

Scientific Documents and Diagrams

Volker Sorge, Dragan Ahmetovic, Cristian Bernareggi and John Gardner

Abstract Scientific documents are a very specialised type of literature not only in terms of their topics and intended audience, but also in terms of their content and how it is presented. They generally use highly topical vernacular, mathematical formulas, diagrams, data visualisations, etc. While any single one of these features on its own poses a considerable accessibility problem, their combination makes the accessibility of scientific literature particularly challenging. However, with nearly all aspects of learning, teaching, and research moving to the web, there is a need to specifically address this problem for science on the web. In this chapter we present an overview of the main challenges that arise when making scientific texts accessible. We will particularly concentrate on the accessibility problem for scientific diagrams and discuss the more common techniques for making them accessible via screen-reading, sonification and audio-tactile presentation. This chapter gives an overview of the current state of the art, sketches some of the technical details on how to create accessible diagrams and closes with some open research questions.

1 Introduction

Since the beginning of this millennium we have seen massive changes in our habits to learn, teach and study the sciences. There has been a trend to online learning

Volker Sorge
University of Birmingham, UK, e-mail: v.sorge@cs.bham.ac.uk

Dragan Ahmetovic
Università degli Studi di Torino, Italy e-mail: ahmetovic.dragan@gmail.com

Cristian Bernareggi
Università degli Studi di, Italy e-mail: cristian.bernareggi@unimi.it

John Gardner
ViewPlus, Inc., Corvallis, OR, USA e-mail: john.gardner@viewplus.com

and massive open online courses (MOOC). And even where traditional teaching methods are still employed, teaching materials are moving more and more to online resources and learning management systems. We no longer research a subject by going to the library and finding a book or reading a paper, but we search the internet for information and read a relevant online article. Even in advanced academia, research is now rarely communicated in traditional paper form, or their electronic equivalent such as PDF documents, but via digital libraries and web documents. While in many countries legislation mandates the accessibility of teaching material, online or otherwise, in primary and secondary education, for advanced scientific material, as used in further and higher education, this is rarely the case. The situation is even worse in advanced academia, where publications are often only aimed at a small number of experts world-wide and accessibility considerations are generally ignored.

Documents in the traditional STEM fields¹ are often a combination of text, tables, formulas, and diagrams, each have to be made accessible with different techniques, often with different systems or in separate workflows. Access to scientific material therefore presents a major challenge for readers with visual impairments — and to some degree for readers with learning disabilities like dyslexia or dyscalculia — that goes far beyond the problem encountered in ordinary document or web accessibility.

Traditionally, scientific literature was made accessible in a manual process, generally on a by-need basis and often restricted to monographs important for teaching a particular subject. Texts would be translated into Braille, with formulas being set in specialised formats, and experts preparing tactile versions of diagrams. Alternatively, subject matter experts would make audio recordings of literature, pronouncing formulas unambiguously, giving detailed explanations of diagrams and illustrations.

In the age of the web, these traditional workflows often fail, mainly for two reasons: firstly, teaching material can now be assembled and customised quickly and easily by everyone. Thus teachers, lecturers and professors prefer to use their own notes for teaching, multiplying the amount of material that needs to be made accessible on a daily basis for teaching alone. However, even if material has been made accessible sufficiently, providing alternative text for images and formulas etc., the second main problem is that web documents are often ephemeral. That is, they can change easily overnight; a formula is updated or a diagram is changed making even the best alternative text description obsolete. Consequently, new ways of making STEM content have to be developed in the age of the web.

2 Scientific Documents and the Web

Scientific documents have a number of particularities that raise the barrier for their accessibility. We briefly give an overview of the particularly challenging components they exhibit, before focusing on diagrams for the remainder of the chapter.

¹ STEM is an acronym for Science, Technology, Engineering, Mathematics.

Highly Specialised Vernacular — Most scientific subjects come with their own particular language, use highly specialised terms that can not be found in ordinary dictionaries or that are taken from other languages, such as Latin. Consequently, using ordinary assistive technologies like screen readers can often lead to mispronunciations or incorrectness (e.g., notations like Greek letters are omitted) that are intolerable for scientific subjects, where precision of expression is often the key.

Setting screen readers to read specific words more slowly or letter by letter can help to work around these problems. However, the obvious drawback is that readers have to spend considerably more time on the text as well as lose the reading flow, which is far from ideal.

Tables — Many sciences rely on presenting data, to convey information or backup experimental results, often in a tabular form. But unlike ordinary tables, where standard row by row or column by column reading is sufficient to comprehend their content, scientific tables often need to be viewed comparatively or holistically. E.g., the distribution of zeros in a table can convey more meaning than the actual numerical values of all the other entries.

While for readers with learning impairments, techniques such as highlighting can be helpful, for readers with visual impairments that rely on extreme magnification or on screen readers, it is nearly impossible to get a picture of a table as a whole and linear exploration will generally not reveal the information as intended by the author. One solution is to employ advanced screen reading techniques such as cursor virtualisation, which can help a reader to jump between different cells of a table. In addition tables can be authored with appropriate *Accessible Rich Internet Application* (ARIA) annotations [15] to guide screen readers to a non-linear navigation.

Formulas — Mathematical, statistical or chemical formulas can be found across the majority of scientific texts. As maths accessibility is a long standing issue, there exists assistive technology specialising on mathematics [47, 8, 48] as well as some support for mathematics in general screen reading technology [45, 3, 49, 53]. However, the reading of complex formulae and the pronunciation of mathematical expressions can vary considerably over different subject areas or STEM disciplines. As a very simple example consider the imaginary number: it is normally represented by i , but in many engineering disciplines j is used, as i denotes current. Simply put, the further advanced or specialised a scientific text, the less likely it is for current screen reading technology to be sufficient to handle formulas correctly.

As formulas play an exceptional role in education and maths accessibility is a research area in its own right, they are treated separately in the next chapter.

Diagrams — Graphical illustrations are an important means of conveying information in STEM subjects and they are ubiquitous in teaching material. While good visualisations are commonly used to great effect for the sighted world, they are practically useless to a visually impaired and particularly a blind audience. Indeed, often diagrams not only complement the exposition in the text, but are used in lieu of an exposition, with the consequence that if one cannot read the diagram, one cannot understand the document.

