Institution: Edge Hill University (EHU)		
Unit of Assessment: B11 Computer Science and Informatics		
Title of case study: Enhancing the decision-making process for software development by using Natural Language Analysis tools		
Period when the underpinning research was undertaken: 1/1/2016 – 31/12/2020		
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:**
Dr Yannis Korkontzelos (YK) Prof Nik Bessis (NB)	Reader in Computer Science Professor in Computer Science	01/12/2015 – to date 14/09/2015 – to date
Period when the claimed impact occurred: 1/1/2017 – 31/12/2020		
Is this case study continued from a case study submitted in 2014? N		

1. Summary of the impact

Research at Edge Hill University (EHU) on Natural Language Analysis (NLA) has produced a set of NLA tools. These tools have contributed to the development of a platform that supports the adoption and reuse of Open-Source Software (OSS) and improved the software development process, increasing productivity and improving software quality. This work was done as part of CROSSMINER, a 3-year EU Horizon 2020 project, addressing the problem of choosing appropriate OSS components for any software project. The NLA tools have been integrated into the CROSSMINER platform, enabling six leading industrial software development partners across Europe, with over 5 million users and a number of major industrial clients, e.g., Allianz, Nokia and Ericsson, to enhance the process of selecting OSS components through sophisticated analysis of communication channels and documentation. The integration of the NLA tools by these companies and organisations has led to improved visualisation of OSS software and its capabilities, enhanced abilities in software selection, decrease in manual effort and ultimately provision of better services and growth of their customer bases.

2. Underpinning research

Bessis and Korkontzelos, with research expertise in Natural Language Processing (NLP), Text Mining (TM) and Big Data Analysis, have carried out research beyond the state-of-the-art in NLP/TM techniques, specifically in question/answer extraction, sentiment analysis and thread clustering for analysing information relevant to OSS. These have been reflected in leading a CROSSMINER work-package concerned with designing and developing NLA tools for enhancing the decision-making process for software development. A set of NLA tools is required to generate and visualise quality indicators that summarise the level of activity in the community of developers and users of an OSS project and the level of support that the community offers to OSS users. These indicators can facilitate evidence-based decision making, as to whether a project is mature enough to rely on and the project's community will continue to support it. The following paragraphs describe underpinning research related to the above need and specifically to document classification tasks (content recogniser, severity classifier, code detector, documentation analyser and license classifier), sentiment analysis (sentiment analyser) and emotion detection (emotion detector).

A document classification task, the classification of online communication messages into content-related classes is a multi-label task, since a message can be assigned one or more class labels. For example, a message on a bug tracker thread may notify that a previously suggested solution worked and also thank the developers for suggesting it. Automatically classifying OSS-related messages is a bespoke task without any ready-to-use systems freely available. A novel multilabel classifier for messages exchanged online about OSS has been proposed [1]. The resulting classification model has been integrated into CROSSMINER as the **content recogniser**.

Another document classification task is severity classification, i.e., identifying the importance of a bug/issue reported on bug trackers/forums linked to OSS projects. Automatic severity classification is beneficial in assessing the quality of OSS projects. Severe bugs may disrupt the usage of a piece of software completely, while less severe bugs usually concern partial errors. Bugzilla, a popular bug tracker, allows users to manually assign severity tags to bugs. A tag can be *blocker*, *critical*, *major*, *normal*, *minor*, *trivial* or *enhancement*. The distribution of tags is skewed towards *normal*, because this is the default setting and users sometimes do not change it. Existing approaches for automatic prediction of a bug's severity tag either do not consider all seven tags or assume that the classes are balanced, resulting in prediction systems that are not directly applicable to real data. A novel hierarchical classification method that considers all seven tags and imbalanced data has been proposed [2]. Imbalance and bias are tackled using a tree consisting of multiple binary classifiers. The resulting classification model has been integrated into CROSSMINER as the **severity classifier**.

Emotion classification, i.e., the identification of emotions in text, is more challenging and less popular than sentiment analysis, i.e., classification of text as positive, negative or neutral. Annotated data sets are small and rare due to their cost, subjectivity and low inter-annotator agreement. Resources, such as lexica, are scarce, due to the annotation cost and the multitude of emotion theories. The task is multi-label, i.e., more than one emotion can be expressed in a block of text. A novel emotion detection method has been proposed, using multi-label classifiers, a dataset of JIRA Issue Tracker comments, one Stack Overflow post and various emotion lexica [3]. Our classification model has been integrated into CROSSMINER, as an **emotion detector**.

