Institution: University of York

Unit of Assessment: 11 - Computer Science and Informatics

Title of case study: Epsilon: An Industrial-Strength Open-Source Toolkit for Model-Based Software Engineering.

Period when the underpinning research was undertaken: 2004 - 2020

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:
Dimitris Kolovos	Professor	2004 - present
Richard Paige	Professor	2001 - present
Fiona Polack	Senior Lecturer	1990 - 2017

Period when the claimed impact occurred: August 2013 – December 2020

Is this case study continued from a case study submitted in 2014? N

1. Summary of the impact (indicative maximum 100 words)

Epsilon is an open-source (freely available to use, modify and build upon) product developed by researchers at the University of York. It provides programming languages and tools for automating model-based software engineering tasks, such as model-based code generation, comparison and transformation of software models, and graphical domain-specific editor design. It is developed under the Eclipse Foundation, which is one of the largest open-source software foundations world-wide, supported by organisations such as IBM, Oracle, SAP and Google. Tools provided by Epsilon have been deployed across the systems and software engineering industry, in organisations such as Rolls-Royce, IBM, Leonardo, BAE Systems, Thales, Siemens, Raytheon, ATOS and NASA, and in more than 40 open-source software projects. The impact of Epsilon has been on increased developer productivity through automation of repetitive and error-prone software engineering tasks, and through enhanced reusability and interoperability.

2. Underpinning research (indicative maximum 500 words)

Epsilon is a family of programming languages and tools for model-based software engineering, which has been developed predominantly at York since 2004. It provides tools for automating repetitive, error-prone and labour-intensive software and systems engineering tasks, such as model-based code generation, comparison and transformation of software models, and graphical domain-specific modelling tool development. These tasks are common in software development projects, and automating them saves developer time and reduces the number of errors and inconsistencies, resulting in a significant impact on development costs and software quality.

Professor Dimitris Kolovos started developing Epsilon in 2004 to address the challenge of automating software development processes that involve heterogeneous models and model management (e.g. validation, transformation) activities. Professor Kolovos commenced the development of Epsilon under the supervision of Professor Richard Paige and Dr Fiona Polack and has furthered its development over a number of funded research and industrial projects since then. The core language of the platform, Epsilon Object Language (EOL), was introduced in 2006 [3.1] and is an imperative language for creating, querying and modifying software models. The Epsilon Merging Language (EML) was also introduced in 2006 [3.6] and is a rule-based language for merging homogeneous as well as heterogeneous models. In 2016, the paper describing EML received the 10-year most influential paper award at the IEEE/ACM 19th International Conference on Model Driven Engineering Languages and Systems (MODELS). The Epsilon Transformation and Generation Languages (ETL and EGL) were developed in 2008 and are languages for transforming one model to another and for transforming models into text, such as executable code or documentation, respectively. EGL is the first model-to-text transformation language to provide incremental execution capabilities [3.3]. A language dedicated to model migration (Flock) followed in 2014 [3.2]. More recently, the platform was complemented with facilities for implementing domain-specific diagram-based model editors (Eugenia) [3.1], with support for managing models constructed using commercial UML software modelling tools [3.5] and the Matlab/Simulink tool-suite [3.4], and with support for multi-threaded model querying and validation.

After attracting significant attention from researchers and practitioners over the first two years of its development as an independent open-source project, Epsilon was invited to become an

incubating project of the Eclipse Foundation in 2006. The Eclipse Foundation is one of the largest open-source software foundations world-wide, supported by leading ICT organisations such as IBM, Oracle, SAP and Google. Epsilon was promoted as a top-level project (https://eclipse.org/epsilon) in 2012 due to significant adoption and application by researchers and practitioners alike.

3. References to the research (indicative maximum of six references)
Papers [3.1-3.3] have been published in Springer's peer-reviewed Software and System Modelling journal, which is the top journal in the field of Model-Based Software Engineering. Papers [3.4-3.6] have been published in the proceedings of the ACM/IEEE MoDELS conference which is the top international conference in the field of Model-Based Software Engineering.

