

<b>Institution:</b> University of Essex		
<b>Unit of Assessment:</b> 11		
<b>Title of case study:</b> Media monitoring platform founded on Essex NLP/IR insights transforms executive decision making by enabling data science start-up to flourish		
<b>Period when the underpinning research was undertaken:</b> January 2008 – April 2019		
<b>Details of staff conducting the underpinning research from the submitting unit:</b>		
<b>Name(s):</b>	<b>Role(s) (e.g. job title):</b>	<b>Period(s) employed by submitting HEI:</b>
Kruschwitz, U	Professor	1997 - 2019
Poesio, M	Professor	2000 - 2017
Chamberlain, J	Lecturer	2017 - 2020
Albakour, D	Research Officer	2010 - 2012
Martinez-Alvarez M	KTP Associate	2014 - 2015
Brill, J	KTP Associate	2019 - 2020
<b>Period when the claimed impact occurred:</b> 2014 – Dec 2020		
<b>Is this case study continued from a case study submitted in 2014?</b> N		
<b>1. Summary of the impact</b>  <p>Essex research provided the crucial insights to create a novel media monitoring platform to monitor global news for business intelligence. The platform produced through research with global executive search and selection firm MBS Group, enables scalable automation of tailored news provision as well as delivering the high quality business intelligence companies demand, more widely. The platform proved so marketable that it enabled a new company, Signal AI to establish its commercial viability with rapid growth supported by Essex research in a collaboration, winning two national KTP awards. [text removed for publication]. Signal AI has since grown from 3 to over 150 full-time employees, [text removed for publication].</p>		
<b>2. Underpinning research</b>  <p>Research led by Kruschwitz in the University of Essex's Natural Language and Information Processing (NLIP) research group demonstrated the capability for industry strength application of state of the art insights/research in Natural Language Processing (NLP) and Information Retrieval (IR). Specifically, research presented in [G1] led to the development of an enterprise-level intranet search engine for real world scalable applications [R1] and provided insights into improving suggestions to guide users [R2]. This demonstrated the capability required for [G1] through which a new media monitoring platform [R3] was developed. Subsequent research [R5, G3] advanced it further.</p> <p>Work on supporting field operatives for British Telecom with contextual search [R1] progressed enterprise search level towards performance quality comparable to mainstream Internet search engines, i.e. those able to process large quantities of information in structured mark-up (typically HTML, but also PDF and Microsoft Word formats). The work in [R1] also provided insights into processing and indexing and then retrieving document collections residing in different data silos, which importantly, is analogous to processing a heterogeneous stream of news articles. In practice, the search engine in [R1] was developed and deployed for technical support staff. This ran live for several years and was accessed by the field force of the large telecommunications company. Kruschwitz et al.'s interactive adaptive query suggestion mechanism proposed [R1] guided users' searches through suggestions (for words and phrases users enter into a search box to obtain a list of results) derived from previous interactions. Further insights into generating such suggestions to enhance search, not of the whole web but smaller and more focused collections [R2] demonstrated: (1) the usefulness of log analysis to extract query modification suggestions; (2) a more fine-grained approach than grouping search requests into sessions allows for extraction of better refinement terms from query log files [R2].</p> <p>Search method insights [R1, R2] resulted in [G2] by evidencing industry applicability. Research by Kruschwitz, Poesio and Martinez-Alvarez to build a scalable architecture for automated information</p>		

