

**White Paper**

**Using Interactive  
Data Visualization to  
Promote an Active  
Learning Experience  
for Engineering  
Students**

# Using Interactive Data Visualization to Promote an Active Learning Experience for Engineering Students

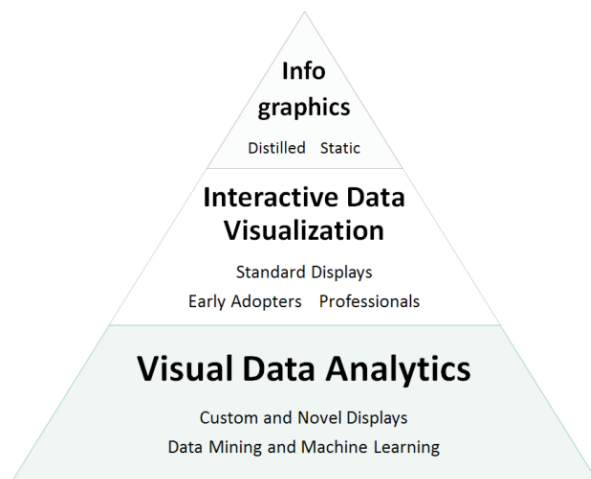
Anselm Spoerri, Rutgers University

---

## Introduction

Data visualization is enjoying a growing interest and adoption in diverse disciplines and business sectors. The exponential growth of data is fueling, in part, the adoption of visualization tools to make the “story” in the data visible. The increased capabilities of web technologies are turning web browsers into sophisticated viewports that make interactive visualizations accessible to diverse audiences. However, before the data can be visualized, it needs to be checked for accuracy and inconsistencies, which can be a time-consuming and complex process. The effectiveness of the resulting “data story” depends on two primary factors: whether there are actually meaningful patterns in the data and if a suitable visualization can be designed for the intended audience.

Methods of varying complexity and ability to deal with large amounts of data are available to visualize data. **Infographics** can be created, using pictograms and simplified graphics, to show major data patterns in a static way, where the designer of the infographic performs the work needed to distill the data. Infographics can be useful to show the meaningful “tip of the data iceberg” in a widely accessible way. **Interactive data visualizations** enable users, such as professionals, to explore more complex and multi-variate data sets to find meaningful patterns. It is interesting to see how major news organizations such as the *New York Times* are incorporating interactive visualizations into their news stories and introducing novel ways of visualizing data to a general audience. When data sets become too large and



complex, **Visual Data Analytics** methods need to be used. This involves the utilization of statistical analysis and machine learning algorithms which are employed to identify subsets or aggregations of the data that warrant further exploration using custom visual displays. These data mining methods, in addition the visual explorations by users, can be coupled to help fine tune the statistical analysis and visual presentation. Standard data displays, such as bar, line, pie, or bubble charts, encode the data in a way that is easy for the human visual system to interpret. While they are familiar to many people, these displays are limited in the number of data variables or data types that can be visualized. Thus, researchers are developing new types of data displays to visualize ever larger data sets or complex data relationships. However, these visualizations tend to have a learning curve and only if the insights to be gained are sufficiently valuable will users have the needed incentive to learn how to use these novel displays.

### Leveraging the Capabilities of the Human Visual System

The overall goal of data visualization is to guide users to interesting and meaningful patterns in a data set. The human visual system first detects “edges” in the visual scene that correspond to significant changes in terms of light intensity,

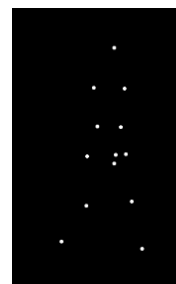
color, texture, and motion. Next, it uses grouping principles, such as proximity, similarity, and continuity, to group these edges into objects. The human visual system is able to infer the three-dimensional structure of the visual scene, although it has as its initial input a two-dimensional array of sensory measurements that are changing over time. It is able to achieve this feat by exploiting the structure of the physical world, in which object surfaces tend to change smoothly or objects move in mostly rigid ways, to compute what we “see” in our conscious mind.