3.1. Kolovos, D.S., García-Domínguez, A, Rose, L.M. & Paige, R.F. Eugenia: towards disciplined and automated development of GMF-based graphical model editors. International Journal on Software and System Modeling (SOSYM), Springer, 16(1): 229-255 (2017). https://doi.org/10.1007/s10270-015-0455-3 [Journal (SOSYM) paper the introduced Epsilon's Eugenia tool] [Returned to REF 2021] [Peer-reviewed publication] [Result of the MODELPLEX EC FP7 project]

3.2. Rose, L.M., Kolovos, D.S., Paige, R.F., Polack, F.A.C. & Poulding, S.M. Epsilon Flock: a model migration language. International Journal on Software and System Modeling (SOSYM), Springer, 13(2):735-755, 2014. https://doi.org/10.1007/s10270-012-0296-2 [Journal paper (SOSYM) that introduced the Flock model migration language] [Peer-reviewed publication]

3.3. Ogunyomi, B., Rose, L.M. & Kolovos, D.S. Incremental execution of model-to-text transformations using property access traces. Software & Systems Modeling, 2018. https://doi.org/10.1007/s10270-018-0666-5 [Journal (SOSYM) paper that presented EGL's incremental model-to-text transformation capabilities] [Peer-reviewed publication] [Result of the MONDO EC H2020 project]

3.4. Sanchez Pina, B., Zolotas, A., Hoyos, H., Kolovos, D.S. & Paige, R.F. On-the-fly Translation and Execution of OCL-like Queries on Simulink Models. In Proceedings of the ACM/IEEE 22th International Conference on Model Driven Engineering Languages and Systems (MODELS), 2019: CORE: A https://doi.org/10.1109/MODELS.2019.000-1 [Conference paper, co-authored with colleagues at Rolls-Royce, which introduced the Epsilon-Simulink bridge] [Peer-reviewed publication] [Result of the SECT-AIR InnovateUK project]

3.5. Zolotas, A., Rodriguez, H.H., Kolovos, D.S., Paige, R.F. & Hutchesson, S. Bridging Proprietary Modelling and Open-Source Model Management Tools: The Case of PTC Integrity Modeller and Epsilon. In Proceedings of ACM/IEEE 20th International Conference on Model Driven Engineering Languages and Systems (MODELS), 2017: 237-247, Acceptance rate: 25%, CORE: A https://doi.org/10.1109/MODELS.2017.18 [Conference paper, co-authored with colleagues at Rolls-Royce, which introduced the Epsilon-PTC IM bridge] [Peer-reviewed publication] [Result of the SECT-AIR InnovateUK project] **[Best paper award]**

3.6. Kolovos, D.S., Paige, R.F. & Polack F. Merging Models with the Epsilon Merging Language (EML). In Proceedings of ACM/IEEE 9th International Conference on Model Driven Engineering Languages and Systems (MODELS), 2006: 215-229, Acceptance rate: 29%, CORE: A https://doi.org/10.1007/11880240_16 [Conference paper that introduced Epsilon's EML language for model merging] [Peer-reviewed publication] [Result of the MODELWARE EC FP6 project] **[10-year most influential paper award]**

4. Details of the impact (indicative maximum 750 words)

Model-based software engineering is the practice of promoting domain-specific models to first-class citizens of the software engineering process, using such models to analyse, simulate and reason about properties of the system under development, and eventually automatically generate from them (parts of) its implementation. Epsilon is an open-source platform [5.1] that

offers an integrated set of state-of-the-art languages and tools for automating many of these tasks.

Epsilon has been adopted across the systems and software engineering industry by a wide range of organisations such as Rolls-Royce, BAE Systems, Codebots, IBM, Blackbelt, DevBoost, NASA, Thales, Siemens, Raytheon, ATOS, Ergon, Talend, JC Chapman, Corpus Solutions and the Develop Group [5.8]. Epsilon is also used as a major component in several open-source projects, more than 40 of which are listed (with supporting evidence) under [5.9]. Components of Epsilon have also been used in the context of software modelling and model-based software engineering teaching in at least 15 higher education institutions worldwide [5.5].