There exist a number of taxonomies to classify diagrams. For the purpose of our discussion we will adopt the following rough division:

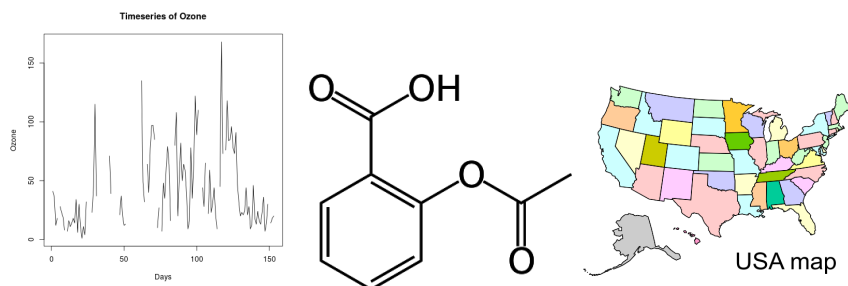


Fig. 1 Some common examples of scientific diagrams: A timeseries plot for the distribution of Ozone; a chemical molecule diagram for Aspirin; and a political map of the USA

Data Visualisations plot a relationship between two or more variables. They include histograms, function graphs, scatter plots, and time series, for which an example is presented on the left in Fig. 1. They are commonly used for the visualisation of statistical data and in other areas of mathematics either over discrete or continuous ranges of values.

Relationship Diagrams illustrate collections of items and relationships between them. Common examples are network diagrams, organisational charts, phylogenetic trees, molecule diagrams. An example of the latter is given in the middle of Fig. 1. They are generally of a graph-like nature and aim for clarity using ordered and clean layout.

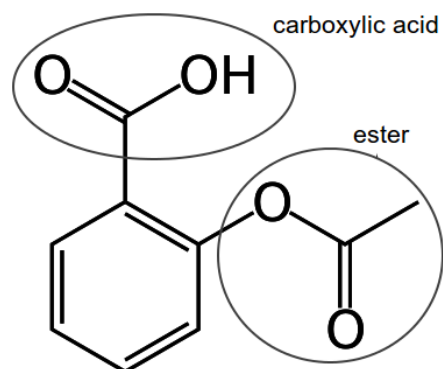
Schematics are the abstract, often simplified graphic depiction of complex systems and relationships. Examples are drawings of machinery or electric circuits in engineering, but also drawings of organisms in the life sciences or maps, such as the example in the Fig. 1. In addition to interspersed descriptive text they often use colour to distinguish and emphasise objects and relationships.

In addition to these types of diagrams, it is not uncommon to find more artistic illustrations in scientific literature, such as botanical drawings of plants. However, as these are becoming less common in advanced scientific material and in particular more modern work, they are of less practical importance.

3 Well-structured and Information-rich SVG

Traditionally diagrams on the web were embedded as bitmap images such as Jpeg, PNG, or GIF. While bitmaps can be made accessible to some degree by providing an alternative text description using the HTML `alt` attribute for image tags, this can at best be viewed as a stop gap solution. Firstly, although alt attributes are voiced by screen readers, some impose a limit on the number of words they will speak. In addition, unlike for regular text, users can generally not interact with alternative texts, such as going back or stepping through it word-by-word. Secondly, and

Fig. 2 Structural overview of the Aspirin molecule consisting of one benzene ring with two attached functional groups: ester and carboxylic acid. Note that single and parallel lines represent single and double bonds, respectively. O and H denote Oxygen and Hydrogen atoms, while carbon atoms together with attached hydrogen atoms are understood to be at junctions of bonds, giving Aspirin 9 carbon atoms altogether



more importantly, complex diagrams are difficult to describe in one-liners or even more extended text. Consider the example of the Aspirin diagram from Fig. 1: simply announcing “Aspirin” would certainly defeat the aim of the diagram to convey the molecule’s structural layout visually. Providing a description of that structure in terms of lines and characters and how they are connected would not only be very long but also make it nearly impossible for a reader to form the correct mental image of the displayed structure.

Thus, to benefit from a diagram like their sighted peers, blind readers must have a means to interact with it, to explore it step-by-step and at their own speed, and possibly to experience the diagram’s structure in alternative formats. To this end, we need to embed semantic information about the diagram into its drawing. The HTML5 web standard [5] offers its own dedicated image format with *Scalable Vector Graphics* (SVG) [11] that offers these possibilities. As a vector graphics format it allows specification of components in terms of shapes and properties with coordinates giving their position instead of black and white or coloured pixels as in bitmap graphics. SVG images are drawn by the browser rather than simply displayed, which has not only the advantage that they scale lossless visually but also that they and their components are elements of the browser’s *Document Object Model* (DOM) and can therefore be made accessible similar to any other part of the document.

One important prerequisite, however, is that the SVG is well constructed: consider the Aspirin molecule from Fig. 1 again. One can construct the SVG simply as a collection of 18 lines and 5 characters with the appropriate coordinates. While this would draw the diagram correctly, it would not capture any of the relations between the components that are so obvious to the sighted reader. But SVG not only offers elements for painting, but also elements that are not rendered and can help to structure and annotate a drawing. In particular, it offers the *container element* \mathcal{g} that permits grouping of related elements together and that can be arbitrarily nested, giving us an easy means to express semantic relations between components. We observe how this can be done with our Aspirin example as depicted in Fig. 2.

Grouping Base Elements and Ordering — A natural initial step for introducing groups is by combining drawn elements that form a single semantic entity. For ex-

ample, while double bonds consist of two lines they should be grouped together into a single entity. Elements can then be arranged in an order that roughly corresponds to their structural layout in the diagram.

Creating Abstract Containers — We can hierarchically combine elements into groups representing semantically interesting components. For example, a chemical molecule is generally composed of one or more sub-components like rings, carbon chains or functional groups. Aspirin consists of three such components: A benzene ring and two functional groups, ester and carboxylic acid. Consequently we can model these using three containers grouping the single bonds and atoms they comprise together. As a result, we get a highly structured SVG consisting of 5 individual elements: the three components and the two bonds connecting each functional group to the ring.