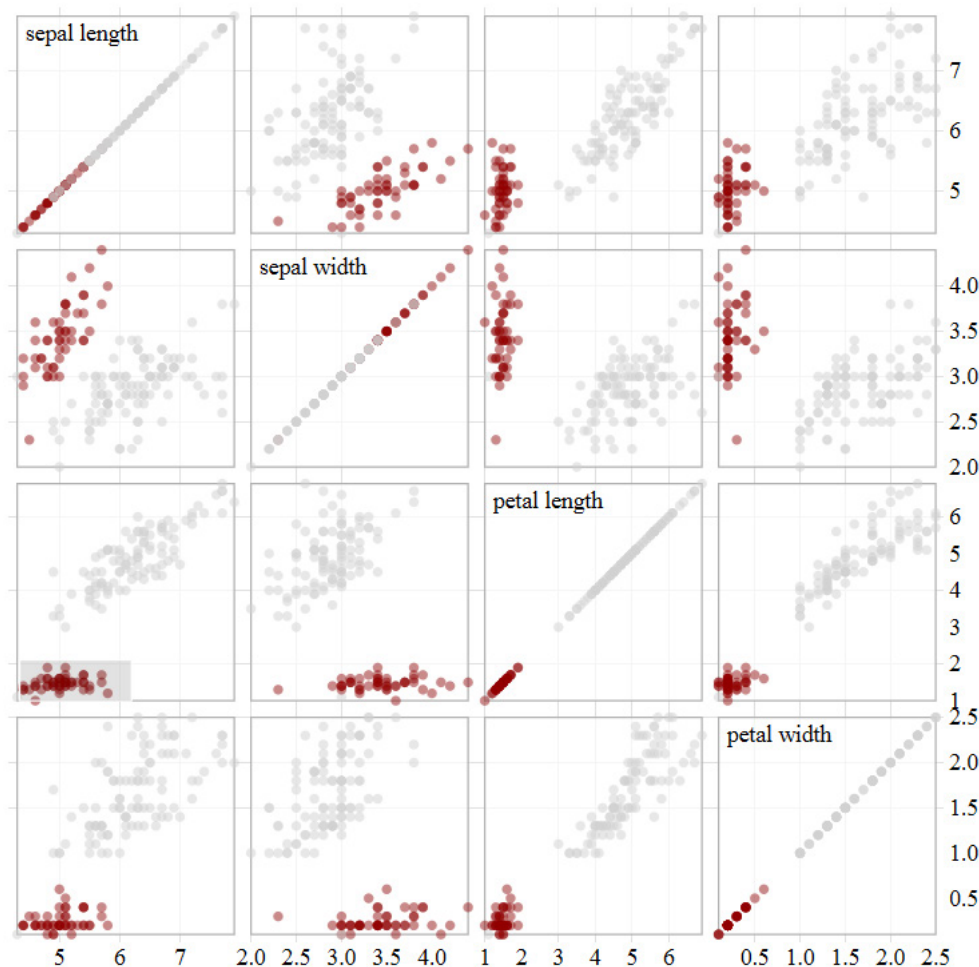
The “Walking Dots” animation illustrates the ability of the human visual system to compute “structure from motion” based on how a set of dots are moving, even when at times a “snow storm” of random dots is shown as well. The human visual system needs to solve the “correspondence problem” by determining where the dots have moved and it cannot just apply the proximity principle. The dots shown in the animation represent the joints (i.e. shoulder, elbow, knee, etc.) of a walking person. For example, upper and lower arms move like a pendulum and so the motion of the joints is constrained to a plane. This constraint and the fact that the human body moves in a semi-rigid way can be used to solve the correspondence problem mathematically. This example demonstrates the amazing computational capabilities of the human visual system to perceive the structure inherent in visual patterns.

When designing data visualizations, it is important to verify whether the resulting visual patterns draw the viewer’s attention to actual key patterns in the data. This becomes progressively more challenging as the number of data variables to be shown in the same display increases.

### Key Design Principles Used in Data Visualization

As the amount of data and the number of data variables to be compared increases, the following