As of December 2020, Epsilon's user forum contains more than 9,700 questions and answers posted by users and developers of the project (more than 4,700 of which have been posted to the forum since August 2013), and is the 21st largest forum (top 5%) among the 395 forums hosted by the Eclipse Foundation. The high-quality support that Epsilon's development team provides to users through the project's forum is recognised as an important contributor to Epsilon's continued adoption and impact [5.6] (quotes #59 - #64).

Below are some examples of industrial uses of Epsilon. Quotes for the first three industrial uses of Epsilon (Codebots, Blackbelt and Leonardo) have been extracted from a report [5.6] commissioned by the University of York. The report was compiled by the independent consulting company Fresh Perspectiv after interviewing several industrial users of Epsilon. The report also contains supporting quotes from organisations such as IBM and Altran and from senior management at the Eclipse Foundation.

Examples of Industrial Value

"Epsilon has a positive impact on our revenue. The core of Codebots projects use Epsilon software. Through our products it also allows our clients to get their products quicker to market." [5.6]

Epsilon is a core underpinning technology [5.6] (quote #22) of the commercial Codebots low-code software development platform (http://codebots.com/), developed by the Codebots company in Brisbane, Australia (Codebots also has offices in London and Singapore). According to the CEO of Codebots, Epsilon reduces the technical debt and costs for the company's customers [5.6] (quote #35) and allows them to ship their products to the market faster. The comprehensive support that Epsilon provides for domain-specific modelling, model validation and code generation also reduces the training effort for software developers in the company: *"Epsilon pulls all the tools into one family. It simplifies the number of software tools I need to train my team on. Epsilon has the full gamut for us."* [5.6] (quote #40). Compared to competing offerings, Epsilon is seen to provide better support for type checking and functional decomposition, leading to *"improved quality, reduced errors, and improved maintainability."* [5.6] (quote #54).

"Without Epsilon, this tool maybe wouldn't exist" [5.6]

Epsilon is also a core technology of the commercial Judo enterprise software development platform (https://www.judo.codes), developed by Blackbelt Technology in Hungary. According to Blackbelt's lead architect, Epsilon is such a fundamental technology of Judo that without it the platform might not exist [5.6] (quote #19). Epsilon's integrated nature has delivered substantial maintainability benefits for the company [5.6] (quote #18), and its openness, ease of use and versatility have also been highlighted [5.6] (quotes #43 and #65) as significant reasons for the adoption of Epsilon in the context of the Judo platform.

"[With Epsilon] we can produce more and better products in a shorter time." [5.6]

At Leonardo UK (https://uk.leonardocompany.com), Epsilon has been used since 2016 to produce software for aircraft-mounted systems which process positioning and sensor data [5.6] (quote #7). The solution developed using Epsilon *"has proven much more robust and scalable than previous ones, has made it easier to check and compare intermediate models, and has allowed the correctness of the specification transcript to be maintained even as source and destination formats evolve"* [5.10]. Epsilon increases productivity and reduces the cost of producing avionics products [5.6] (quote #25), by reducing the amount of code that developers need to write manually [5.6] (quote #52) thus increasing profit margin [5.6] (quote #17); the company estimates savings of GBP200,000 per annum as a result of this application of Epsilon [5.10]. The use of Epsilon is also seen as an improvement to the software engineers' quality of life as it eliminates repetitive and dull work and allows engineers to spend more time on solving interesting problems [5.6] (quote #68).

At JC Chapman (https://www.jcchapman.com), multiple components of Epsilon were used in 2015 to develop a platform for rigorous specification and automated translation of constrained natural language rules in the financial regulation domain (predominantly rules published by the Prudential Regulatory Authority of the Bank of England). This work was carried out in the context of a Knowledge Transfer Partnership (KTP) co-funded by the company and Innovate UK. The Epsilon-based solution developed in the KTP was able to capture 97% of the regulatory rules it was evaluated against, with ~50% less code on average compared to SQL [5.7], leading to significant productivity and maintainability benefits for customers of the company.