provision resulted in a novel information search and filtering platform [R3]. This aggregates, analyses and classifies news articles so that they can be matched against a client's bespoke search profile. The cloud based architecture formed by an analytics pipeline comprising: *Document Stream > Summarisation > Named Entity Recognition (NER)/ entity recognition and disambiguation (ERD) > Topic Classification > De-duplication > Clustering > Email Delivery to Client*. Each one of the 5 text analytics modules between document stream and Email delivery has a different component. Each document is processed through all the components, extending the information available for it. For instance, after the summarisation component, the system has access to the summary of the document. Near-duplication detection addresses the problem caused by republished articles in written media. Moreover, if not considered duplicates, often there are tens or hundreds of articles focused on the same information. These are addressed using clustering and event detection mechanisms. The pipeline uses a queuing system between components, allowing them to scale independently. This characteristic provides a scalable solution while minimising the complexity of the architecture. This also allows focus on specific solutions for each one of the components in order to improve the quality of the system over time [R3]. This addressed information filtering systems' growing problem of information overload, in particular to obtain insights to inform strategic decision making from disparate sources. The media monitoring platform [R3] allowed information search, filtering and summarisation on a scale and at a speed that was previously not possible as a single business service to analyse all the news of the world in real-time.

Essex NLIP researchers addressed the media monitoring platform's next stage in development in [G3]. The new approach for scalable visualisation of sentiment and stance, which addressed scalable visualisation of planning data in e-government, [R5] was presented to Signal AI, demonstrating capability to address the problem in [G3]. Moreover, the first published paper into the problem of identifying whether a news article can be identified as topical or an aggregation of different news stories [R5], showed that the process can be broken down into a two-stage approach: first segment an article into smaller units, examine possible topic shifts and then apply a neural-network-based approach akin to image processing to identify whether the article/image resembles an aggregate or a topical piece. This advanced topic detection and filtering summary articles – key to the media monitoring system's analytics modules. Narrow searches may exclude important information. Broader scope was exceeding existing monitoring products. Some articles, e.g., summaries should be excluded [R2]. Further work by Chamberlain and Brill on [G3] addressed turning a large feed of news articles into a digest of a small number of articles summarising and covering the key insights in those articles. They developed IR algorithms (based on clustering and ranking) for this (as noted in [S5]). [text removed for publication].

### 3. References to the research [can be supplied by the HEI on request]

Following peer-review: [R1] was published by Springer and has been downloaded over 11,000 times; [R2] was published in a leading international information science journal (Q1, H124, IF2.7); [R3] (demo) and [R5] (full paper) were accepted at the annual European Conference on Information Retrieval (ECIR), the premier European forum for new research in the field of Information Retrieval. [R4] was presented at LREC, an important conference in the NLP community.

[R1] M-D. Albakour, G. Ducatel, and U. Kruschwitz. The Role of Search for Field Force Knowledge Management. In *Transforming Field and Service Operations: Methodologies for Successful Technology-Driven Business Transformation, Theory and Applications of Natural Language Processing*, p117–132. Springer, 2013. DOI: 10.1007/978-3-642-44970-3\_8

[R2] U. Kruschwitz, D. Lungley, M-D. Albakour, and D. Song. (2013) Deriving query suggestions for site search. *JASIST*, 64: p1975-1994. DOI:10.1002/asi.22901.

[R3] M Martinez-Alvarez, U. Kruschwitz, W. Hall and M. Poesio. (2015) Signal: Advanced Real-Time Information Filtering. In: Hanbury A., Kazai G., Rauber A., Fuhr N. (eds) *Advances in Information Retrieval. ECIR 2015. Lecture Notes in Computer Science*, vol 9022. Springer, Cham. [https://doi.org/10.1007/978-3-319-16354-3\\_87](https://doi.org/10.1007/978-3-319-16354-3_87)

[R4] J. Chamberlain, U. Kruschwitz & O. Hoeber. (2018) Scalable Visualisation of Sentiment and Stance. Proc. of LREC 2018, Myazaki, Japan. <https://www.aclweb.org/anthology/L18-1660.pdf>

[R5] M. Fisher, M-D. Albakour, U. Kruschwitz, and M. Martinez (2019). Recognising summary articles. In Proceedings of ECIR'19, 2019. DOI: 10.1007/978-3-030-15712-8\_5.

[G1] EPSRC Grant. "Automatic Adaptation of Knowledge Structures for Assisted Information Seeking (AutoAdapt)". Udo Kruschwitz (PI), 2008 to 2012, £278,271.

