REF2021

Institution: University of the West of England, Bristol

Unit of Assessment: 11

Title of case study: CRISTAL-ISE: Enabling Traceability and Change Management for Large Data Systems

Period when the underpinning research was undertaken: 2000 - 2016

Details of staff conducting the underpinning research from the submitting unit:

Name(s):	Role(s) (e.g. job title):	Period(s) employed by submitting HEI:
Richard McClatchey	Professor	August 1991 – present

Period when the claimed impact occurred: 01.08.2013 – 2020

Is this case study continued from a case study submitted in 2014? No

1. Summary of the impact

The CRISTAL software system was originally developed by UWE at CERN, through research led by Professor McClatchey, to manage complexity in the construction of one of the Large Hadron Collider experiments. UWE Researchers subsequently collaborated with business users including French company Agilium, to develop the software for commercial application. The new CRISTAL-ISE software enables users to trace the origins of data and describes how the data have been processed and transformed. With wide-ranging commercial applications, users have become more agile and responsive to customer needs, improved their efficiency and cost-effectiveness, and generated business growth, implementing systems across multiple business processes and sectors, Europe-wide.

2. Underpinning research

Organisations increasingly operate within unpredictable and rapidly-changing environments. The systems they use must therefore be able to evolve dynamically in response to changes in technology and user requirements. In order to manage the impact of these changes, there must be full traceability of data and activities between design and the evolving system specification. Providing greater accessibility and traceability of elements of data or processes in large volumes of data, often Cloud-resident, is an essential requirement in the post-Big Data era.

The UWE team developed the initial version of the CRISTAL software between 1999 and 2005 (**R1, G1**) in response to the requirements for dynamic system evolution and traceability of data that had emerged from the construction of the Compact Muon Solenoid (CMS) at CERN's Large Hadron Collider. This self-describing software system supported the CMS construction by managing over 200Tb of data while CMS requirements evolved, providing 24/7 management of key scientific data and associated workflows.

Between 2005 and 2013, further research was conducted by the UWE team at CERN to develop 'CRISTAL-ISE', a service-oriented version of the original software, based on a so-called 'description-driven framework' (**R2**). This ran on the Cloud and provided support for legacy data and their evolution. It provided a hitherto unachievable level of system flexibility and discoverability, and a novel 'provenance' data framework. An early version of the software was also introduced to Agilium to enable them to benefit from its dynamic and

traceability properties to support the development of their business process management software.

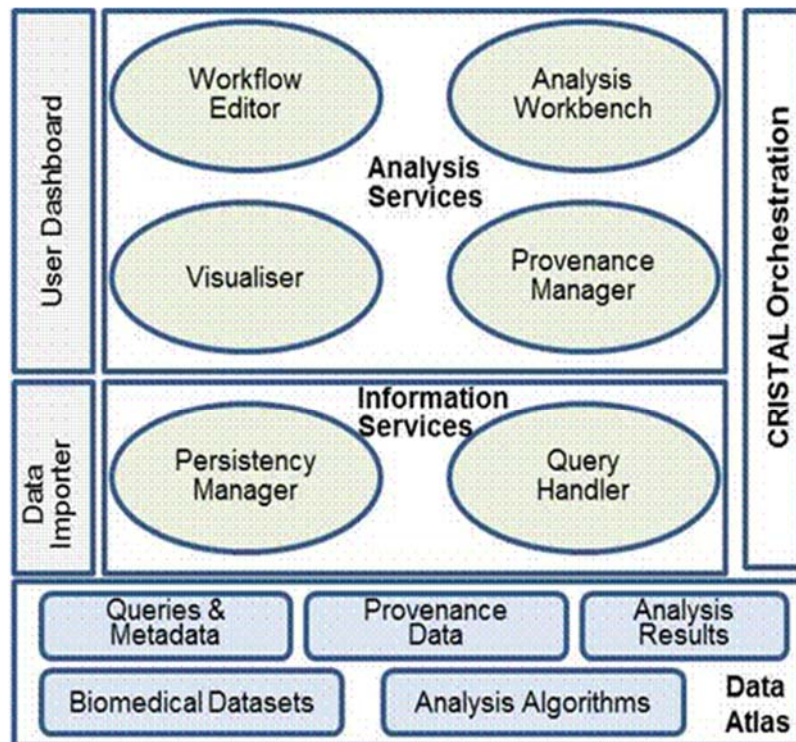The current version of the software, CRISTAL-ISE, was then developed to meet the needs of industry and clinicians by providing full design traceability, flexibility and dynamic reconfiguration of workflows. It was used to develop a Virtual Research Environment (VRE) (see Figure 1, **R3**) for a set of EU Framework projects (**G2**, **G3**) - Health-e-Child, neuGRID and neuGRID for Users (**R4**, **R5**). In the VRE, sets of biomedical data (images, laboratory results, patient records, epidemiological analyses etc) and the workflows used to process those data, together with their provenance data, are captured in the CRISTAL-ISE software. In a practical context, this has allowed clinicians to track data and process usage in a shared manner, thus promoting collaboration, analysis-informed studies, the reproducibility of results and the development of analysis gold-standards.