At BAE Systems (https://www.baesystems.com), Epsilon has been used since August 2013 (started from 2010) in the context of a model-based product for automated consistency checking of Tactical Data Link (TDL) specifications, that are used to specify the communication capabilities of defence platforms. The process of migrating from a document-based to a model-based development process using Epsilon, and the benefits realised from this transition were presented in a joint paper between academics at York and engineers at BAE Systems [5.2]. According to the engineers at BAE Systems [5.2] *"the Epsilon language components serve to provide a parsimonious implementation of a complex domain"* with the company estimating productivity benefits *"of an order of magnitude, given similar skilled levels of practitioners"*, which also extend into the maintenance phase of the product.

At Rolls-Royce (https://www.rolls-royce.com), Epsilon has been used extensively since 2017 to support model-based engineering tasks such as model validation and model-to-text transformation in the context of an in-house domain-specific model-based engineering solution, CaMCOA Studio [5.3]. CaMCOA Studio will be used to architect and integrate the software for all future Rolls Royce engine control and monitoring systems [5.4]. In the absence of Epsilon, Rolls-Royce would have needed to use a mix of independently-developed and syntactically inconsistent languages, which would hinder reuse and increase development and maintenance costs. Epsilon is seen as an integral part of Rolls-Royce's modelling toolkit and currently around 20 engineers in the company are using the tools and languages it provides. Professor Kolovos is the PI in an ongoing 3-year KTP with Rolls-Royce, co-funded by InnovateUK, to support further embedding of Epsilon technologies in CaMCOA Studio. According to the company, *"… during the development of CaMCOA Studio, several process improvement events have shown a 96% reduction in the number of man-hours compared to the process CaMCOA Studio replaces. This could result in an initial cost avoidance figure of approximately GBP2,500,000 per annum when deployed"* [5.4].

5. Sources to corroborate the impact (indicative maximum of 10 references)

5.1.   https://eclipse.org/epsilon [Epsilon's website]

5.2.   Ajit, S., Holmes, C., Johnson, J., Kolovos, D.S. & Paige, R.F. Model-based tool support for Tactical Data Links: an experience report from the defence domain. Software and System Modeling, 16(2):559-586, 2017. [Paper describing the use of Epsilon at BAE Systems]

5.3. https://www.his-2019.co.uk/session/stuart-hutchesson [Talk bySoftware Architect at Rolls-Royce discussing CaMCOA]

5.4. Letter of support from Rolls-Royce, Chief of Technology, Intelligent and Optimised Systems.

5.5. https://www.eclipse.org/epsilon/users/education/ [List of universities that use Epsilon to teach model-based software engineering, with supporting evidence]

5.6. Fresh Perspectiv. Impact Case Study: Epsilon [Report compiled by Fresh Perspective following interviews with representatives from organisations that use Epsilon.]

5.7. Barmpis, K., Kolovos, D.S. & Hingorani, J. Towards a Framework for Writing Executable Natural Language Rules. In Modelling Foundations and Applications - 14th European Conference, ECMFA 2018, Held as Part of STAF 2018, Toulouse, France, June 26-28, 2018, Proceedings, pages 251-263, 2018. [Paper describing the use of Epsilon at JC Chapman]

5.8. https://www.eclipse.org/epsilon/users/ [List of known industrial users of Epsilon, with supporting evidence]

5.9. https://www.eclipse.org/epsilon/users/open-source [List of known open-source projects that build on top of Epsilon, with supporting evidence]

5.10. https://my.sharpcloud.com/html/#/story/9517f066-5610-4308-bf0b-4be1eace4bd4/element/7eefc3ed-0968-4e0a-840b-3383500b2626 [Report from the SECT-AIR project, stating Leonardo's benefits from the use of Epsilon]