Embedding Alternative Structures — Often, hierarchically structuring an SVG is not enough to embed all the semantic relations we would like. For example, chemical molecules are in most cases complex graph structures. However an SVG is a tree and therefore does not have the same expressive power. It is already impossible to indicate that the benzene container is a ring using a simple structuring approach. Once we have structures where two or more abstract containers need to share one or more components, e.g., two functional groups share atoms and bonds, we need to use an auxiliary structure to hold this information.

In this case, one can represent the graph structure as a separate XML element that is either embedded into the SVG in a different namespace or as an invisible structure in the DOM of the web page. The elements in the XML graph are related to the SVG components via their unique element ids. This technique was first introduced in [50] in the context of chemical diagrams. [46] uses a similar approach to introduce an auxiliary XML structure to represent information on animated physics simulations.

Inserting Textual Information — In addition to expressing semantics via grouping, SVG offers ways to embed textual information with the `title` and `desc` elements. The former can provide short information on an element that browsers can display as tooltip on mouse hovering. The latter is used to provide more detailed description for an element. As each container or graphics element in an SVG can have associated `title` and `desc` elements, we can annotate drawings very detailed.

As an example consider the map of the USA in Fig. 1. We would annotate each state with a `title` element containing its name, while the `desc` element could hold additional information, such as the state's capital or its population.

4 Audio-Tactile Access

Converting content into tactile formats (e.g., braille notation) is a traditional way to provide access for visual impaired people to documents and graphics. Although it is generally assumed that blind people all read braille and understand graphics by feeling tactile replicas, this is not necessarily the case. Reliable statistics are not available, but estimates are that 10-15% of blind people read braille and only 2-3%

are comfortable with tactile graphics. This number decreases for people who develop a visual impairment later in life. In STEM, the percentage is almost certainly larger, but it is still small. This is compounded by the fact that it is often difficult to convey tactilely all the information that is readily available visually: tactile resolution is considerably smaller than image resolution, making it difficult to clearly separate features in crowded diagrams; colours can only be modelled to a limited extent by different textures before they become indistinguishable; text in graphics can not always be fitted as Braille and needs to be abbreviated or supplied in an extra key. All this makes pure tactile graphics often difficult and cumbersome to read.

4.1 Overview on Audio-Tactile STEM Graphics

Audio-tactile diagrams try to solve this problem by complementing the tactile experience with audio feedback that can give information to components or explain elements that are difficult to represent tactually. The concept was first introduced by Parkes in 1988 as audio/touch technique [40, 41] as a way that even non-braille readers and people with other print disabilities can access graphical information. The user obtains the two dimensional overview of a tactile graphic and hears information spoken when they indicate a text label or some graphical object.

The potential of using audio-tactiles in teaching was realised quite early [34, 33] and a particular emphasis was given to their use in teaching and examining mathematics and sciences in primary and early secondary education [18, 31]. Another major application is the creation of audio-tactile maps [37]. Technically, talking graphic uses an interface (e.g. a touch screen or finger-detecting camera) that sends spatial information to a phone, tablet, computer, or other device that then can provide spoken or non-speech audio information about whatever is at the indicated position on the graphic. Diagrams have to be represented either in proprietary file formats [31] or standard SVG format appropriately enriched with meta-data [20] as described in the previous section. The latter approach makes it particularly useful for adaptation to the Web and we will concentrate on it in the remainder of this section.

Audio-tactile diagrams are also closely related to audio-haptic approaches, where sound and haptic feedback are combined [29]. For example [25] is an approach to teach mathematics using tablets that provide normal audio feedback together with vibration when graphical components are touched.

4.2 Creating Audio-Tactile SVG

When Creating good-enough audio-tactile replicas of more complicated STEM graphics, care has to be taken to create both good tactile diagrams and speech data. In the US, the Braille Authority of North America (BANA) has established guidelines [39] for tactile graphics. Only trained transcribers can make diagrams that meet

these standards, so officially-sanctioned tactile graphics are very expensive. “Good enough” tactile graphics can be made by sighted author using common drawing software and printed using a graphics-capable embosser. For example embossers from the ViewPlus line [54] by default, emboss images as a “tactile gray scale”. Dark regions are embossed with tall dots and lighter regions with progressively lower dots. Simple line drawings, block images, and some colour images with good colour contrast can directly produce acceptable tactile images.

To add audio information, we need diagrams as SVG files with appropriate metadata. They need a meaningful title and a description explaining what the image is supposed to convey. In addition, we need to add `title` elements to all the semantically-meaningful components. If the title is not sufficiently descriptive, or if the author wants to convey related information, it should be given in the object description (i.e., the `desc` element). SVG also permits graphical features or text to be fitted to invisible overlay objects with titles and descriptions. This feature allows us to make even more informative tactile diagrams. Consider again our Aspirin diagram from Fig. 1 semantically enriched with abstract containers and descriptions for functional groups as discussed in Sec. 3. In that form a reader can only get audio feedback for the drawn elements, i.e., bonds, double bonds, oxygen, etc. To allow users to find the omitted carbon atoms we add invisible rectangles to the SVG at the junctions of the bonds with a title of “Carbon”. Similarly, we add invisible bounding polygons around the containers representing the functional groups. This has the effect that, if the reader touches, for instance, inside the ring structure “Benzene ring” will be spoken, while touching one of the ring’s boundary lines, the corresponding bond will be announced. Note that the resulting audio-tactile image still retains the original visual properties, making a graphic accessible to both sighted and non-sighted readers at the same time.

5 Screen Reading

While audio-tactile graphics present an ideal means for readers to engage with diagrams, they have two major drawbacks: firstly, reading them requires running additional, often proprietary software outside a web browser, that might not be available on all platforms. Secondly, they are relatively costly, due both to the price of embossers and to the time it takes to emboss a tactile graphic, which is a consideration if one only wants to glance briefly at a diagram when reading an article.

Consequently, there are attempts to enable screen reading and interaction with graphics directly in browsers similar to working with ordinary text and to some extent mathematics. We have already commented on the disadvantages bitmap graphics have, due to the limitations of alternative text. However, although SVG effectively offers all the technical specification that can enable effective presentation of graphical material to visually impaired readers, support for working with SVG in main stream screen readers is still relatively poor. One reason is the late adoption of SVG as an official HTML5 standard and in particular its implementation in all ma-

for browsers; in Internet Explorer SVG was unsupported until version 9. Moreover, screen readers often have problems with highly nested structures that require non-linear progression through DOM elements. Nevertheless, there are some successful approaches to making complex STEM diagrams web accessible with general screen readers, either by using ARIA constructs to guide screen readers or by turning SVG images effectively into rich web applications using JavaScript functionality.

