Institution: University of Birmingham

Unit of Assessment: UoA 11, Computer Science and Informatics

Title of case study: Providing accessibility to scientific documents to visually impaired readers through new technology

Period when the underpinning research was undertaken: 2007–present

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:
Prof. Volker Sorge	Professor of Computer Science	2002–present

Period when the claimed impact occurred: 2014–present

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

1. Summary of the impact

Professor Sorge's research on scientific document analysis led to the creation of a Speech-Rule-Engine (SRE) which in turn has led to **improved educational inclusion of visually impaired groups** engaged in mathematical subjects. Overcoming the inability of screen readers to meaningfully interpret mathematical notation has significantly enhanced **equality of access** for these groups.

The innovative SRE accessibility features **stimulated entrepreneurial activities** through being incorporated into **new products and online services** by companies and organisations **worldwide**, including Google's ChromeVox, learning management systems (Moodle, WeBWorK), and commercial educational support providers (Benetech, Texthelp). Major Learned Societies and STEM publishers (e.g. IEEE, Elsevier, OSA) **adopted** Mathjax with SRE Assistive Technology Extension as the software of choice, **enhancing and improving delivery** of accessible online mathematics to **millions of readers globally**.

2. Underpinning research

The fundamental contribution of research underpinning this impact case study was the definition and construction of a generic representation of mathematical formulae that enables them to be accessible to visually and print-impaired readers over a wide range of platforms and in commonly used web browsers.

Working with Google's Accessibility Engineering Group in 2012/13, Sorge pioneered a rule-based approach to speaking mathematical expressions on the web and implemented the **Speech-Rule-Engine (SRE)** system that is now embedded as a core feature of ChromeVox [R1]. On the strength of this research, Sorge was invited in 2015 to join the MathJax Consortium as the lead of assistive technology: "We were impressed by Volker's research ideas and work, and so after he finished at Google and returned to his post at Birmingham University we recruited him part-time to help integrate his Speech-Rule Engine […] into an extension to MathJax that would make MathJax output accessible to users with visual impairments" [MathJax Lead Developer, S1]. Sorge has since been carrying out further development of SRE as an open-source standalone system [R2].

A **semantic tree transformation** for `MathML elements` is the key concept originally designed and implemented by Sorge in the context of making mathematical notation accessible in the screenreader ChromeVox [R1]. It is a homogeneous, semantically correct and syntactically

unambiguous representation of a mathematical expression derived from its original syntactic markup by the use of pattern recognition techniques [R1, R3]. It overcomes the limitations of `Presentation MathML` which is insufficiently expressive to enable *accessible* rendering of formulae. The semantic tree provides a sufficiently rich structure for generating various forms of accessible output such as speech descriptions, Braille, simplifications, and summarisation.

Sorge further developed a novel **semantic enrichment** procedure [R3] to embed the semantic tree of a mathematical expression into its original syntactic representation. This yields semantically rich documents without altering their syntactic markup and subsequent display. It allows different applications to generate a variety of outputs (speech, tactile output, magnification, highlighting) from the same internal document format.

As a direct application of the above concepts, and with the aim of supporting users with reading disorders and visual impairments, Sorge has devised and developed the **Assistive Technology Extension for MathJax**. The impact of this work comes from the fact that mathematical web content is automatically converted into fully accessible documents regardless of their source (LaTeX, AsciiMath, MathML) and of rendered display (SVG, HTML, CSS), in the majority of web browsers. Embeddings into the rendered markup are implemented as `data attributes`, providing fast and standardised means of retrieval fully consistent with the HTML5 and ARIA standards, thus ensuring platform independence [R3]. The SRE-based speech module translates both single symbols and complex expressions. The embedded semantics enable the provision of diverse assistive features in MathJax including the following [R4 and R5]:

- An innovative responsive rendering of mathematical content through interactive collapsing and exploration of subexpressions;
- Aural rendering, including changes in intonation that provide a more natural reading experience;
- Meaningful exploration of mathematical content, providing multiple highlighting features, magnification and synchronised aural rendering;
- Personalisation, enabling the user to choose preferred styles of magnification, voicing, exploration, etc;
- Providing on-the-fly switching of speech rules using *automatically selected* domain-specific heuristics for mathematics, physics, computer science, or logic;
- Tactile rendering, enabling Nemeth Braille output on connected Braille displays.

Without SRE, the equation

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

is spoken in a standard screen reader (in e.g. Safari) as "*x equals b plus or minus b 2 4 a c 2 a displace tile equals x b pm sqrt b 2 minus 4 over 2 a*"; with SRE using ClearSpeak, it is "*x equals the fraction with numerator negative b plus or minus the square root of b squared minus 4 a c and denominator 2 a*" [R4].