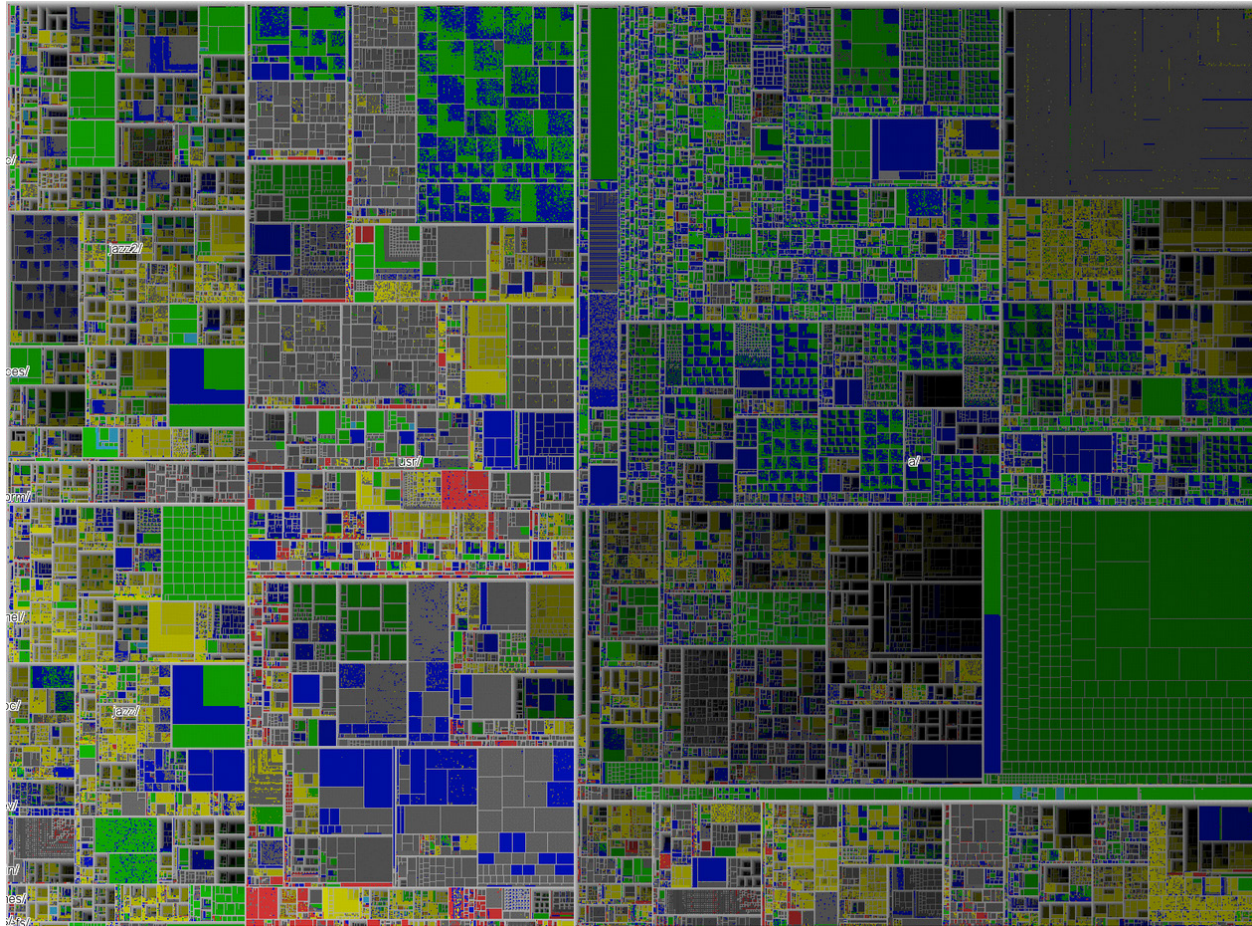
Scatterplot matrix (<http://mbostock.github.io/d3/talk/20111116/iris-splom.html>)

design principles can be employed to create effective data visualizations.

1. **Interactivity** enables users to explore large and multi-variate data sets by allowing them to easily dive into the data by interacting with sliders, checkboxes, and other ways to filter the data. For example, in the “[Delegate Calculator](#)” created by the *New York Times*, users can explore how changing the percentage of votes specific candidates receive would affect which candidate is able to reach the needed delegate count and when.
2. Going hand in hand with interactivity, a data display needs to provide **immediate feedback** to users when they interact with the data filters or the display. As more sophisticated data

visualizations are becoming accessible via web browsers, this forces developers to design applications that optimally use browser and server resources.

3. As the number of data variables to be explored increases, multiple displays need to be used and **linked** so that users can easily see how a selection in one display “plays out” in the other displays. For example, in a [scatterplot matrix](#), if users select a subset of spatially clustered points in one scatterplot, then the other linked scatterplots will immediately show whether or not these points have similar spatial patterns. Thus, linked displays make it easy for users to explore a multi-variate data space in a coordinated fashion.



