

The Complete CRIS – Cayuse, a True Advancement in CRIS Software

INTRODUCTION

A common managerial approach - merger and acquisition (M&A), may be utilised to expand a business into new areas or to gain access to resources that might otherwise be unobtainable (Calipha, et al., 2010). In the technology industry this may present itself as the merging of multiple solutions, each providing a distinct piece of functionality, which together meets the requirements of the intended audience. The research technology industry is no exception, with acquisitions and mergers commonplace and used to build up a large suite of solutions benefitting Higher Education Institutions (HEIs) and their researchers. M&A will always require an answer to the boundary question: where does one solution end and another begin? Many software vendors answer this by breaking their offerings up into individual products which defines the boundary between them and works very well when functionality is conceptually distinct.

The technical undertaking to develop a Current Research Information System (CRIS) meeting all modern requirements is vast, and typically these systems were created many years ago to meet the needs from that time. Continuous development of software leads to improvements in the functionality and potentially broadening of the scope of the system, however without careful design this scope can be incredibly restricted and create unintended silos. These restrictions, loss of competitive advantage, and general changes in the market, could lead organisations and software vendors to acquire/merge with other providers (Calipha, et al., 2010). This is a strong approach to expand the offering of a business however in a closely related system, such as a CRIS, answering the boundary question (“where does one solution end and another begin?”) becomes very difficult. It is reasonably likely that when integrating a repository or bid costing solution into a CRIS the relationship between them would ideally appear borderless, and this can be closely replicated with careful UX/UI design and efficient usage of the interfaces between them. This approach may present itself as a single system to end users, however they are likely moving between the solutions reasonably frequently (depending on the coupling). The end user may not be able to tell the difference and could appear satisfied, with the system meeting or even exceeding their needs, but could these artificial barriers be damaging the potential collaboration and reducing administrative oversight within these integrated solutions? (Fuller & Kruchten, 2021). Cayuse took a very different approach – one where a combination of products engage with a single platform interface for all information exchange, facilitating the creation of a true advancement in CRIS software.

The broad purpose of a CRIS is to “store data on research activity, such as scientific publications and other outputs, researchers, research budgets and projects, research institutions, research funding applications and reviews” (Science Europe, 2016). Contrary to this definition the storage of an institution’s outputs is the responsibility and primary function of an Institutional Repository (IR). Many CRISs have introduced relationships with IRs because they “share a similar interest on the scientific publication in the academic environment” (Castro, et al., 2014). This has resulted in the gradual merging of responsibilities with institutions expecting to be able to leverage data from across the

research lifecycle, including outcomes, to get an accurate understanding of the research undertaken at the institution and any impact generated.

This is where Cayuse is advancing the quality of CRISs, by breaking down the implicit barriers resulting from merging multiple solutions. As a company Cayuse utilises software development alongside M&A to expand the offering of its suite of research administration products, with a few fundamental differences to the usual model. The most relevant in this context is that Cayuse onboards new products by integrating them with the "Platform", a single place for data exchange, while splitting the solutions up into distinct products. This is a hybrid of two common approaches: looking for clear integration points between the individual solutions while splitting them into products and attempting to appear as a single system. Both individually bring with them a few fundamental flaws however the Cayuse method ensures fluid information exchange between solutions with a robust framework for expanding its suite, rather than simply adding these new solutions using smoke and mirrors to blur the boundaries between them.

THE COMPLETE SYSTEM

The Cayuse Platform is an information management and exchange system with well-designed APIs, providing great flexibility with the different solutions that can be integrated. The broad purpose of a CRIS can be simplified to just research information management which is one of the reasons this platform allows the Cayuse suite to form an excellent modern CRIS without barriers. This suite of products was built over many years providing an abundance of knowledge and experience relevant to both the research industry as well as in integrating disjoint solutions. A recent example of the integrations with the Cayuse platform was the "Haplo Suite of Solutions". Cayuse merged with Haplo in early 2021, and since then their modules have been undergoing the onboarding process of integrating with the Cayuse platform.

Prior to entering the CRIS domain, Haplo was known for developing a postgraduate degree management application, PhD Manager. The first step towards the CRIS was implementing a system into which researchers would deposit their outputs which would then be passed on to a repository for publication to facilitate wider dissemination. Building a strong relationship with early clients led the team to understanding some of the larger issues with legacy repository software. These issues were assessed and in response the team built, and open sourced, the Haplo Repository¹. Following on from Repository, with further understanding of the pain points faced by institutions the team went on to build ethical approval, funding, and finally bid costing/budgeting modules.