To assist learning and **editing mathematics in LaTeX**, Sorge extended Emacspeak Audio Desktop by providing facilities to write and rearrange expressions and hear the effect, and, by using MathJax's error reporting mechanism, to indicate incorrect expressions and by enabling interactive exploration of rendered expressions on the fly, while editing. Syntax highlighting (e.g. bold typeface or section heading) is provided aurally and SRE summarisation features are exploited to simplify aural summaries of the input [R6].

3. References to the research

R1. Sorge, V., Chen, C., Raman, T.V., Tseng, D. Towards making mathematics a first class citizen in general screen readers. In *Proceedings of the 11th Web for All Conference* 7 April 2014 (p. 40). ACM. DOI: 10.1145/2596695.2596700.

R2. Speech-rule-engine, GitHub speech rule engine [accessed 15 October 2020].

R3. Cervone, D., Krautzberger, P., Sorge, V. Employing semantic analysis for enhanced accessibility features in MathJax. In *13th IEEE Annual Consumer Communications & Networking Conference* (CCNC) 9 January 2016 (pp. 1129–1134). IEEE. DOI: 10.1109/CCNC.2016.7444948.

R4. Cervone, D., Sorge, V. Adaptable Accessibility Features for Mathematics on the Web. In *Proceedings of the 16th Web For All 2019 Personalization-Personalizing the Web* 13 May 2019 (p. 17). ACM. DOI: 10.1145/3315002.3317567.

R5. Cervone, D., Krautzberger, P., Sorge, V. Towards universal rendering in MathJax. In *Proceedings of the 13th Web for All Conference* 11 April 2016 (p. 4). ACM. DOI: 10.1145/2899475.2899494.

R6. Sorge, V. Supporting visual impaired learners in editing mathematics. In *Proceedings of the 18th International ACM SIGACCESS Conference on Computers and Accessibility* 23 October 2016 (pp. 323–324). ACM. DOI: 10.1145/2982142.2982212.

4. Details of the impact

Accessibility tools provided by SRE have **improved educational inclusion** and **improved equality of access to educational opportunities**. They have been **adopted by global companies** and thus **contributed to innovation** and to the **improvement of standard industry practices in science publishing**. This collectively benefits millions of impaired users around the world. As a WeBWorK co-founder stated, "I cannot think of any developments in recent years that have had a bigger impact on STEM education and on-line mathematics communication in general" [S2].

Improving inclusion through access to educational and research STEM resources for visually impaired users by provision of free, open-source tools

MathJax [S3.1] is an open-source JavaScript platform for rendering mathematics in web browsers. SRE [S3.2] forms a core feature of its inbuilt Assistive Technology Extension [S3.3]. SRE's inclusion in MathJax [S3.1] led to it becoming a preferred accessibility solution. It is used in educational tools (e.g. WeBWorK, Moodle), online scientific publishing (see below), informational websites (e.g. Wikipedia, StackExchange), and blogging sites (e.g. PhysicsForums). The MathJax Lead Developer and Project Manager stated that "Because of [Sorge's] work, all of these are accessible with no extra work on the part of the site administrators; if they are using MathJax, the accessibility comes automatically. The importance of this technology to both education and scientific research cannot be overstated" [S1]. The American College Board, based in USA and associating educational institutions in over 85 countries worldwide, recommended the use of MathJax and its Assistive Technology Extension for students with disabilities as an accessibility solution for stay-home exams during the Covid-19 pandemic.

WeBWorK is an open-source online homework system for STEM courses used at over 1000 institutions worldwide [S3.4]. Accessibility for mathematics is provided by ChromeVox and MathJax. The co-founder of the The WeBWorK project states: "The work of Dr Sorge [...] has undoubtedly provided a vital benefit for STEM students with visual disabilities. [The] **accessibility options are now ubiquitous for on-line mathematics material**. The widespread advantage for mathematicians with visual impairments is enormous and continues to improve" [**S2**]. WeBWorK is supported by Mathematical Association of America and the National Science Foundation.

Bringing innovation into commercial educational products and online services for the visually impaired through the adoption of new technology

Businesses with global reach have recognised the value of the SRE to their markets and incorporated it into their accessible products.

SRE has been incorporated as a core element of **ChromeVox**, a built-in screen reader for **Google Chromebooks**, which are widely used in education. They have an estimated 40 million users worldwide and over 20% market share in Europe. Google estimates that in the US around 1.5 million of visually impaired Chromebook users benefit from improved general educational inclusion and access to mathematics in the STEM subjects [S4]. As Chromebooks currently dominate primary and secondary education in the US, **SRE is the dominant solution to make maths accessible** for visually impaired or dyslexic high school students.

Texthelp is a market leader in assistive learning solutions whose technology is used by millions of people around the world [S5]. It uses SRE in its mathematics editor EquatIO®: "SRE is not an ordinary screen reader — it 'understands' formulae [...] Our users say that this makes a real difference [...] All this gives us competitive edge in the market. We [...] can confirm that the software has been deployed to over 3 million Google accounts" [S5].

