Data Visualisation for Systems Engineers

John Welford[1]

[1] WSP | Parsons Brinckerhoff,
john.welford@pbworld.com

June 2016

A case is made for data visualisation as an essential systems engineering skill. Examples of systems engineering visualisations used within the rail industry are given, along with discussion of the philosophy behind them. Finally, a comment is made on the relationship between the visualisations presented and model-based systems engineering.

Accessibility: Beginner upwards
Application: Good practice
Topic: MBSE, Communication

Background

The art and science of data visualisation has been developing since the work of William Playfair in the early nineteenth century; however, as with much of science and technology, its progress has accelerated rapidly in the last few years. We now live in an age of abundant data, more than we can really comprehend, and terms like *big data*, *infographics* and *analytics* are starting to enter the common parlance. Increasing use of the internet for communication, improvements in computer processing power, along with the prevalence of smart devices and the internet-of-things, enables presentation techniques that were not previously possible.

In the context of this paper we will use the terms such as *visualisation*, *data presentation* and *diagram* interchangeably to describe non-text-based documentation, produced with the intent of communication, and as an aggregation of raw data into useful information. This includes all kinds of graphs, models, designs, dashboards and maps. They may be produced in a variety of mediums, from a sketch on a whiteboard through to an interactive software application. Visuals can be completely static and unchanging, or highly dynamic, responding to changes in the background data and interacting with user input.

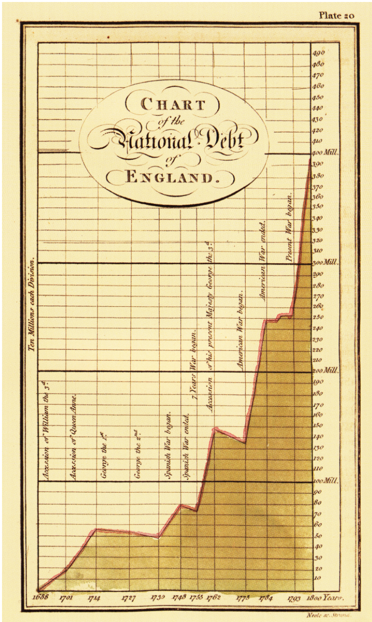

Figure 1: William Playfair's chart of the national debt of England

Static

The design of static visualisations is a well-documented subject, with good design practices being laid out by proponents such as Edward Tufte, John Tukey, Stephen Few and Naomi Robbins.[2] Much of this practice is also supported by academic research, including the early work of William Cleveland along with more recent developments.[3]

[2] Edward R. Tufte. *The Visual Display of Quantitative Information*. Graphics Press, 2001; John Tukey. *Exploratory Data Analysis*. Pearson, 1977; Stephen Few. *Information Dashboard Design*. Analytics Press, 2013; and Naomi B. Robbins. *Creating More Effective Graphs*. Wiley-Blackwell, 2005

[3] William S. Cleveland. *Visualizing Data*. Hobart Press, 1993; Jeffrey Heer and Michael Bostock. Crowdsourcing graphical perception: Using mechanical turk to assess visualization design. In *28th international conference on Human factors in computing systems*, pages 203 – 212, 2010; and Heike Hofmann, Lendie Follett, Mahbubul Majumder, and Dianne Cook. Graphical tests for power comparison of competing designs. *IEEE Transactions on Visualization and Computer Graphics*, 18(12), December 2012

Dynamic

Development of dynamic visualisations is a very active area, with the art-of-the-possible being continually extended through the development of new tools and libraries. Web-based technologies are popular in this area, due to the versatility of the HTML Document Object Model and the ubiquity of compliant web browsers. One of the most popular javascript libraries for visualisation on the web is D3 (standing for *Data-Driven Documents*), developed by Mike Bostock.[4] Outside of the web-browser the Computable Document Format (CDF) is being developed by Wolfram Research to provide a very extensive feature set for offline interaction.[5] There are also visionaries such as Bret Victor working towards improvements not only in visualisation output, but also in the process of visualisation creation.[6]

[4] `https://d3js.org`

[5] `http://www.wolfram.com/cdf`

[6] `http://worrydream.com/#!/DrawingDynamicVisualizationsTalk`

The current state-of-the-art for visualisation would therefore not only be built upon technologies that allow for dynamic content, with useful and coherent interactivity, but also following best design practices laid out by the experts referenced above. It is beyond the remit of this paper to go into further detail on visualisation specifics; however the reader is encouraged to review the advice of experts in this area and consider its application to their own work.