5.1 Accessible SVG using ARIA

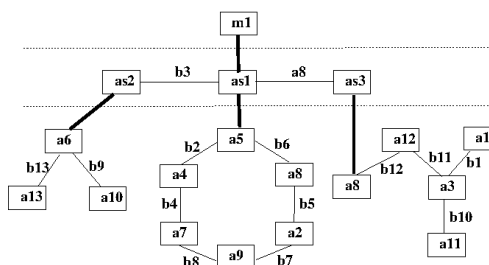
Since SVG elements implement the same interface as regular DOM elements, they are amenable to ARIA technology [15] for creating Accessible Rich Internet Applications. One such approach is implemented in the Highcharts library [38] for generating data visualisations and maps. It allows authors to put a simple accessibility layer over visualisations. Technically, this is done by giving components of a diagram an entry in the tabbing structure of the web site via the `tabindex` attribute. Speech content that we would otherwise embed as `title` or `desc` elements, can instead be added via the `aria-label` attribute. For example, when considering the US states map in Fig. 1, each SVG element representing a state gets a `tabindex="-1"` and an `aria-label` with state name and any other information we want to convey. This allows the reader to browse through the different states once the map is focused, using left and right arrow keys. The screen reader will read the information embedded in the `aria-label`. Further visual effects like highlighting can be added by change of element styling on focus events.

The advantage of this approach is that it can be employed without the need for any special application or hardware, using any screen reader on any platform. If JavaScript is available in the display engine, events can be exploited for visual style changes. However, care has to be taken on the order in which elements of the SVG are arranged as the tab order determines the order in which they can be presented to the user. This highlights the major drawback of this approach: the limitations it imposes on the freedom for the user to explore the diagram. For example, the states on the map can only be explored in a fixed order, not giving the reader the choice of seeing what neighbouring states lie north, south, east or west. Similarly, it is difficult to present information in a hierarchical manner by giving users the option to dive deeper into particular components, such as exploring cities or provinces of a state. Since the SVG elements can only be presented linearly; it is either all or nothing.

5.2 SVG as Rich Web Application

This problem can be solved by using navigation structures via bespoke XML annotations as introduced in Sec. 3. Examples of this approach are navigatable molecule diagrams [50] and data visualisations [17] as well as animated physics simula-

Fig. 3 Abstract Navigation graph for Aspirin: The labels a_i and b_i on vertices and edges correspond to ids of the drawn elements for atoms and bonds in the SVG. $as1, as2, as3$ represent the benzene ring, ester, and carboxylic acid components, respectively, and $m1$ the entire molecule



tions [46]. The basic idea is to combine the SVG with a corresponding graph data structure and use JavaScript code to drive the navigation as soon as the user inspects the diagram. The SVG itself is turned into a rich web application, by annotating it with the ARIA `role="application"` attribute. This allows a user to switch from normal reading mode into diagram exploration mode (often with keystroke `Enter`). This hands control from the screen reader to the JavaScript application until explicitly left (usually with the `Escape` key). Further communication with the screen reader is achieved by using an ARIA live region; this is a DOM element that screen readers monitor and, whenever it is updated, its new content is voiced.

We observe how this works in practice using the Aspirin molecule as an example. Fig. 3 presents the graph data structure that implements navigation on the molecule in three layers: the entire molecule ($m1$), its three main components ($as1, as2, as3$), and for each of these, the layer of atoms and bonds. When a reader enters the molecule, they can use the arrow keys to navigate the graph: initially the entire molecule will be announced; that is, “Aspirin” would be pushed to the live region and voiced by the screen reader. Keystroke `Down Arrow` will enter the next lower level and the benzene ring $as1$ will be spoken. Keystrokes `Right Arrow` and `Left Arrow` navigate to the functional groups ester ($as2$) and carboxylic acid ($as3$), respectively, while `Down Arrow` keystroke will allow the reader to walk around the single atoms of the ring.

Note that the navigation structure is connected to the SVG via ids of the drawn elements, which allows effects like synchronised navigation highlighting, zooming and magnification. Having full control via a JavaScript application also makes it easy to offer other options to the reader, like toggling expert and novice explanations of the chemistry, changing navigation styles, or switching languages. But the reliance on JavaScript leads also to a disadvantage: in environments where JavaScript is not available, for instance in some ePub readers, this approach is not suitable.

6 Sonification

Up to here our techniques to make STEM diagrams accessible relied on verbal explanations of their content. An alternative for non-visual exploration and editing of graphical representations is *sonification*, the transformation of any data relation into

non-speech sound [27, 43]. Data sonification can be achieved thanks to the ability of the human auditory system to identify even slight changes in a sound pattern so that, the amount of information conveyed through an auditory representation can be, in some cases, very close to the visual equivalent [4, 26]. Based on this perceptual capability, many different solutions have been developed, in particular: sonification systems to explore any kind of visual scenarios (including images in STEM subjects) and sonification techniques to explore exclusively images in STEM subjects.

6.1 Background

Sonification techniques can be classified into two main categories, according to the exploration method adopted [58]: *scanning* methods, in which image data is sonified in a predefined sequence, and *probing* methods, where the user can dynamically and interactively choose the portion of image to be sonified.

Examples of systems for exploring any kind of visual scenario, based on the scanning method, are “The vOICe” [36], where frames captured through a camera are sonified by mapping the vertical position of each pixel to frequency, the horizontal one to time and brightness to loudness [12], and “EyeMusic” [1], that represents colour images through pleasant sounds. The position is mapped to a note in a pentatonic scale, the luminance to the attenuation of the note and the colour to distinct musical instruments. The evaluations of both systems with blind subjects show that they can be used to explore basic geometric shapes. However, it is not clear whether they are adequate for exploring more complex representations.