[G2] KTP InnovateUK Grant with MBS (Moir Benigson Executive Search LLP). "To build a scalable technology architecture that will enable automated information provision and creates a new revenue stream." Kruschwitz (PI), Poesio (CoI), 2013 to 2015, £120,566.

[G3] KTP InnovateUK Grant with Signal Media. "Develop insight extraction and visualisation techniques to convert a stream of unstructured textual documents." Kruschwitz (PI), Chamberlain (CoI) 2019 to 2020, £127,675.

#### 4. Details of the impact

##### **Development of a novel media monitoring platform contributed to business success for MBS and the rapid expansion of Signal AI**

Company executives require regular, up-to-date news tailored to their business or sector to inform decision making. MBS Group (MBS), a leading global executive search and selection firm, found its manually produced newsletter was not providing the width and depth of coverage or the high quality, relevant, personalised external information required by its senior executives. [text removed for publication] Essex's NLIP researchers demonstrated their research was applicable to this [R1, R2, G1]. Consequently, [text removed for publication] they began a project [G2] in 2013 to build the architecture required. Their resulting media monitoring platform is a state-of-the-art, cloud-based information processing architecture solution [R3] that not only enables scalable automation of providing tailored news but proved capable of delivering the high quality business intelligence companies demand more widely. The platform analyses world news, filtering out noise and focusing on trends and strategic information, users are provided with a structured stream of documents relevant to their business needs. [text removed for publication]

Importantly, through ongoing collaboration with Essex's NLIP researchers, the underlying technology for a new company, Signal AI, was established [text removed for publication]. According to Signal AI, the collaboration with the University of Essex '*created tangible value for Signal AI by allowing us to bridge the academic and industrial ecosystems in a seamless way through application of the University of Essex's research. The first project allowed us to set the foundations of our Natural Language Processing (NLP) pipeline, Signal AI's platform, by working in collaboration with Prof Udo Kruschwitz and Prof Massimo Poesio to apply their NLP expertise*' [S2]. **The project transformed the business so significantly and led to winning in 2015 the best Knowledge Transfer Partnership award of the year** [S3]. Essex research was further embedded in Signal AI through Albakour becoming Data Scientist (in 2015) and Kruschwitz continuing as Signal AI's Senior Advisor to '*advise and guide Signal AI in all research matters*' [S4].

Essex's research was an essential part of Signal's rapid expansion. Signal AI generated its first substantial sales to several well-known brands and organisations [text removed for publication]. This project '*helped the company secure investment of over £5.8m. This played a pivotal and transformative role in the development of Signal's AI-powered media monitoring capabilities as well as the company's growth*' [S4]. In the second Innovate UK project with the company (started January 2019), Signal AI worked with Kruschwitz (until June 2019) and, leading thereafter, Chamberlain '*who provided his expertise and advice, based on his research at Essex, not only on NLP but also on Human Computer Interaction (HCI) in order to maximise the value of the new developments in our product. This project started a new product line in the company*' [S2]. Signal AI adds '*the collaboration provided us the support to move one level higher in the value chain,*

*looking at how to provide more strategic insights to our clients based on our data*. [text removed for publication]

Commenting on the company's growth, Signal AI note *'During the course of these collaborations, Signal AI has gone from three people working in a garage to over 150 full-time employees now based in offices in London, New York and Hong Kong'* [S2]. They add that Signal AI's collaborations with the University of Essex, as well as the features delivered as a result, have been a key selling point in discussions with investors *'Signal AI has obtained more than £50,000,000 of investment to date, with the most recent Series C funding round of almost £20,000,000 completed in 2019'* [S2]. [text removed for publication] Signal AI developed a new feature, Signal Briefings, which provides clients with an executive level summary of a large feed of news to save time and keep clients updated with industry news [S5, S6]. [text removed for publication]. In reporting the next generation of AI technology in the Signal AI platform in February 2020, [S6] highlights *'Signal AI, one of the leading companies transforming how business leaders make sense of the world's information'* in 2019 *'saw unprecedented success, more than doubling its customer base, international expansion into the US and Asia Pacific, and a successful Series C funding'*. Signal conclude that *'Working with Essex's natural language processing and text analytics researchers gave our start-up the technical expertise needed to differentiate ourselves in a competitive market'* CEO, Signal [S2]. This collaboration *'helped us drive innovation in the company'* [S2].