The remainder of this paper is split into three key areas:

1. The case that data visualisation should be a key skill for practicing systems engineers.
2. A demonstration of data visualisation within systems engineering in the rail industry.
3. A discussion of how data visualisation relates to model-based systems engineering (MBSE).

Importance to systems engineers

As a cross-cutting discipline, systems engineering is often concerned with interfaces. These can exist between systems, engineering disciplines and stakeholders. Effective management of interfaces requires that they be recognised, defined, communicated and agreed; effective communication is therefore an important skill for any practicing systems engineer. Often the task of ensuring that different disciplines effectively communicate and collaborate can be a full-time role; particularly where the integration of multiple systems and sub-systems forms a significant part of the project.

> "What is to be sought in designs for the display of information is the clear portrayal of complexity . . . to give visual access to the subtle and the difficult . . ."
>
> Edward R. Tufte. *The Visual Display of Quantitative Information*. Graphics Press, 2001

Communication can take many forms, including oral, written, and graphical; typically it is a combination of these. Within engineering, communication typically needs to be formal and precise, as it will form the basis of decision making and may be recorded and referenced at a later date. Engineering communication is also important as safety, budget and timescales can be at stake. Communication skills therefore typically form part of an engineer's formal education, often supplemented with additional training courses throughout their career. Graphical communication may be covered during this education and training, however often not explicitly and generally in an ad-hoc manner. This is unfortunate as graphical visualisation occurs frequently in engineering, for example:

work breakdown structures and Gantt charts are used in project planning; 2D and 3D design images are generated in engineering computer-aided design (CAD); flowcharts and systems models are produced for process control; bar, scatter and pie charts are used in reporting; and dashboards and control panels are designed based on human-factors. Within systems engineering there are many further views of a system that can be represented graphically, including: requirements traceability; functional/product/system breakdown structures; interface matrices; cause and effect diagrams; and sequence diagrams. Whilst in some cases the production process for these is straightforward, and may be automated within appropriate software, the engineer must still select the correct options, and determine what outputs are appropriate.

Engineering has successfully transitioned from a situation where documents were both produced and consumed directly on paper into a situation where production is almost entirely performed on computer. It is currently transitioning into a situation where document consumption is increasingly through a digital screen, be that on monitors, projectors or handheld devices. Unfortunately until this transition is complete, and we operate in the frequently prophesied *paperless office*, the outputs produced continue to be thought of as static documents. Retaining a focus on *printability* misses out on a lot of the rich interactive features that are possible within a digital medium. In fact, 3D and executable models really cannot be reduced to a 2D printed medium without a significant loss in information content.

This transition is not unique to engineering, and can be seen at various states of progress in sectors such as the media, education, science and medicine. Given the importance of effective communication within systems engineering, perhaps its practitioners should be leading the way in producing innovative high quality examples of data visualisation? Although it is hard to measure objectively, systems engineering may not be any more advanced than other engineering disciplines in this regard.

It is contended that:

1. Graphical/visual information presentation is an important mechanism for communicating within a systems engineering setting, ensuring that complex concepts are accessible to a wide audience; therefore these skills should be a key tool in a systems engineers skill-set.

2. Graphical/visual information presentation should be more widely taught and understood in engineering education and training.

Application to systems engineering within the rail industry

Data visualisation presents different challenges depending upon the discipline and domain in which it is applied. In this section we demonstrate the application of data visualisation techniques to systems engineering within the rail industry. Systems engineering practices are still relatively new to rail and infrastructure, and WSP | Parsons Brinckerhoff has been a major player in promoting their use within the industry, primarily through demonstrating the benefits that they can bring within programmes, along with making a commitment to exploiting digital technologies to enhance delivery across all its projects.[7] These practices are defined within a framework which provides a general systems process which can be followed to ensure project success.[8] This framework proposes the generation of several project artefacts, such as documents, drawings or diagrams, which aid communication and generate a common understanding between stakeholders. The generation of these artefacts has evolved into reusable semi-automated tools that allow faster and more efficient delivery.

Client feedback on these artefacts has led to many being produced as visualisations, as these are simpler to understand and faster to digest than written documents or tabulated data. These began as bespoke drawings and have since developed into more dynamic data-driven diagrams. Typically they are produced using Microsoft Visio, which allows for effective manual layout, whilst still providing stencils to speed up drawing and promote reuse. Visio Professional provides functionality to connect data stored in Excel or Access with diagram shapes; this shape data can then be used to determine attributes such as colour, position or text content. The *data links* are retained, such that the diagram can be automatically updated when the external data is changed. Through the use of sophisticated Visio functions, complex and interactive visualisations have been produced. These visualisations are tailored to the rail domain, and take into account several important aspects of it; not least, the fact that systems engineering practices are still not commonplace and therefore ease of understanding is paramount.