Treemap of folder hierarchy of one million computer files (<http://www.cs.umd.edu/hcil/millionvis/>)

4. An effective visualization provides users with a compact **overview** of the data, which then can be explored using **zooming + filtering** interactions and provide **details-on-demand**. For example, this “**treemap**” visualization provides an overview of the complete hierarchical structure and types and sizes of over one million files on a computer hard disk. The size of the rectangles is proportional to the file size and the color related to the file type. Like a Russian nesting doll, the rectangles reflect the nesting structure of the folders, where the main folders have large white borders.
5. When users interact with display or filter controls, it helps them assimilate the resulting changes if the visualization supports **animated transitions**. In the *New York Times* visualization of the 2013 US federal budget, users can explore different ways to organize the individual budget items. When users select a display option, they can visually track how the budget items are moved into a new spatial arrangement, which reduces the cognitive work required to make sense of the new visual layout.
6. The overall goal of visualization is to show as much data as possible and **increase the information density** so that most or all of the available screen real estate is used to visualize the data. The treemap of over one million computer files is an example of a novel visualization developed by researchers that strives to maximize the information density by using a space-filling layout.

### Enhance Material Science and Engineering Education Using Data Visualization

Students initially learn about materials and their properties by consulting tables in handbooks or searching online, using resources such as [MatWeb](#). Students tend to get preoccupied with “looking up values” and do not get to explore why the major classes of materials have the ranges of values they do. What is missing for students is the ability to easily see material properties in context. More importantly, engineering is all about making trade-offs between different design requirements, and students need an active learning experience that helps them explore material property patterns and make informed decisions.

I have been part of a yearlong design and development process to help McGraw-Hill Education create [AccessEngineering’s DataVis](#), the interactive data visualization tool that lets faculty and students visualize property data across a wide

range of materials. The iterative design of the prototype and final tool was guided by interviews with educators and students to help us better understand their specific needs and which types of data displays would be most suitable. The temptation could have been great to create a “flashy” tool that uses novel visualizations, such as [parallel coordinates](#), but such displays are unfamiliar to most users and they tend to have a learning curve. Instead, we decided to use more standard and familiar visualizations – dot and scatter plots – where meaningful data patterns result in visual patterns that are easy for the human visual system to detect and interpret.

In a **dot plot**, as shown in Figure 1a, the x-axis shows the major material categories, such as metals, polymers, etc., as distinct columns of dots and assigns a unique color to each major category, using a color-blind safe color palette. The y-axis shows the values for the selected material

Figure 1: (a) Dot plot of density values for the major material categories; (b) the metal category expanded into its subcategories.