### **Signal AI's [text removed for publication] global clients benefit from the Signal media-monitoring platform**

[text removed for publication] The Signal AI platform has gone on to change the practice of its business clients in areas such as, tracking press coverage, monitoring global news, monitoring policy change and content generation and business intelligence [S7]. Public relations teams use Signal AI's platform to track news coverage of their clients, for both promotional campaigns and to monitor negative news stories for reputation management and competitor performance. The platform has enabled MBS themselves, as well as leading international law firms such as Bird and Bird and Simmons and Simmons, to save time enabling them to process more information and adding value to their business. Bird & Bird's PR team saved over an hour a day using the platform and note *"We have halved our annual expenditure on media monitoring, but expanded and improved the speed of our coverage tracking."* Head of Marketing and Communications, Bird & Bird [S7]. Simmons and Simmons, save 4-5 hours a month tracking news coverage, improving efficiency and enabling more information to be processed. The platform is also used to monitor enterprise-level legal policy and law. Deloitte, the largest professional services firm in the world, report that the platform enables over 1,200 of its clients to stay informed about changes proposed and enacted in law and regulation transforming the way they monitor and assess changes to tax law and regulation [S7].

Beyond business, the platform is used by NGOs such as Scholars At Risk, Amnesty International and the Disasters Emergency Committee to monitor and verify reports of human rights violations, and to deliver efficient responses to global disasters [S7]. Scholars At Risk, an organisation, which aims to protect threatened academic scholars around the world, report a 50% reduction in time spent monitoring global news for human rights violations, potential attacks on higher education institutes and to strategically plan campaigns and advocacy. The product has measurably helped their cause *'time saved is people saved'* Senior Program Officer, Scholars At Risk [S7]. Amnesty International is using the platform to verify eyewitness reports of human rights violations happening globally. Within the first few weeks of using this, they were able to quickly corroborate an eyewitness video of an extrajudicial execution in Mexico by locating a local Spanish language news story. Amnesty International was then able to compile sufficient evidence to call upon the authorities to perform an independent investigation. *'We were able to discover relevant news from sources we didn't know existed'* Head of Digital Verification Corps at Amnesty International [S7]. The Signal product, enables the Disasters Emergency Committee, which coordinates the delivery of efficient disaster responses in the world's poorest countries *'to have unlimited clippings during busy times without extra cost [...] the flexibility of controlling our media monitoring service from our desktops allows valued agility to keep up with unpredictable and fast moving crises'* CEO, Disaster Emergency Committee [S7].

In conclusion, the CEO of Signal AI notes '*Our partnership with the University of Essex has been long and fruitful, and it has helped shape Signal into the innovative media company it is today.*' CEO, Signal AI [S2].

#### 5. Sources to corroborate the impact

[S1] [text removed for publication]

[S2] Testimonial from Chief Data Scientist and quotes from CEO, Signal AI

[S3] Best KTP Award 2015

[S4] Signal AI Webpages Compilation

[S5] [text removed for publication]

[S6] Signal AI Introduces the Next Generation of Artificial Intelligence With the Launch of AIQ:

<https://www.businesswire.com/news/home/20200219005066/en/Signal-AI-Introduces-the-Next-Generation-of-Artificial-Intelligence-With-the-Launch-of-AIQ>

[S7] Compilation of client case studies from Signal AI website:

<https://www.signal-ai.com/customer-stories>