Probing techniques try to leverage both sonification and proprioception² to facilitate the selection of the portion of image to be sonified. Generally, the image is presented on a touchscreen and exploration is achieved by sliding the finger over the touchscreen. The portion of image pointed by the finger is sonified. For example, “EdgeSonic” [59], uses two sonification modalities: local area and distance-to-edge. In the former, the user slides the finger over the touchscreen and once a line is touched, a sound is played. In the latter, a pulse train signal is used to represent the finger’s distance to the closest line. The game “Invisible Puzzle” [22, 21, 35] requires the player to recognise many different shapes (e.g., geometric shapes, daily life objects) through both probing and scanning methods, using 6 novel image sonification techniques. Evaluation of both systems have demonstrated that users are able to recognise geometric shapes after a few minutes training.

In addition, there has been considerable research specific for application of sonification in STEM. It is worth noting that research has investigated solutions both for all scientists and for blind people: Many different sonification techniques have been investigated to assist scientists in data analysis. E.g., to mine databases of time series through musical melodies [32], monitor network traffic [16] or to monitor real-time data streams [28]. While all of these models can be adapted for integra-

² Proprioception, in this context, refers to the awareness of the reader of their relative position in the graphic that is explored.

tion into assistive technologies [56], there has also been work specifically targeting support for blind people to sonify data visualisations — especially function graphs — and relationship diagrams, in particular graphs (in the graph theoretical sense).

In particular, the work on data visualisations has found its way into commercial products. For example, *Audio Graphing Calculator* [19] sonifies function diagrams by mapping coordinates to sound frequency. *SAS Graphics Accelerator* [44] enables blind people to interactively explore bar charts, time series plots, heat maps, line charts, scatter plots and histograms and *Desmos* [14] leverages proprioception and sonification to explore function diagrams on touchscreen devices.

The evaluation with blind people gives evidence of the effectiveness of this sonification especially for understanding the trend of the function as well as maximum or minimum points. However, it also suggests that sonification on its own is not enough. Consequently, approaches that complement sonification with verbal explanations have been developed. For example, *Audio Functions* [52] combines scanning, probing and speech. A user can explore function graphs on a tablet by listening to the sonified diagram as a whole or by sliding the finger over the touchscreen following the sonified curve or along the x-axis and listening to the sonified function value. Further information about the function diagram (e.g., concavity and point coordinates) are optionally provided by speech. Analogously, *graCALC* [24] is a graphing calculator that adds an overview containing a sonification of the diagram and an automatically generated description in order to help the blind person in the initial navigation.

For relationship diagrams, exploration of directed and undirected graphs through sonification was notably analysed in the PLUMB system [10] (later extended also to support other data structures [7]). It presents a graph on a tablet, and a continuous sound varying in pitch and loudness guides the finger tip along edges to discover nodes. Names of edges and nodes are communicated by speech. Some events relevant for exploration, such as (e.g., entering/exiting a node) are signalled through auditory icons. A similar approach, but also including editing of structures, is followed in the Graph Sketching tool (GSK) [4].

6.2 Sonification on the Web

As opposed to sonification in general, sonification for the web has been explored only relatively recently, mainly because the technical prerequisites were not available: a dedicated `audio` tag was not introduced in HTML5 until 2014. The Web Audio API [2] that allows developers to add audio effects to web pages, has only been formalised since 2011 and implementations in major browsers are not all yet fully compatible. And even then, practical obstacles remain, such as sounds often being played with delays in a browser, which makes real time sonification difficult.

Nevertheless, there is already some work exploiting these features initially for function graphs as well as other STEM graphics. Both the Desmos graphics calculator [14] and the Highcharts [38] library support the sonification of x-y plots on the

web using the Web Audio API. While the latter uses SVG representations as discussed in the previous sections, the former draws programmatically inside `canvas` elements [6]. [57] goes beyond simple plots by also sonifying charts and maps, in particular floor maps. When tracing graphics on a touch device the distance to the nearest wall are modelled with increasing and decreasing sounds. [9] introduces the auditory representations of real-time particle collision data from high energy physics. While this application is not aimed at a blind audience only, it illustrates the potential for sonification for complex data visualisation and simulations.

Finally, as a practical example, we consider sonification of a timeseries plot like the one presented in Fig. 1 that can be created using the SVG and the Web Audio API. First we note that the data in the plot is created with one or more SVG `path` elements, which is effectively a set of (x,y) -coordinate pairs $\{(x_0,y_0), \dots, (x_n,y_n)\}$, each representing the endpoint of a line. We use simple *oscillators* to create sound. An oscillator represents an audio source that generates a periodic waveform (e.g., a sine wave) for a given frequency. We then turn the coordinates of the path into frequency over time as follows: decide on a base frequency f_0 ; for simplicity we take $f_0 = \sigma(y_0)$, where σ is a suitable scaling function. Then map all remaining y coordinates to frequencies $f_i = \sigma(y_i)$. Similarly, create time n intervals $[t_{i-1}, t_i]$, $i = 1, \dots, n$ by setting $t_i = \tau(x_i)$, where τ is a suitable scaling function. We then create n oscillators o_i , $i = 1, \dots, n$ with frequency f_i . A sonification is then created by running oscillator o_i in interval $[t_{i-1}, t_i]$. Note, that the created sound will be rather crude. More pleasing results can be achieved by applying smoothing methods or modulating externally loaded sounds instead of using oscillators.

While this approach gives us a scanning method for sonification, we can easily see how one could create a probing method by combining this technique with the interactive diagram exploration from the previous section, if we allow the reader to step through recursively decreasing parts of the data plot and add a sonification to each sub-interval separately.

7 Discussion

The techniques discussed in this chapter demonstrate that there are already ways to give full access to scientific web documents to all readers, regardless of their disabilities. The evolution and availability of current web technologies (HTML5, SVG, ARIA, Web Audio, etc.) can ultimately ensure seamless non-visual exploration of graphical content dynamically generated on the web, without the need for specialist software tools or browser extensions. However, an important pre-requisite for all of the presented techniques to allow meaningful non-visual presentation of the content is that the SVG diagrams are sufficiently semantically rich, which leaves the question of how the relevant information can be obtained and embedded. There are effectively three ways to do this:

Authoring — The straight forward method is to annotate components during authoring. SVG drawing tools like Inkscape[30] permit adding title and descriptions

to each object’s properties. Similarly, the Highcharts library [38] for generating data visualisations, maps, etc. offers methods to allow authors to annotate SVG drawings programmatically.