[7] Steven Turner and John Welford. System integration in a fragmented rail industry. `http://network.wsp-pb.com/article/system-integration-in-a-fragmented-rail-industry`, September 2016; and WSP | Parsons Brinckerhoff. Digital life, digital legacy. `http://www.wsp-pb.com/PageFilesn/55557/Digital%20Life%20Digital%20Legacy%20online.pdf`, 2014

[8] Malcolm Thomas, Paul Carter, and Alan Knott. SE for different industries: one size fits all? In *INCOSE UK Annual Systems Engineering Conference*, 2013

Stakeholder engagement

In industries such as aerospace and defence, where systems engineering techniques were developed and proven, stakeholders have generally been on the development journey alongside the systems engineers, and therefore the highly-evolved products of systems engineering in these areas tend to be well accepted; relatively complex notations such as the Systems Modelling Language (SysML) are understood and recognised. These types of highly-evolved notation are less well accepted in rail systems engineering, and therefore may not assist in stakeholder communication. Additionally there are some important aspects of rail and infrastructure systems engineering that require special treatment, as detailed below.

Extant systems

Due to rail projects in the UK almost always being an incremental upgrade of existing infrastructure, it is often necessary to model the current system before considering any additions or changes occurring within the project. These extant systems may have been originally designed decades or even centuries ago, but must interface directly into modern system upgrades. This presents major challenges to the project that must be reflected in the systems engineering approach taken.

Simple *rich-picture* approaches to physical systems modelling have been developed; these show the interfaces between various elements of the railway and highlight system upgrades and additions. This has helped clients such as London Underground and Network Rail achieve a clear understanding of how changes will occur as their programme of work progresses.

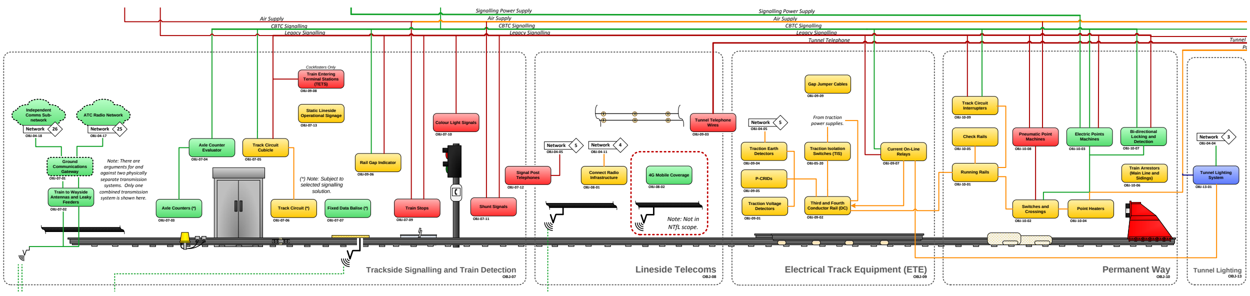

Figure 2: Part of a rich-picture showing physical London Underground systems, colour coding highlights sub-systems to be added or removed

Temporal considerations

Although the multi-year timescales seen in rail and infrastructure are not uncommon in other industries, rail perhaps differs in the phased introduction that is typically seen. In contrast to a product, that will progress through several prototypes and culminate in a final fully-tested design prior to being released to its end-users, the railway must go through multiple phases of development, whilst remaining operational throughout. This leads to an increased focus on design evolution for the rail systems engineer. Typically this has been handled through the use of *migration planning* within WSP | Parsons Brinckerhoff.

A tool has been developed to automatically generate a migration plan diagram, based on data on project completion dates and how they feed into the key configuration states of the programme. This allows the effect of delays or project issues on the overall programme plan to be quickly assessed. It also gives a convenient single-page plan showing how the various aspects of a project come together to form an upgraded railway system. This format has proven itself useful across a number of different rail programmes in the UK and internationally.