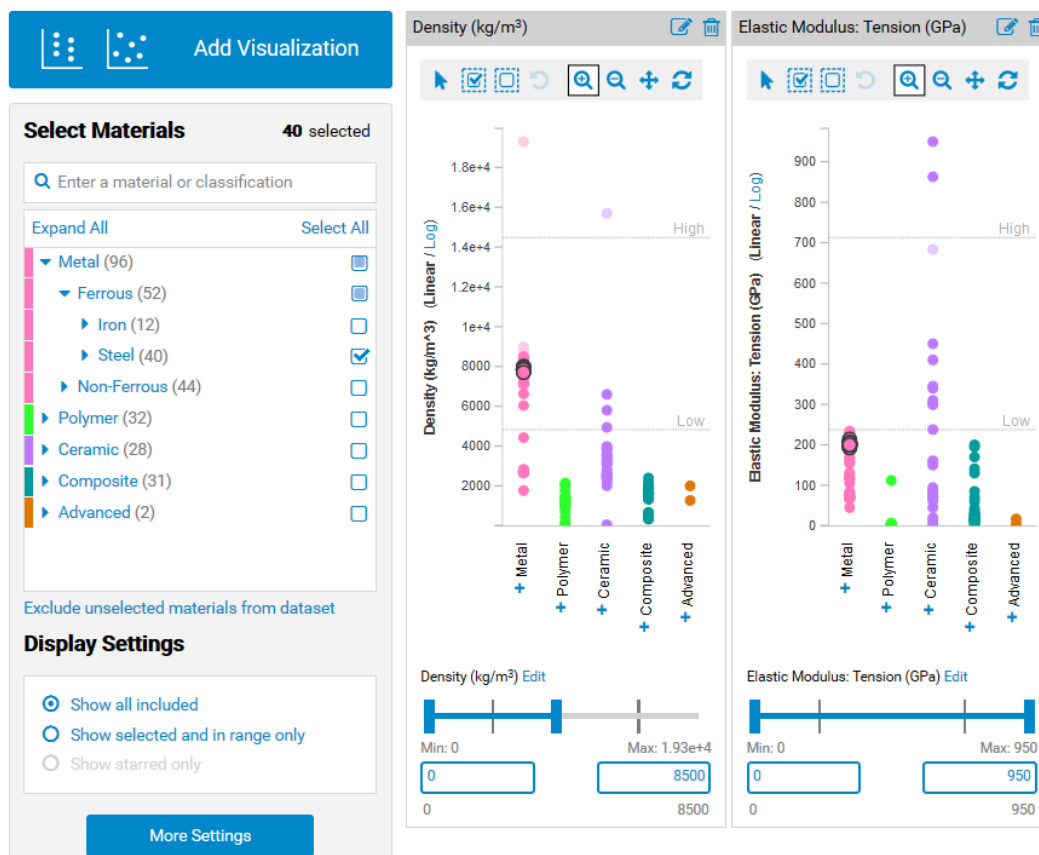


property (Figure 1 shows values for density), and users can easily spot outliers and compare the ranges of values for the different material categories. Figure 1a shows that metals have a great range of density values, where gold stands out with its very high density value, whereas polymers have a much narrower range of density values. Further, using an animated transition, a material category can be expanded into its subcategories by clicking on the blue plus signs on the x-axis. For example, Figure 1b shows that the density values for ferrous metals are spatially close and thus very similar, whereas a greater range of values exists for non-ferrous metals. In a **scatter plot**, as shown in Figure 4, users can explore the relationship

between two material properties, where the color coding helps users to see the differences in the correlation patterns between the major material categories. Both the dot and scatter plots provide an overview of all the material property data being compared on the same page. Users can easily zoom and pan inside the individual plots and a brief animated transition is used when users zoom or reset the display. Users can click on a dot to see details on demand about the material, such as its material classification and structure, as well as the source for material property value.

A key design goal was to make it easy for faculty and students to compare multiple material properties using linked dot or scatter plots in a

**Figure 2: Linked dot plots of density and elastic modulus: tension properties with only steels selected, and a range filter applied.**



step-by-step fashion; this allows users to easily add or change the material properties to be explored. Figure 2 shows how density and elastic modulus: tension can be compared side-by-side. Users can compare up to five material properties on the same page, as shown in Figure 3, and create and save projects that contain multiple pages. This makes

it possible for faculty to introduce students in a step-by-step way to key properties in materials science, starting with a single property on the first page and ending with a page that contains the key material properties that need to be considered, for example, when designing an aerospace structure, as shown in Figure 4.

**Figure 3: Four linked dot plots, where titanium is selected and starred, and steels and magnesium alloys are also selected. A range filter specifies that only materials with density less than 8500 are “in-range.”**