Automatic Generation — Of course, authors often care more about visuals than accessibility, so it is always preferred if annotations can be generated automatically. Since the majority of SVG images are not drawn individually by hand but generated programmatically as output of some scientific calculation program, diagrams could be generated accessibly at point of production.

[17, 23] presents a way to produce data visualisations like the time series shown in Fig. 1 from the statistic software package R [42] that generates fully accessible diagrams directly from the statistical model. Speech rules are triggered during SVG output that add `title` and `desc` elements to each of the drawn components such as the ticks on the axes, as well as to grouped elements like the x and y axes and most importantly to the data elements.

Retrofitting — Since many diagrams are created “naively”, either by authors or software, and there is a need to make legacy content accessible, the final alternative is to retroactively add the semantic information. Specialised software like IVEO Transformer [55] can assist in transforming bit map images to good tactile copies. It can optical character recognition on text, replace colour by automatically-generated tactile patterns and add Braille for tactile images. Users can elect to put lines at edges of coloured objects and to reduce or eliminate colour fill. A “pencil” allows objects to be added or removed. Finally, Transformer provides a way to improve the tactile image without changing its visual.

Much more challenging is to automatically generate semantic information for an already existing diagram, by means of image analysis and pattern recognition. For example, [50] presents an approach that analyses bitmap images of molecule drawings like our Aspirin example, automatically recognises the chemical structure and semantically interprets it. It cannot only correctly describe all atoms and bonds, but also recognises the intermediate components, i.e., the ring and the two functional groups, and names and describes them correctly. This results in an SVG together with an external annotation structure as mentioned in the previous section, which can be used to make the molecule diagram accessible either by screen readers or as audio-tactile graphics. While this method has the advantage that even poorly authored images and legacy documents can be made accessible, its drawback is that, in case of recognition errors, incorrect descriptions might be generated.

8 Authors’ Opinion

It is our opinion that the future for fully accessible scientific documents lies in the web and its technologies. Its ubiquity and transferable standards will help to eliminate the need for creating bespoke, often platform specific specialist software tools for accessing material in different subjects. This will not only enable more robust and future-proof documents but in particular will relieve readers from the need to

buy, use and learn new software tools for every subject matter. There might be some need to extend the ARIA standard to allow for more flexible exploration techniques to avoid the need for screen reading software to become expert systems on each scientific subject. However, we should not fall into the trap of designing subject specific representation standards for the web. The example of the web standard for mathematics MathML, which has existed for many years and is yet to be implemented in most browsers, demonstrates that these efforts can easily become cul-de-sacs. Instead we should acknowledge that the representation standards are already there, and we only have to fill them meaningfully. In other words the most important task is to close the semantic gap between representation and accessibility, by retrieving, generating and fitting the necessary information into visual content ideally without manual intervention.

Unfortunately, it would be too much to hope for authors to change their ways of preparing scientific documents to be accessible out of the box. However, many components like diagrams, formulas, tables are generated with scientific software in which this information is readily available but often ignored when the visual output is generated. There is therefore a need to both educate system developers about accessibility needs and to provide bespoke accessibility extensions. Scientifically there is plenty of demand for developing techniques in image analysis, pattern recognition and semantic analysis as well as exploiting advances in machine learning and AI to recognise, understand, classify and semantically enrich existing visual components to make documents fully accessible.

9 Future Directions

Immediate future work should consist of porting existing accessibility techniques to the web. For example, audio-tactile enabled SVGs can be embedded into websites, but to take full advantage of their capabilities they still need to be opened in special software applications. It is ongoing work to enable the audio feedback directly from within browsers. And this is by no means purely derivative work: Creating general methods that work on all software platforms and devices — desktops, tablets, phones — simultaneously are of higher scientific value, than building a bespoke system for one particular operating system, only.

Although many accessibility techniques have been investigated and evaluated with blind and visually impaired students, large-scale longitudinal studies aimed at measuring the effectiveness and the acceptance of the presented techniques in the context of STEM education have not yet been conducted [13]. These studies, however, would be of great value to aid visually impaired students in choosing suitable tools in their scientific curriculum.

It is clear that well designed, semantically rich diagrams are an important prerequisite for making scientific content accessible. To that extent effective and reliable methods for the automatic, or automated, conversion and semantic enrichment of existing diagrams are needed, combining areas of image analysis and document under-

standing. Likewise better support for editing information-rich diagrams are needed, ideally on the web. While, for example, techniques for editing graphical representations assisted by sonification exist, they are still very limited. Touchscreen and wearable devices able to localise the position of many body parts (e.g., fingers, arms, head) even in a three-dimensional space are becoming more and more widespread among visually impaired people. These devices are proving to be suitable platforms to investigate editing of graphical representations driven by sonification.

Finally, the medium of the web enables us to go beyond what is possible on static media like paper, leading to more reader engaging approaches to present science, by providing animated graphics as well as interactive simulations. For an example of interactive statistics see [51] and for physics simulations see [46]. Both projects serve as a research base for advanced accessibility techniques in interactive graphics.

10 Conclusions

Accessibility to scientific material, and in particular diagrams, is one of the most challenging tasks in accessibility research. It is not merely a niche concern, as access to education is a human right, and restricting disabled students from learning scientific subjects due to lack of accessible material is a clear discrimination. In fact the societal importance of this work can not be overestimated as the need to participate in the information society where, for instance, data visualisation commonly occurs in news, sports and even in the social media. Complex data-intensive content must be accessible by all, everywhere and on any device.