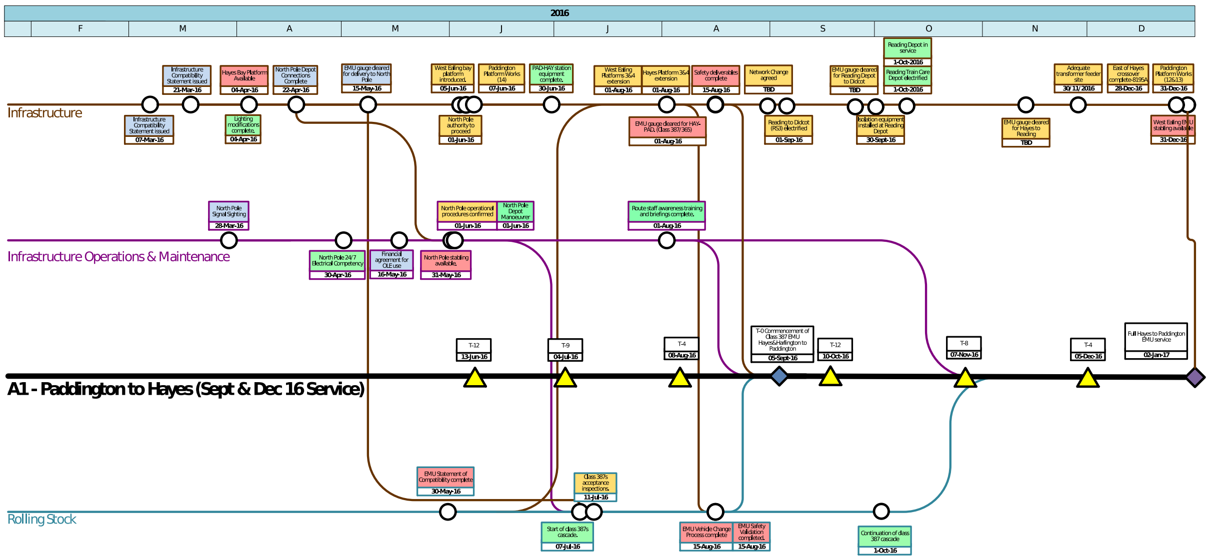

Figure 3: Migration plan showing the temporal inter-dependencies between projects on a section of the Great Western route

Geographic locations

Rail projects tend to be widely distributed geographically, but also require relatively detailed design in specific locations. This requires an approach that falls somewhere between classic model-based systems engineering and geographic information systems (GIS). By taking a data-driven approach to tool development we have been able to create semi-geographic rail network layouts that highlight various aspects of the system. For instance, through the use of colour coding we have been able to produce a visualisation of performance along the Great Western route, based on Network Rail recorded data. This has provided important insight into where system delays were occurring that was of considerable use to the Great Western Route Modernisation programme.

Similar approaches have been pursued to predict track loading based on timetable forecasts; to consider what platforms require lengthening to enable longer trains to follow a planned route; to show the status of line electrification across an area; and to show the location of projects and highlight their potential interfaces.

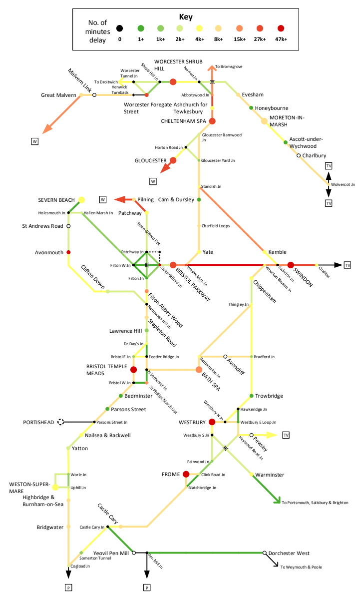

Figure 4: Historic cost of delays on the West-of-England section of the Great Western rail network, shown as a heat-map

Data-driven benefits

Taking a data-driven approach to producing diagrams has several advantages. Where the data exists in a suitable format it can substantially speed up the process of producing a visualisation. For some diagrams it can be as simple as clicking a button to completely generate it, for others manual layout will be required, however all the components can be automatically generated ready for positioning.

Alongside simplifying initial generation, taking a data-driven approach also makes document management and update much more straightforward. For example, the HS2 System Architecture Diagram (SAD) displays the geographic location of assets, such as access points, and features, such as tunnels and listed buildings, along a straightened representation of the track. This allows engineers to evaluate proposed changes to the railway layout and the impact they will have on other aspects of the design. The original version of the SAD was manually drawn, and required careful update and adjustment whenever any minor changes were made to asset positions, particularly if the same asset appeared on more than one page of the multi-page document. By rebuilding this diagram as an automatic drawing tool, the location of assets is now handled through an Excel spreadsheet and updates are a much more straightforward process.