Supervised classification methods can suffer from **data imbalance and outliers**, in specific domains where data classes are imbalanced and annotated data is rare due to high annotation costs. In addressing these issues, we experimented with existing data curation processes in meaningful combinations and order. We found that a particular combination and order of curation processes performs significantly better than others [4], and used this finding to improve NLA tools, such as the **content recogniser** and the **licence detector**.

To scale up to industrial-level user needs, NLA tools in general are required to provide **high throughput/efficiency**. To investigate this requirement, we selected the task of spam detection on social media, leveraging the abundance of available posts. State-of-the-art approaches for this task largely rely on historical posts of users to extract features, generating an extra overhead for the detection system, or limited features learnt by unsupervised techniques. We employed a number of machine-learning models and features related to the user, their profile on social media, or to the pairwise engagement between users to propose a novel real-time method [5]. The robustness and discriminative power of features has been ascertained using recursive elimination. In comparison to earlier studies, these features offer more discriminative power and perform consistently across different learning models. These conclusions have been considered while designing NLA tools, e.g., the **sentiment analyser** and the **code detector**.

3. References to the research

[1] Campbell, D., Cabrera-Diego, L. and Korkontzelos, I., 2019. What is the message about? Automatic multi-label classification of content found in messages relating to open source software. In: Hassanien AE., Azar A., Gaber T., Oliva D., Tolba F. (eds) Proceedings of the International Conference on Artificial Intelligence and Computer Vision (AICV2020). AICV 2020. Advances in Intelligent Systems and Computing, vol 1153, pp 520-531. Springer. https://doi.org/10.1007/978-3-030-44289-7_49

Published in a rigorously peer-reviewed international conference.

[2] Nnamoko, N., Cabrera-Diego, L., Campbell, D., and Korkontzelos, I., 2019. Bug Severity Prediction Using a Hierarchical One-vs.-Remainder Approach. In: Métais E., Meziane F., Vadera S., Sugumaran V., Saraee M. (eds) Natural Language Processing and Information Systems. NLDB 2019. Lecture Notes in Computer Science, vol 11608, pp 247-260. Springer. https://doi.org/10.1007/978-3-030-23281-8_20

Published in a rigorously peer-reviewed international conference.

[3] Cabrera-Diego, L., Bessis, N. and Korkontzelos, I., 2020. Classifying emotions in Stack Overflow and JIRA using a multi-label approach. Knowledge-Based Systems, vol. 195, 105633. https://doi.org/10.1016/j.knosys.2020.105633

Published in a rigorously peer-reviewed international journal

[4] Nnamoko, N., and Korkontzelos, I., 2020. Efficient treatment of outliers and class imbalance for diabetes prediction. Artificial Intelligence in Medicine, vol. 104, 101815. https://doi.org/10.1016/j.artmed.2020.101815

Published in a rigorously peer-reviewed international journal, this paper has attracted more than 10 citations in a year.

[5] Inuwa-Dutse, I., Liptrott, M., and Korkontzelos, I., 2018. Detection of spam-posting accounts on Twitter. Neurocomputing, vol. 315, pages 496-511, Elsevier. https://doi.org/10.1016/j.neucom.2018.07.044

Published in a rigorously peer-reviewed international journal, this paper has attracted more than 65 citations in two years.

4. Details of the impact

Context

OSS is free-to-use software developed in communities of independent developers. Developers have access to the source code and can improve it or modify it freely. The OSS Market was valued at USD 12.68 Billion in 2018 and is projected to reach USD 66.05 Billion by 2026, growing at a Compound Annual Growth Rate of 22.94%. Many commercial software developers are using OSS components to reduce costs. However, investment on OSS involves higher risk than proprietary software, as it does not come with any user support or quality guarantee. The CROSSMINER platform is a complete software suite for assessing the quality of OSS projects. The set of NLA tools form a key element that provides information about the liveliness of the community of developers and users of an OSS project and the level of user support it can offer. This is useful because software development companies need to be reassured that the OSS projects that they use in their software are of high quality, well-maintained and the community of developers supports and will support it and address related user issues. CROSSMINER is used by six leading industrial software development partners across Europe, who have also helped in the requirements collection and evaluation phases of the project. CROSSMINER's NLA capabilities have revolutionised the way that industrial partners are analysing OSS and, in turn, has led to benefits, such as simplification of OSS component selection, reduction of manual effort, better visibility of OSS projects and potential to increase customers. OSS communities, with over 5 million members, have benefitted from the use of these NLA tools, streamlining the OSS decision process during software design and development.

