

# insurance day

www.insuranceday.com



TECHNOLOGY FOCUS



## A wealth of information

Our guest columnist today is, among many other things, a philosopher who used to write gags for a syndicated comic strip.

David Weinberger is hard to pigeonhole and so, as he explains, are the vast amounts of information the world is now generating. It is not just the data that prove hard to manage and understand. We are also producing new kinds of knowledge by finding patterns amid the growing data supplies.

How can we make sure this wealth of information produces real wealth? Our businesses and communities cannot profit from the networked age unless the data is somehow networked too.

Weinberger introduces the concept of linked data – a simple, elegant approach to connecting data that promises to build more value into published data.

Linked data can also serve to point data users to standard definitions, thereby anchoring networks of data in agreed meanings.

Weinberger uses Wikipedia pages as an example of anchor points, but community standard models such as the ACORD framework will play the same role.

It is also interesting linked data has the concept of relationships at its heart. Each “triple” demonstrates a typical relationship between data items and so reflects a real-world relationship between items of interest to business.

The classic relationships we see in insurance information, such as customer-owns-policy, translate easily to this new idiom. Again, such relationships abound in the ACORD framework.

The more meaning we can build into our growing information world, the more value we can all derive from it.

Gregory Maciag is president and chief executive of ACORD and can be reached at gmaciag@acord.org



# Data becomes a network

The internet age has transformed the way we collate and use data



David Weinberger  
Technologist and author

Data used to be points. Now it is a network. In fact, just about everything we used to think about data in the information age has reversed in the internet age.

This is not the first time the concept has undergone a transformation. The term “data” comes from the Latin for “the given.”

For example, in maths, data were the terms an operation starts with. In 1946, as information science was beginning to take shape, data became the stuff the early computers were given – typically the punchcards with holes (or lack of holes) that stood for the smallest units of information.

Computer data became a resource, something that could be stored and re-used, very unlike data in the old mathematical sense.

And we began to worry we had too much of it.

### Swamped by data

Worse, businesses felt they were not getting enough value from the huge quantities of data they were stockpiling. So we began to think about information as a refinement of the data “ore” that found valuable relationships within it.

Still, data remained as it had been: a single element, crisply defined, separated from all other data by clear boundaries.

The internet age is changing that. The first change is in quantity. For all our 1950s concerns about being swamped by data, at the same time we strictly limited how much data we recorded about each employee or transaction because our machines could not handle more.

Now, of course, you have more data in your pocket than entire companies maintained.

And when we talk about big data, we are overloading information overload far beyond what anyone ever imagined.

### A more holistic picture

The excitement about big data is due to the correlations we find that would not otherwise emerge.

For example, a 2007 study reported in *Time* magazine discovered if you become obese, your friends are 57% more likely to become obese. Likewise, a recent study of lots of data reported in *The Los Angeles Times* found HDL (“good cholesterol”) may have little to do with heart health. We may not know the causes but at least we see the correlations.

And when data sets can be integrated, we can get a more holistic picture of a world we rather arbitrarily fragment into discrete research areas. We know diet can affect cancer rates, but how about weather patterns? Is it conceivable “fracking” might affect educational achievement? Indeed, entirely new subject areas can now be studied, such as formerly unpredictable social behaviours.

### Linked data

Even more exciting possibilities are being opened by the change in data’s nature itself, as it moves from discrete to networked. The clearest example is the rise of the linked data standard advanced by Sir Tim Berners-Lee, the creator of the world wide web.

Linked data (an enabler of the semantic web also championed by Berners-Lee) provides a standard way to make online data computable by computers.

There are two components of the linked data standard. First, it says

you should make your data available in the form of “triples”: X has a relation to Y. Here’s an example: [1] Platypuses [2] inhabit [3] Tasmania. This makes it easy for a computer to connect the dots if it comes across a triple that says [1] Platypuses [2] have [3] venomous claws. Now the computer knows that there’s a creature with venomous claws in Tasmania. What were atomic facts now begin to form networks of knowledge.

### Cut through the differences

The second feature of linked data is especially brilliant. It says whenever possible, express each element of a triple as a hyperlink (a URL) pointing to a public source. If, instead of using strings of letters that differ (for example, some triples might refer to a platypus by its scientific name), the triple uses URLs that point to, say, the same page about platypuses in the online Encyclopedia of Life (or Wikipedia, or some other source), the computer can cut through the linguistic differences to see the triples are in fact talking about the same thing.

Big data when expressed as linked data becomes far more usable because the data have literally become links.

Because of that, data clouds become data networks. And our online ecosystem becomes much smarter... especially if we make all this data openly available to anyone with an idea to pursue. Networked data want to be free. ■

David Weinberger is a technologist and commentator best known as a co-author of *The Cluetrain Manifesto*. His latest book is *Too Big to Know: Rethinking Knowledge Now That the Facts Aren’t the Facts, Experts Are Everywhere and the Smartest Person in the Room Is the Room*.



**For all our 1950s concerns about being swamped by data, at the same time we strictly limited how much data we recorded about each employee or transaction because our machines could not handle more. Now, of course, you have more data in your pocket than entire companies maintained.**