References

- [1] Abboud S, Hanassy S, Levy-Tzedek S, Maidenbaum S, Amedi A (2014) Eye-music: Introducing a “visual” colorful experience for the blind using auditory sensory substitution. *Restorative neurology and neuroscience* 32(2):247–257
- [2] Adenot P, Toy R (2018) Web audio api. W3c working draft, World Wide Web Consortium, <http://www.w3.org/TR/webaudio>, Accessed 01 Sep 2018
- [3] Apple (2018) Voiceover. <http://www.apple.com/accessibility/osx/voiceover>, Accessed 01 Sep 2018
- [4] Balik S, P Mealin S, Stallmann M, Rodman R, L Glatz M, J Sigler V (2014) Including blind people in computing through access to graphs. In: *ASSETS14 - Proceedings of the 16th International ACM SIGACCESS Conference on Computers and Accessibility*, ACM, pp 91–98
- [5] Berjon R, Faulkner S, Leithead T, Navara ED, O’Connor E, Pfeiffer S, Hickson I (2013) Hypertext Markup Language (html) version 5.0. W3c candidate rec-

- ommendation, World Wide Web Consortium, <http://www.w3.org/TR/html5>, Accessed 01 Sep 2018
- [6] Cabanier R, Mann J, Munro J, Wiltzius T, Hickson I (2015) Html canvas 2d context. W3c recommendation, World Wide Web Consortium, <http://www.w3.org/TR/2dcontext>, Accessed 01 Sep 2018
- [7] Calder M, Cohen RF, Lanzoni J, Landry N, Skaff J (2007) Teaching data structures to students who are blind. In: Proceedings of the 12th Annual SIGCSE Conference on Innovation and Technology in Computer Science Education, ACM, New York, NY, USA, ITiCSE '07, pp 87–90, DOI 10.1145/1268784.1268811, URL <http://doi.acm.org/10.1145/1268784.1268811>
- [8] Cervone D, Krautzberger P, Sorge V (2016) Towards universal rendering in mathjax. In: Proceedings of the 13th Web for All Conference, ACM, p 4
- [9] Cherston J, Hill E, Goldfarb S, Paradiso JA (2016) Sonification platform for interaction with real-time particle collision data from the atlas detector. In: Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems, ACM, CHI EA '16, pp 1647–1653, DOI 10.1145/2851581.2892295, URL <http://doi.acm.org/10.1145/2851581.2892295>
- [10] Cohen RF, Yu R, Meacham A, Skaff J (2005) Plumb: Displaying graphs to the blind using an active auditory interface. In: Proceedings of the 7th International ACM SIGACCESS Conference on Computers and Accessibility, ACM, New York, NY, USA, Assets '05, pp 182–183, DOI 10.1145/1090785.1090820, URL <http://doi.acm.org/10.1145/1090785.1090820>
- [11] Dahlström E, et al (2011) Scalable vector graphics (svg) 1.1. W3c recommendation, World Wide Web Consortium, <http://www.w3.org/TR/SVG>, Accessed 01 Sep 2018
- [12] Dallas Jr SA, Erickson AJ (1983) Sound pattern generator. US Patent 4,378,569
- [13] Davison BK (2012) Evaluating auditory graphs with blind students in a classroom. ACM Transactions on Accessible Computing 102:4–7, DOI 10.1145/2140446.2140447, URL <http://doi.acm.org/10.1145/2140446.2140447>
- [14] Desmos (2017) Graphing calculator. <http://www.desmos.com/calculator>, Accessed 01 Sep 2018
- [15] Diggs J, McCarron S, Cooper M, Schwerdtfeger R, Craig J (2017) Accessible rich internet applications (wai-aria) 1.1. W3c recommendation, International Digital Publishing Forum, <http://www.w3.org/TR/wai-aria>, Accessed 01 Sep 2018
- [16] Emsley I, De Roure D, Chamberlain A (2017) A network of noise: Designing with a decade of data to sonify janet. In: Proceedings of the 12th International Audio Mostly Conference on Augmented and Participatory Sound and Music Experiences, ACM, New York, NY, USA, AM '17, pp 36:1–36:5, DOI 10.1145/3123514.3123567, URL <http://doi.acm.org/10.1145/3123514.3123567>

- [17] Fitzpatrick D, Godfrey AJR, Sorge V (2017) Producing accessible statistics diagrams in R. In: Proceedings of the 14th Web for All Conference, ACM, p 4
- [18] Gardner J (2005) New technologies for accessible tactile math and accessible graphics. Proceedings of the National Federation of the Blind Goals for Achieving Math Accessibility Summit
- [19] Gardner JA (2002) Access by blind students and professionals to mainstream math and science. In: 8th International Conference on Computers for Handicapped Persons, Springer, pp 502–507
- [20] Gardner JA (2016) Universally accessible figures. In: 15th International Conference on Computers Helping People with Special Needs, Springer, pp 417–420
- [21] Gerino A, Picinali L, Bernareggi C, Alabastro N, Mascetti S (2015) Towards large scale evaluation of novel sonification techniques for non visual shape exploration. In: Proceedings of the 17th International ACM SIGACCESS Conference on Computers & Accessibility, ACM, New York, NY, USA, ASSETS '15, pp 13–21, DOI 10.1145/2700648.2809848, URL <http://doi.acm.org/10.1145/2700648.2809848>
- [22] Gerino A, Picinali L, Bernareggi C, Mascetti S (2015) Eyes-free exploration of shapes with invisible puzzle. In: Proceedings of the 17th International ACM SIGACCESS Conference on Computers & Accessibility, ACM, New York, NY, USA, ASSETS '15, pp 425–426, DOI 10.1145/2700648.2811335, URL <http://doi.acm.org/10.1145/2700648.2811335>
- [23] Godfrey AJR, Murray P, Sorge V (2018) An accessible interaction model for data visualisation in statistics. In: 16th International Conference on Computers Helping People with Special Needs, Springer, pp 590–597
- [24] Goncu C, Marriott K (2015) Gracalc: An accessible graphing calculator. In: Proceedings of the 17th International ACM SIGACCESS Conference on Computers & Accessibility, ACM, New York, NY, USA, ASSETS '15, pp 311–312, DOI 10.1145/2700648.2811353, URL <http://doi.acm.org/10.1145/2700648.2811353>
- [25] Gorlewicz JL, Burgner J, Withrow TJ, Webster III RJ (2014) Initial experiences using vibratory touchscreens to display graphical math concepts to students with visual impairments. *Journal of Special Education Technology* 29(2):17–25
- [26] Harada S, Takagi H, Asakawa C (2011) On the audio representation of radial direction. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, ACM, New York, NY, USA, CHI '11, pp 2779–2788, DOI 10.1145/1978942.1979354, URL <http://doi.acm.org/10.1145/1978942.1979354>
- [27] Hermann T (2008) Taxonomy and Definitions for Sonification and Auditory Display. In: Proceedings of the 14th International Conference on Auditory Display (ICAD 2008), IRCAM
- [28] Hermann T, Hunt A, Neuhoff JG (2011) *The Sonification Handbook*. Logos Publishing House