Impacts on software development processes

CROSSMINER has had a range of impacts that streamline and improve software development processes. These impacts can benefit a wide range of businesses and not-for-profit OSS communities. The main impacts from the inclusion of NLA tools in CROSSMINER and its subsequent adoption on software development are:

1. **Increased ability to evaluate OSS projects and make decisions about using them as software components:** the NLA tools measure the maturity of the community of developers and users of OSS projects and the support that the community provides to users. These measurements provide decision makers with evidence to choose the most suitable OSS projects and also estimate the readiness level of projects for the market.
2. **Improved software quality:** the NLA tools allow software developers and engineers to constantly monitor the development activity. Shortcomings and miscommunication are easier to spot and can be fixed more quickly, leading to the production of better software.
3. **Reduction in manual effort and cost:** NLA tools automatically monitor software for changes and inform developers about changes in software that they are using, saving the time and effort in doing this process manually.

4. **Higher visibility for OSS projects:** NLA tools can generate evidence about the health and maturity of OSS developing communities, helping software developing organisations to enrich their offering with new services and to boost **user satisfaction** and **retention**.

Impacts of CROSSMINER adoption on the European software development sector

CROSSMINER has been used and evaluated across six European software development partners and subsequently embedded into the standard business practices.

Bitergia, an SME in Spain, provides advanced quantitative analytics for software development. Bitergia has integrated our NLA tools into their visualisation tool. The output of the *sentiment analyser* and the *emotion detector* are plotted over time, showing their evolution. This visualisation offers a live health-check, an approximate psychological profile of a project, offering project managers a feature that was not available before. The *semantic clustering tool* is used to identify topics in user discussions. They can now dynamically monitor the communication of project developers and users, the liveliness of the community, the support provided to newcomers; and also compare projects based on metrics of interest. NLA tools allow Bitergia users to automatically extract objective recommendations for OSS projects, based on the topics discussed in their forums, and the sentiment and emotions expressed [S1]. This has **increased** the level of **confidence** in deciding to use particular OSS components for the development of new software, leading into **improved software quality**. Since the integration of CROSSMINER Bitergia have been able to 'provide additional services related to code quality status and evolution over time, community behaviour for fair play analysis, related projects benchmarking to understand how competitors are performing on engaging contributors and users, and finally wider community analysis based on dependent projects to have a more complete view of the project footprint' [S1]. These *additional services* have **strengthened** the company's **competitive position** within the software analytics market.

The ***Eclipse Foundation (EF)*** is a non-profit organisation offering a popular Integrated Development Environment and hosting hundreds of OSS projects. Today, EF handles 375 open-source projects, managed by 1600 committers that together deliver 195 million Lines of Code per year. Eclipse also has 5 million active users, an average of 1.5 million downloads per month, and more than 2 million unique visitors per month. EF uses our NLA tools to analyse mailing lists and forums, Bugzilla issues and Git commit messages. The *sentiment analyser*, the *emotion detector* and the *semantic clustering tool* are appreciated by the users of Eclipse project repository, since they provide rare insights into the project's community. The metrics generated by these components, displayed in dedicated dashboards, enable developers and consumers of Eclipse projects to get better insights on the project's dynamics and capabilities. According to EF, *sentiment analysis* and *emotion detection* are long-awaited features that cross the gap between the low-level, classic metrics and higher-level indicators [S2]. These features made EF's project repository unique in what they offer to the OSS community base, leading to **increased user satisfaction** (70% satisfaction amongst users of the dashboards), adding value and improving EF's **competitive advantage**. 'For the community and EF at large, CROSSMINER technologies enable EF to **increase quality**, **reliability** and **usability** of Eclipse projects' [S2].