Figure 1 : CRISTAL as a Virtual Research Environment

Using a design based on system 'descriptions', which enriches data models with meta-data and focuses the design process on re-use, the CRISTAL-ISE software has further promoted controlled system evolution, and improved systems development, continuous maintainability, and the integration of new data with legacy systems. CRISTAL-ISE's ability to manage data and their provenance at the terabyte scale, with full traceability over extended timescales (based on its description-driven approach), has provided the adaptability required to future-proof dynamically evolving software.

The CRISTAL software was redesigned for use in new application domains via the CRISTAL-ISE project (2011-2015, **R6**, **G4**). Its V3.0 kernel was released Open Source (OS) in 2014 and the OS version was developed further over 2014-18 with refinements to data integration and business process harmonisation. The latest release, V4.0, was rolled-out to its user base in 2019-2020.

3. References to the research

R1 Estrella, F., Kovacs, Z. Le Goff, J-M., McClatchey, R., Solomonides, N. (2003) Pattern Reification as the Basis for Description-Driven Systems. *Journal of Software and System Modeling*, vol 2, no 2, pp 108-119. https://doi.org/10.1007/s10270-003-0023-0

R2 Branson, A., McClatchey, R., Le Goff, J-M., Shamdasani, J. (2014) CRISTAL: A Practical Study in Designing Systems to Cope with Change. *Information Systems*, *42*, pp 139-152. https://doi.org/10.1016/j.is.2013.12.009

R3 McClatchey, R. (2017) Data Provenance Tracking as the Basis for a Biomedical Virtual Research Environment. *Proceedings of Science 2017, Open Access journal of SISSA, Italy.* https://doi.org/10.22323/1.293.0015

R4 McClatchey, R., Branson, A., Anjum, A., Bloodsworth, P., Habib, I., Munir, K., Shamdasani, J., Soomro, K. and the neuGRID Consortium. (2013) Providing Traceability for Neuroimaging Analyses. *International Journal of Medical Informatics*, 82, pp 882-894. https://doi.org/10.1016/j.ijmedinf.2013.05.005

R5 Munir, K., Hasham Ahmad, K., McClatchey, R. (2015) Development of a Large-Scale Neuro-Imaging and Clinical Variables Data Atlas in the neuGRID4You (N4U) Project. *Journal of Biomedical Informatics,* vol 57, pp 245-262. https://doi.org/10.1016/j.jbi.2015.08.004

R6 Hasham, K., Munir, K., and McClatchey, R. (2017) Cloud Infrastructure Provenance Collection and Management to Reproduce Scientific Workflows Execution. *Future Generation Computer Systems,* vol 86, pp 799-820. https://doi.org/10.1016/j.future.2017.07.015

Evidence of the quality of the underpinning research

G1 McClatchey, R. *CRISTAL: Cooperating Repositories and an Information System for Tracking Analysis Lifecycles*, CERN, 1998 – 2002, CHF 470,000.

G2 McClatchey, R. *neuGRID*, European Commission, 2008 – 2011, €700,000.

G3 McClatchey, R. *NeuGRID for You - expansion of neuGRID services and outreach to new user communities (N4U)*, European Commission, 2011 – 2014, £569,900.

G4 McClatchey, R. *CRISTAL-ISE 2 (Marie Curie IAPP)*, European Commission, 2011 – 2015, £720,601.

4. Details of the impact

Between 2014 and 2016, CRISTAL-ISE has been exploited by companies working in the areas of business process management and production management systems: Agilium in France, and iCube and AdAcAp in Switzerland.

Business process management by Agilium
In 2011, UWE and Agilium were awarded EUR1,350,000 under the Marie Curie Industry-Academic Partnership Pathways (IAPP) programme (**G4**), a scheme specifically designed to encourage technology transfer and to rapidly promote the commercial possibilities of research. Driven by the needs of Agilium's customer base, this partnership enhanced CRISTAL-ISE with distribution, traceability and semantic features, taking advantage of a new modular plugin system to exchange server components. As a consequence, Agilium was able to reduce their development costs over the period 2014-2020, reported by Agilium's CEO as '*a direct cost-saving effect in our R&D effort*' (**S1**); they rapidly developed a business process management tool, first released in 2015, that enabled their customers to use CRISTAL-ISE for workflow and process traceability, and for the integration and co-operation of multiple business processes, especially in their business-to-business applications.