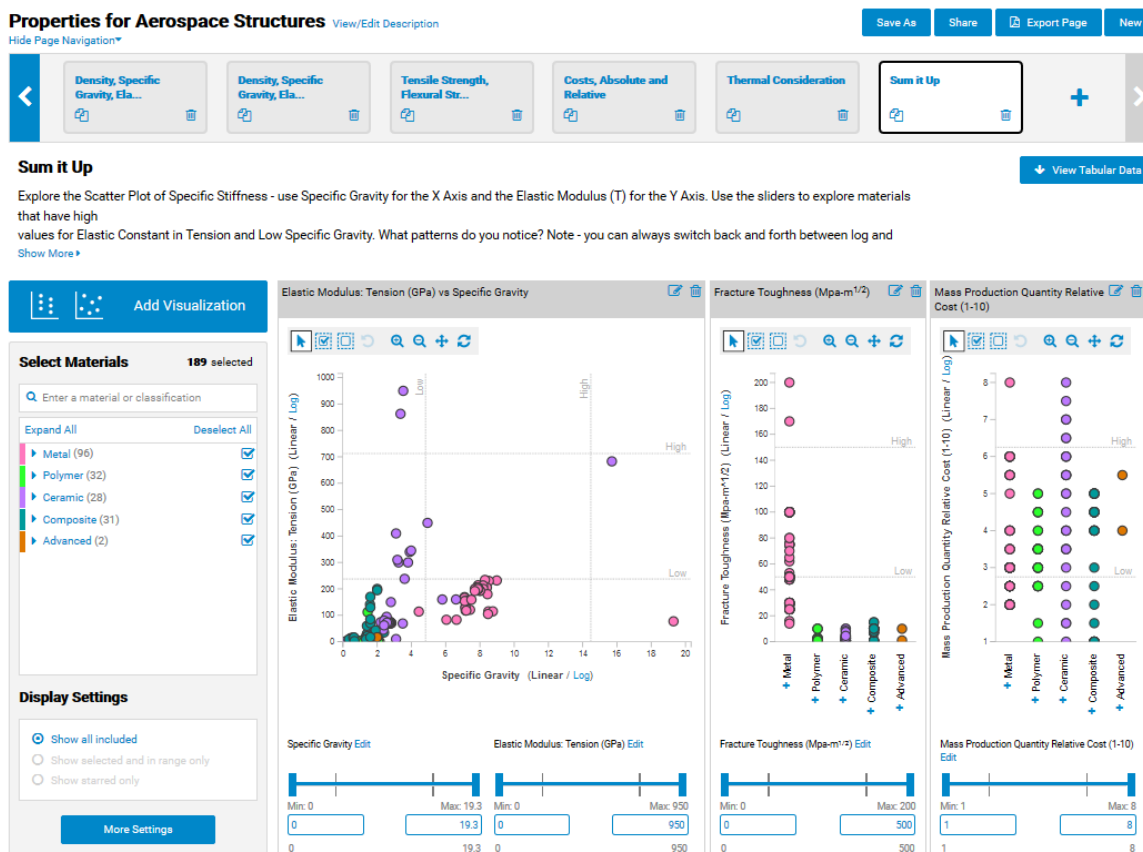


A curated and representative dataset of almost 200 materials and more than 60 properties was compiled by materials scientists, where the materials are hierarchically organized based on their material classification. Using a filter control, users can specify the selection status of materials based on their categories and subcategories; the dots representing selected materials are shown with a black border and unselected materials have no border. For example, if users select only steels, then the linked dot plots in Figure 2 show that steels have

similar density and elastic modulus: tension values, since the dots with black borders are tightly clustered in both displays. Further, users can specify the range of material values of interest by interacting with the range sliders—the dots that fall outside the specified range become dimmed. The visual coding used for controlling the appearance of the dots allows users to distinguish between materials that do—or do not—belong to the selected material (sub)categories, and see which materials are within the specified data ranges.



Figure 4: One page of a predesigned, customizable project, featuring a scatter plot and two dot plots.



In summary, AccessEngineering's DataVis allows faculty to create customized active learning projects that promote discovery and conceptual understanding of material properties through interactive data visualizations. As a result, students become engaged in an active and structured learning experience since they can interact with the data using hierarchical filters and range sliders to create in a step-by-step fashion their own visualizations that can be easily shared with and edited by others.

## Conclusion

The barriers that once made it difficult to enable students to explore rich data sets using sophisticated visualizations are disappearing. Representative data sources are increasingly

accessible online and powerful web-based visualization tools are becoming available to a general audience. Interactive data visualizations enable educators to present complex information in visual ways that can be easily interpreted, provided the appropriate displays are used to make the meaningful patterns in the data visible. Instructors can introduce students to key facts and concepts in an engaging way without having to sacrifice the richness of these concepts, and do so at an earlier stage of their education. Interactive visualizations enable students to learn to look for patterns, ask questions, and investigate the data in a purposeful way that promotes conceptual understanding and can make learning more enjoyable. The increased availability of modern data visualization tools has the potential to transform education in major ways.

### About the Author

Anselm Spoerri has a PhD from MIT and is a faculty member at the School of Communication and Information at Rutgers University. He has conducted research in the field of information visualization for the last 20 years and teaches students how best to visualize data. A recent co-authored paper about how

to visualize what is most controversial in Wikipedia received widespread media attention, including articles in *Wired*, *MIT Technology Review*, *The Washington Post*, and *The Economist*. Dr. Spoerri leveraged his data visualization expertise to guide the visual display and interactive functionality that are key to AccessEngineering's DataVis.

DataVis is built into AccessEngineering—your all-in-one solution for professional engineering content integrated with dynamic teaching and learning tools.

For information regarding a free trial of AccessEngineering, pricing, or placing an order, please contact us at:

[mheducation.co.uk/accessengineering](http://mheducation.co.uk/accessengineering)