Mathshare, Benetech's interactive environment for mathematics in high schools, has employed SRE as its integral part since 2018 [S6]. SRE and MathJax also form the back-end of MathML Cloud platform — a free, open-source, web-based tool used by major publishers: "People from at least 81 countries, including India, UK and Germany, have accessed and used MathML Cloud. Over 5000 users have visited the site" [S6]. SRE-based technologies are also available via Benetech's DIAGRAM Center, an R&D organisation disseminating best-practice accessible educational tools: "We believe that a proportion of [disabled] population has already benefited from the accessibility tools such as provided by SRE by gaining fair and equal access to scientific materials" [S6].

Improving standard practices in STEM publishing through the adoption of new technology to enhance provision for the visually impaired research community

At least 19 STEM publishers provide accessible mathematics for some of their online journals through the MathJax enabled browsers. They include the Institute of Electrical and Electronics Engineers (IEEE: as of 2018, the world's largest association of technical professionals with more than 419,000 members in over 160 countries), The Optical Society of America (OSA), Elsevier, Oxford University Press, Institute of Physics (IOP) Publishing, and 17 others [S7].

IEEE provides web access to over 3 million documents with more than 8 million downloads each month. MathJax's Assistive Technology tools "have become an invaluable, irreplaceable part of our publishing ecosystem, facilitating inexpensive and unmatched quality in online presentation of accessible math for our journals. There is no better solution to fully accessible math-on-the-web" [S8].

All **OSA** journals support accessible mathematics. MathJax and its Accessibility Extensions "provide tools and features that lend themselves well to the presentation of the math [...] making the math notation easier to read and understand. We found MathJax Accessibility Extensions particularly effective when we developed our EPUB article format, which was aimed specifically at providing improved visual and aural navigation for visually impaired users" [S9]. OSA readership spans 218 countries, with over 10,000 known institutions accessing OSA content.

Elsevier has been employing MathJax on its ScienceDirect platform since 2013. The Product Manager for SEO and Accessibility stated, "While we do not collect data on the use of accessibility features, we know that improved rendering and accessibility via MathJax is particularly valued by our customers and benefits millions of users on our site. On ScienceDirect, over 550 journals contain more than 1,000 articles with MathML elements. In 2019, those titles saw 288M visits from 95M unique visitors, for a total of 514M article page views" [S10].

The following recent statistics (13 October 2020) illustrate the scope and continuing active interest in the MathJax and its Assistive Technology Extension [S3.5]:

- 1.7 billion monthly hits for all versions of MathJax on jsDelivr (a Content Delivery Network);
- 9th most popular package on jsDelivr;
- 7 million hits for the A11y extensions of the latest version only of MathJax v2.7.5;

and in the Speech-Rule-Engine alone:

- 380 million monthly hits for all versions of SRE on jsDelivr [S3.6];
- 45th most popular package on jsDelivr [S3.6];
- 29,300 monthly downloads of the SRE software via Facebook's package manager Yarn [S3.7];
- over 8,000 weekly downloads of SRE from Node Package Manager (NPM), 1,738,160 total downloads since SRE was first published on NPM [S3.8].

5. Sources to corroborate the impact

S1. MathJax: Testimony from MathJax Lead Developer and Project Manager.

S2. WeBWorK: Testimony from co-founder of the The WeBWorK project, Professor, Department of Mathematics, University of Rochester (dated 5 February 2020).

S3. Sources evidencing the scope and uptake of software:

3.1 MathJax Accessibility Features [accessed 13 January 2021]
3.2 GitHub speech rule engine [accessed 13 January 2021]
3.3 MathJax Accessibility Extension [accessed 13 January 2021]
3.4 WeBWorK Sites [accessed 30 November 2020].
3.5 JSDelivr MathJax [accessed 16 October 2020]
3.6 JSDelivr speech rule engine [accessed 13 October 2020]
3.7 Yarn speech rule engine [accessed 13 October 2020]
3.8 NPMJS speech rule engine [accessed 13 October 2020]

S4. Google: Testimony from Director of Google AI, formerly Google Director of Accessibility Engineering (dated 26 May 2020)

S5. Texthelp: Testimony from CTO and co-Founder (dated 11 February 2020).

S6. Benetech: Testimony from Director of Educational Research and Development.

S7. STEM publishers providing accessible mathematics via MathJax [accessed 6 January 2021].

S8. IEEE: Statement from IEEE Senior Director of Publishing Technologies [accessed 6 January 2020].

S9. OSA: Testimony from Senior Director, Publishing Production and Technology (dated 24 February 2020).

S10. Elsevier: Testimony from Product Manager SEO and Accessibility (dated 10 November 2020).