Interoperability is at the core of design within the Cayuse suite, which is highlighted not only by the shared platform but also in its utilisation, consider the data management plan functionality. Utilising the RDA common standard for machine actionable DMPs allows for information around ethical issues, and costs involved with the research, to be stored as part of the plan (Miksa, et al., 2021). The Cayuse suite offers DMP functionality which can retrieve information from across the system to preload data into the plan, specifically regarding any ethical issues or costing/funding for a project. Once a dataset has been created within the institutional repository it can be linked to the project and then linked as a distribution of a dataset from the plan, this can be seen in Renner (2020). This interoperability was straightforward to develop as each solution has access to links between

¹ <https://github.com/haplo-org/haplo-repository>

records throughout the system, rather than requiring some external API to request the information from disjoint solutions.

All solutions within the Cayuse Suite, including the recently joined Haplo solutions, maintain a borderless approach to data exchange through the platform while separating the functionality out into products. The Cayuse approach to M&A has led to a “Complete” CRIS – one with all information accessible between solutions, providing simple management and oversight of the full research lifecycle.

LESSONS LEARNED

Management of the full research lifecycle is made immeasurably easier when the assisting system works with end users and administrators. It’s simpler to provide full context and information across all aspects of the research lifecycle when that information is immediately available and shared between the related components. Research assessment exercises, such as the REF in the UK, involve complex reporting to provide an insight into the quality of research at an institution. These exercises typically cross multiple areas including researcher information, outputs, projects and funding, and research impact. The ability to share information simply across each module cannot be understated as an advantage when undertaking critical reporting exercises. A complete system provides a single place to ensure compliance with relevant policies, considering all related information. While this approach could be seen as necessary to consider at design stage of the software, this approach is rather simple to add to an existing system. The core of the approach is a single place for information exchange, sometimes referred to as a data warehouse. While requiring careful technical planning and design, the approach can be retrofitted to any existing software system.

Cache invalidation is a common problem within computer science, and this is always considered in system design. Outdated, and potentially misleading, information become a risk when solutions aren’t implicitly sharing their data. The main advantage of a complete system is simply that all data within it is fully accessible by any solution within it. Hypothetically we could make any existing system complete by using a shared schema for information exchange between components within the application. The technicalities of this may differ between implementations but at its core, any component within the CRIS should be able to retrieve any existing piece of data from another. This data can then be manipulated into the component’s native schema including only relevant information to allow any necessary actions to be taken.

CONCLUSION

The industry-wide demands for more functionality within research systems has paved the way for vendors to provide a CRIS built through the merging of multiple software solutions. In a world of growing security concerns and a desire for increased accountability regarding data breaches, institutions are moving away from in-house and legacy systems and towards vendor supplied ones. The relationship between components in a merged system implicitly adds barriers to communication, increases the surface for a potential cyber-attack, and causes the system as a whole to attempt to perform actions for which it wasn’t originally designed. There are some key advantages for research cooperation and usability when going further than just integrating solutions. The careful and modular design of the Cayuse suite has led to individual products that can meet the demands of a modern CRIS. These modules function in their respective roles exceptionally however, because the relationship

between them is carefully managed through the Platform, together they allow Cayuse to shine as a true advancement in CRIS software.

BIBLIOGRAPHY

Calipha, R., Tarba, S. & Brock, D. M., 2010. Mergers and acquisitions: A review of phases, motives, and success factors. *Advances in Mergers and Acquisitions*, 21 December, Volume 9, pp. 1-24 [https://doi.org/10.1108/S1479-361X\(2010\)0000009004](https://doi.org/10.1108/S1479-361X(2010)0000009004).

Castro, P. d., Shearer, K. & Summann, F., 2014. *The gradual merging of repository and CRIS solutions to meet institutional research information management requirements*. s.l., Science Direct, p. 41 <https://doi.org/10.1016/j.procs.2014.06.007>.

Fuller, R. & Kruchten, P., 2021. Blurring Boundaries: Toward the Collective Empathic Understanding of Product Requirements. *Information and Software Technology*, Volume 140, p. <https://doi.org/10.1016/j.infsof.2021.106670>.

Miksa, T. et al., 2021. Application Profile for Machine-Actionable Data Management Plans. *Data Science Journal*, 20(1), pp. 32 <http://doi.org/10.5334/dsj-2021-032>.

Renner, T., 2020. *Turbo-charging Data Management Plans*. Dublin, s.n., pp. 15-18 <https://doi.org/10.5281/zenodo.3673058>.

Science Europe, 2016. *Position Statement on Research Information Systems*, s.l.: s.n.