An updated product, Agilium New Generation (Agilium-NG), introduced in 2016 uses a full CRISTAL-ISE database and development framework. This enables multiple, heterogeneous processes to be integrated with each other and have their workflows tracked in the database

- improving the harmonisation of business processes. Agilium-NG has been used both by existing (pre-2014) and new customers of Agilium. Pre-2014 customers include the Ville de Lyon (France), who now use Agilium-NG for managing its operational procedures, and the specialist aerospace painting/sealing company, STTS, who use Agilium-NG for managing its internal business process controls. New customers of Agilium include Bouygues Immobilier, who use Agilium-NG for logistics management, and SUEZ, who use Agilium-NG for knowledge management in the French utilities and waste/water industries (**S1**). In addition to enabling the work of Agilium's customers, Agilium-NG itself has been a commercial success. Agilium's CEO noted that that '*the revenue since 2013 is growing rapidly, and we target to reach EUR2,000,000 per year in 2020, that is directly attributable to our adoption of the CERN-UWE developed CRISTAL-ISE software*' (**S1**).

By allowing novel methods for data integration via its description-driven design (**S2**), Agilium-NG has also united the disciplines of business process management and enterprise application integration. Existing customers now using Agilium-NG also include NEXO, GDP Vendome, Dynastar, Photowatt Technologies and the ski resorts of Tignes and Meribel. New customers using Agilium-NG include the Isere-based French transportation Group VFD and the Imprimerie Nationale Group in France (**S3**).

Tracing the production and use of radio-isotopes by iCube

CRISTAL-ISE also provided the foundation for a new application designed by Technoledge (2012-2017, **S4**) and subsequently distributed an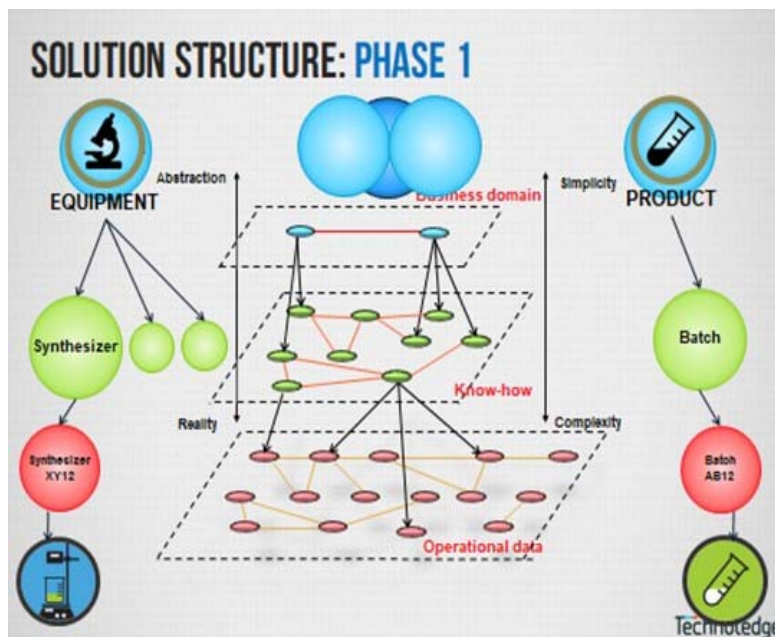d supported by iCube. The application uses the CRISTAL-ISE software to record and track the information coming from the production, delivery and use of short half-life radioisotopes used for the treatment of cancer (see Figure 2). The Associate Director of iCube, Frederic Despont, commented in 2019 that CRISTAL-ISE '*…has allowed iCube to support customers in the area of medical isotope production, and also for production in the dairy industries in Switzerland for the past six years, yielding over CHF500,000 annually in revenue during that period...*' (**S5**).

Figure 2: iCube use CRISTAL-ISE to track production and use of radioisotopes

The application has also been exploited throughout Europe by a customer of iCube - AdAcAp Smart Pharma (**S6**). AdAcAp expanded rapidly from ten employees in 2012 to over 150 by 2016, recorded income of more than EUR100,000,000 in 2016/17. AdAcAp commented '*it is very important to take prompt decisions based on data we trust … the Cristal software framework allowed us to have a secure and verified process for the collection of production and quality control data and thus simplify the product release. The integrated reporting capabilities allows us to detect out of trend and out of expectation as*

soon as they appear and make the completion of the Annual Production Report straightforward' (**S7**).

5. Sources to corroborate the impact

S1 Testimonial from CEO of Agilium

S2 McClatchey, R., Branson, A., Shamdasani, J. and Emin, P. (2017) Evolvable Systems for Big Data Management in Business. Proc of the 4th International Symposium on Big Data Principles, Architectures & Applications (BDAA 2017) at the 15th International Conference on High Performance Computing & Simulation (HPCS 2017), Genoa Italy July 2017 https://doi.org/10.1109/HPCS.2017.14

S3 Agilium website list of customers (bottom of home page) https://agilium.com

S4 Testimonial from CEO Technoledge

S5 Testimonial from Associate Director of iCube

S6 AdAcAp website https://adacap.com/our-business/

S7 Testimonial from Qualified Person for Radiopharmaceuticals at Adacap Smart Pharma, cited on iCube website