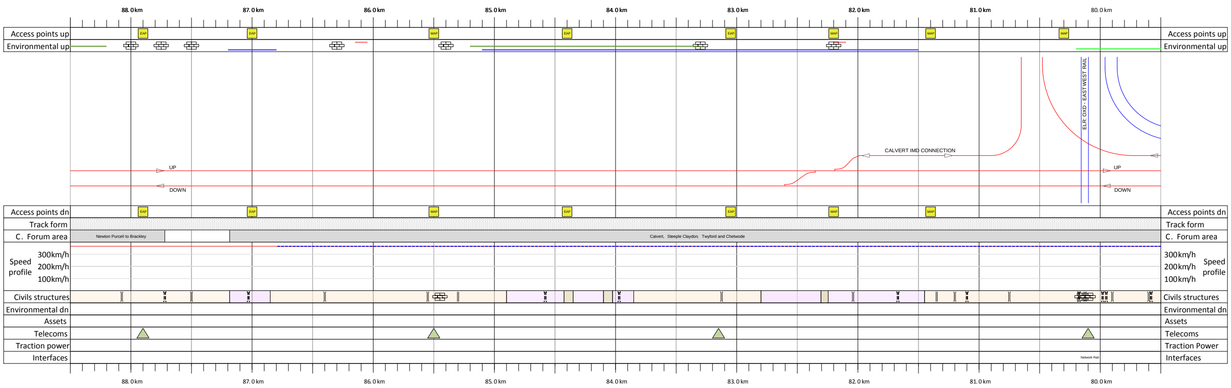

Figure 5: Icons automatically placed according to their linear position alongside the HS2 track

Interactive diagrams

It is often the case that there are multiple pieces of data about a diagram element, or that parts of the diagram are only relevant in specific circumstances. In a static environment this can be dealt by producing either multiple diagram versions, or a complicated diagram showing everything at once. This situation frequently arises where the system is changing over time, or where different stakeholders are interested in different aspects of the system. An elegant solution to this is to produce a single interactive diagram, that allows the user to select the particular view of the data that interests them. This approach has been used for the Great Western performance diagram, allowing the heat-map to be changed to show either: average time delayed; cost incurred; or incidents involved. The map can also be overlaid with manually identified hotspot locations that are known to cause performance issues.

Relationship to MBSE

The data-driven approach to developing systems engineering artefacts has delivered a set of reusable tools that enable various different views of the system to be produced. These have been well received by clients and are helping to progress the use of systems engineering within the rail industry. The tools themselves continue to be developed and extended; however initiatives are also underway within WSP | Parsons Brinckerhoff to take them one stage further.

Although data-driven, the tools currently have independent data sources. This data is also separate from that held in project management, document storage and building information modelling (BIM) software. Ongoing work is looking at pooling the background data into a centralised repository, whilst also allowing import/export from other software tools. This will provide a *single source of truth* for the project engineers and will effectively make this a full MBSE approach.

This is perhaps coming at MBSE from a different direction than some other organisations. In the context of Holt and Perry's MBSE ontology for architectures,[9] the visualisations being discussed here represent views, and the tools used to generate them define viewpoints; whilst an ontology effectively exists, embedded within each tool, there is not yet any higher-level architecture or architectural framework – in an explicit sense.

[9] Jon Holt and Simon Perry. *SysML for Systems Engineering – a model-based approach*. IET Publishing, 2013

By generating a view that end stakeholders find valuable, one of the criticisms sometimes levelled at SysML – that its end products are not of use – is avoided. Interestingly, current work by the Object Management Group (OMG) on the development of SysML version 2 separates model visualisation from model entry and editing, with the intention of addressing this issue.[10]

[10] Hedley Apperly. Sysml 2.0 update. `http://www.incosewiki.info/Model_Based_Systems_Engineering/Files/2/2c/PTC_SysML_2.0_Update_%28May_2016%29.pdf`, May 2016

Discussion

This paper has put forward the view that data visualisation should be an essential skill for practicing systems engineers, and that these skills should be formally included within engineering education and training. It has also presented examples of systems engineering visualisations that are proving effective within the rail industry. Finally it has described these visualisations within the context of MBSE.

In many cases an off-the-shelf solution may be adequate for the task at hand; the bespoke visualisations shown here are tailored to specific requirements within a specific industry. The production of bespoke visualisations is not trivial and therefore the time and costs involved in their production need to be traded against the project benefits that they bring, including the substantial risks that they mitigate around miscommunication and insufficient analysis.

Acknowledgments

The work shown in this paper was carried out by the WSP | Parsons Brinckerhoff Systems Development team. Particular credit goes to Joseph Hughes and Luke Fischer for development of several of the diagrams shown.