress is better than MS Access, Oracle, Alta Vista, and Folio Views

1999 Ke Software is willing to provide assistance for the "Year 2000" problem

2007 Growing adoption of KE Emu by a wide variety of organizations: museums, art galleries, historical societies, libraries, universities and colleges, corporations, botanical gardens and herbaria.

2014 KE EMu purchased by Axiell, Swedish company

2019 KE Emu currently hosts 200 sites in 13 countries

MUSE

1978-1981 Julian Humphries develops the idea of MUSE while obtaining a graduate degree at the University of Michigan

1988-1992 3-year NSF grant received by Julian Humphries to further improve the software

1992-1995 3-year NSF grant awarded to adapt Muse to new hardware platforms

1993 Adopted by more than 20 NHMs nationally

1996-1999 2.5 year grant from NSF to Julian Humphries and Jim Beach (founder of Specify) for development and support of the MUSE

Specify

1999 Specify 1.0 is available

1999-2002 3-year NSF grant awarded to Specify to create tools that assist in research analysis and synthesis of natural history collections

2000-2003 Specify awarded another 3-year NSF Grant for $897,508.00 to keep Specify current with commercial software components

2002-2205 NSF awards Specify another 3-year grant for $860,181.00. It is stated in the proposal that the software used by 55 collections.

2005-2010 Software supports 112 biological research collections. Awarded a 5-year NSF grant for $2,085,906.00.

2010-2016 Awarded a 6-year NSF grant for $1,858,718. Supports 189 U.S. collections located at 90 institutions.

2016 Software is used by 284 U.S. research collections to digitize, manage and publish specimen holdings data.

2016-2019 Awarded a 3- year NSF grant to derive and implement a business model to financially support its core software engineering and technical support activities.

2019 Specify still open source, now managed by a consortium that charges for collection management

Symbiota

1999-2003 NSF Awards Arizona State University a 3-year grant for Networking our Research Legacy; Infrastructure to Document, manage, and Access Ecological Data.

2009 Idea of Symbiota presented at various biological conferences nationally

2008-2012 Arizona State University awarded a 3-year grant to develop a computer-based identification system for plants and animals.

2013-2017 Symbiota adds a variety of modules to their platform, making it much more suitable to the needs of their community

2018 NSF funding for Specify ends

2018-2021 NSF awards Symbiota a 3-year grant for $283,231.00 to assist in the development of Symbiota2

1965 1970 1975 1980 1985 1990 1995 2000 2005 2010 2015 2020 2025

References

Thomer, AK, Rayburn, A, Tyler, ARB (2020). Three approaches to documenting database migrations. 15th Annual International Digital Curation Conference.

Thomer, A. K., Weber, N. M., & Twidale, M. B. (2018). Supporting the long-term maintenance of natural history museum databases. *Proceedings of the Association for Information Science and Technology*.

Acknowledgements

This project was funded by IMLS grant # RE-07-18-0118-18. Thanks to our study participants for their time and insights in this project.