- [29] Iglesias R, Casado S, Gutierrez T, Barbero J, Avizzano C, Marcheschi S, Bergamasco M (2004) Computer graphics access for blind people through a haptic and audio virtual environment. In: *Haptic, Audio and Visual Environments and Their Applications, 2004. HAVE 2004. Proceedings. The 3rd IEEE International Workshop on, IEEE*, pp 13–18
- [30] inkscape (2017) Inkscape 0.92 — draw freely. <http://inkscape.org>, Accessed 01 Sep 2018
- [31] Landau S, Russell M, Gourgey K, Erin JN, Cowan J (2003) Use of the talking tactile tablet in mathematics testing. *Journal of Visual Impairment and Blindness* 97(2):85–96
- [32] Last M, Gorelik A (2008) Using sonification for mining time series data. In: *Proceedings of the 9th International Workshop on Multimedia Data Mining: Held in Conjunction with the ACM SIGKDD 2008*, ACM, New York, NY, USA, MDM '08, pp 63–72, DOI 10.1145/1509212.1509220, URL <http://doi.acm.org/10.1145/1509212.1509220>
- [33] Loetzsch J, Roedig G (1996) Interactive tactile media in training visually handicapped people. *Institut National de la Sante et de la Recherche Medicale: colloques et seminaires* pp 155–160
- [34] Löttsch J (1994) Computer-aided access to tactile graphics for the blind. In: *4th International Conference on Computers for Handicapped Persons*, Springer, pp 575–581
- [35] Mascetti S, Gerino A, Bernareggi C, Picinali L (2017) On the evaluation of novel sonification techniques for non-visual shape exploration. *ACM Transactions on Accessible Computing* 9(4):13:1–13:28, DOI 10.1145/3046789, URL <http://doi.acm.org/10.1145/3046789>
- [36] Meijer PB (1992) An experimental system for auditory image representations. *Biomedical Engineering, IEEE Transactions on* 39(2)
- [37] Miele JA, Landau S, Gildea D (2006) Talking tmap: Automated generation of audio-tactile maps using smith-kettlewell's tmap software. *British Journal of Visual Impairment* 24(2):93–100
- [38] Moseng Ø, et al (2018) Highcharts — make your data come alive. <http://highcharts.com>, Accessed 01 Sep 2018
- [39] of North America BA (2010) Guidelines and standards for tactile graphics. <http://www.brailleauthority.org/tg>, Accessed 01 Sep 2018
- [40] Parkes D (1988) Nomad, an audio-tactile tool for the acquisition, use and management of spatially distributed information by visually impaired people. In: *Proc. of 2nd Int. Symposium on Maps and Graphics for Visually Handicapped People*
- [41] Parkes D (1991) Nomad: enabling access to graphics and text based information for blind, visually impaired and other disability groups. In: *Technology for People with Disabilities*, vol 5, pp 690–714
- [42] R Core Team (2017) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, URL www.R-project.org

- [43] Sarkar R, Bakshi S, Sa PK (2012) Review on image sonification: a non-visual scene representation. In: *Recent Advances in Information Technology (RAIT), 2012 1st International Conference on*, IEEE, pp 86–90
- [44] SAS (2017) *Sas graphics accelerator*. <http://support.sas.com/software/products/graphics-accelerator>, Accessed 01 Sep 2018
- [45] Scientific F (2018) *Jaws*. <http://freedomscientific.com/Products/Blindness/JAWS>, Accessed 01 Sep 2018
- [46] Smith TL, Greenberg J, Reid S, Moore EB (2018) Parallel dom architecture for accessible interactive simulations. In: *Proceedings of the Internet of Accessible Things*, ACM, p 4
- [47] Soiffer N (2005) *Mathplayer: web-based math accessibility*. In: *Proceedings of the 7th int. ACM SIGACCESS conference on Computers and accessibility*, ACM, pp 204–205
- [48] Sorge V (2018) *Speech rule engine version 2.2.2*. <http://github.com/zorkow/speech-rule-engine>, Accessed 01 Sep 2018
- [49] Sorge V, Chen C, Raman T, Tseng D (2014) Towards making mathematics a first class citizen in general screen readers. In: *11th Web for All Conference*, ACM, Seoul, Korea, pp 40:1–40:10
- [50] Sorge V, Lee M, Wilkinson S (2015) End-to-end solution for accessible chemical diagrams. In: *Proceedings of the 12th Web for All Conference*, ACM, pp 6:1–6:10
- [51] Stark PB (2018) *Sticigui*. <http://www.stat.berkeley.edu/~stark/SticiGui>, Accessed 01 Sep 2018
- [52] Taibbi M, Bernareggi C, Gerino A, Ahmetovic D, Mascetti S (2016) *Audio-functions: Eyes-free exploration of mathematical functions on tablets*. In: *15th International Conference on Computers Helping People with Special Needs*, Springer, pp 537–544
- [53] *Texthelp* (2018) *Equatio*. <http://www.texthelp.com/en-gb/products/equatio>, Accessed 01 Sep 2018
- [54] *ViewPlus* (2018) <http://viewplus.com/viewplus-embossers-at-a-glance>, Accessed 01 Sep 2018
- [55] *ViewPlus* (2018) <http://viewplus.com/product/tiger-software-suite>, Accessed 01 Sep 2018
- [56] Walker BN, Mauney LM (2010) Universal design of auditory graphs: A comparison of sonification mappings for visually impaired and sighted listeners. *ACM Transactions on Accessible Computing* 2(3):12:1–12:16, DOI 10.1145/1714458.1714459, URL <http://doi.acm.org/10.1145/1714458.1714459>
- [57] Wörtwein T, Schauerte B, Müller KE, Stiefelhagen R (2015) Interactive web-based image sonification for the blind. In: *Proceedings of the 2015 ACM on Int. Conference on Multimodal Interaction*, ACM, ICMI '15, pp 375–376, DOI 10.1145/2818346.2823298, URL <http://doi.acm.org/10.1145/2818346.2823298>

- [58] Yeo WS, Berger J (2005) A framework for designing image sonification methods. In: Proceedings of International Conference on Auditory Display
- [59] Yoshida T, Kitani KM, Koike H, Belongie S, Schlei K (2011) Edgesonic: image feature sonification for the visually impaired. In: Proc. of the 2nd Augmented Human Int. Conf., ACM, p 11