Integrating our NLA tools benefits OSS maintainers, by showing them which issues can be targeted directly at a reasonable cost to address the shortcoming of an OSS project [S2]. As a result, maintainers can work faster, and users get a product of **higher quality** 'the sentiment detection and community polarity analysis….are long-missing features that really cross the gap between the low-level, classic metrics and higher-level indicators' [S2]. Similarly to Bitergia, EF reported that the visualised output of our NLA tools **helps decision makers** in justifying technical choices. These visualisations demonstrate how a product is fit for a specific purpose, leading to objective and factual decision-making [S2]. The impact of using our NLA tools is that decision makers are now **more confident** in making **decisions**, regardless of their level of previous experience and their familiarity with the application domain.

OW2 Consortium, a global non-profit OSS community organisation hosting 100 projects approximately, integrated CROSSMINER into their code management tool. OW2 is using our NLA tools to analyse the issue trackers of projects they host. The output provides an approximate psychological profile of a project and its evolution, allowing more than 40

organisations and more than 2.5K individual users to spot progress and thus **simplifying** the process of **selecting projects for adoption.** NLA tools were combined with other tools that OW2 is using, to score projects for Market Readiness [S3].

Our tools enable OW2 to offer detailed project and community activity evidence to project users and leaders, as decision makers. NLA tools offer a unique way to visualise the psychological state of the project community and its evolution, based on factual evidence out of forums or mail exchange threads. 'Natural Language Analysis component to projects' assessment **improves** the **ability** for project leaders **to evaluate** the psychological situation of the project community at a glance.' [S3] This was not practically possible before [S3]. Like Bitergia and Eclipse, OW2 decision makers report that they are more confident in making decisions, regardless of previous experience. 'The NLA component, provided OW2 irreplaceable input to … **measure** the **Market Readiness** Level of all mature projects in the OW2 code and evaluate them from a management perspective' [S3].

SoftTeam, a French enterprise of 1000 employees, has integrated our NLA tools into their object-oriented software modelling tool, developed in-house, that is used by 80K customers approximately and has an annual development budget of 1M Euro. In contrast to the previous practice of using low-level code-based metrics to monitor source code quality only, SoftTeam now also uses the *sentiment analyser,* the *emotion detector* and the *content recogniser* to **help** software architects **select** OSS **projects for integration** into their modelling tool. Software architects are now able to spot shortcomings that relate to the community of developers and users, such as bugs that are not addressed and communication issues or delay. This was not possible before [S4]. Spotting these risks early in the development process, can increase **customer retention** and **decrease maintenance costs**. 'Using CROSSMINER technologies, SoftTeam gains competitive advantage. It improves the quality of its IT projects and facilitates maintenance and support services.' [S4]

Unparallel Innovation, a high-tech spin-off company in Portugal that commercialises ICT research, has developed IoT-Catalogue, a decision support tool for customers choosing Arduino components for specific applications. The company is using our NLA tools to locate associated software libraries and assess their quality. In particular, Unparallel Innovation is using the *sentiment analyser* to analyse bug tracker threads to **assess** the level of **support** offered to the community and the *documentation analyser* to **assess** the **quality** of documentation. NLA 'metrics provide users evidence to reach an **informed decision** based on code maturity of libraries and the provision of support' [S5]. NLA tools allowed Unparallel Innovation to enrich the IoT-Catalogue with useful data, which, in turn, provides its customers with accurate information about products and libraries [S5]. 'NLA components automated the discovery of new products, components and libraries. This supports the enrichment of the IoT-Catalogue with new assets' [S5]. The impact of these IoT-Catalogue enhancements is **increased customer satisfaction**, **retention** and **growth**.

FrontEndART, a software quality assurance SME in Hungary, provides a tool for static source code analysis, as a standalone and as a plugin for a popular code quality inspection framework. The plugin needs to be updated frequently to support the latest version of the framework. Our NLA tools help developers by periodically and automatically checking the framework's documentation sites, forums and blogs and extracting information about updating FrontEndART's plugin. As a result, the 'NLA component has led to **significant reduction** in the **time** that developers need to invest when **updating** our … plugin' [S6]. The ultimate impact of this is **lower maintenance costs** and **increased customer satisfaction**.

5. Sources to corroborate the impact

[S1] Factual statement, Chief Executive Officer, ***Bitergia***
[S2] Factual statement, Research Relations Director, ***Eclipse Foundation***
[S3] Factual statement, Chief Executive Officer, ***OW2 Consortium***
[S4] Factual statement, Head of Modelio Research, ***Softeam***
[S5] Factual statement, Chief Executive Officer, ***Unparallel Innovation***
[S6] Factual statement, R&D Advisor, ***FrontEndART***