

Get 4 FREE trial issues &gt;&gt;

Subscribe

Student offers

The  
Economist

Log in | Register | My account

Newsletters | RSS | Subscribe | Classifieds

Friday July 16th 2010

Search



Home

World ▾

Business &amp; Finance ▾

Science &amp; Technology

Economics ▾

Culture

Site Index

Print Edition



Technology

Babbage

## The Difference Engine: The wisdom of crowds

Jul 15th 2010, 13:56 by N.V. | LOS ANGELES

EVER noticed how opinionated the world has become? It's not just all those product reviews on websites like Amazon, Expedia or eBay. From Facebook to Twitter, personal opinions litter the blogosphere—sometimes in narratives apropos of nothing in particular; other times embedded in comments on the news; and frequently, sad to say, as flagrant plugs by shills masquerading as innocent bystanders.



Once we sought advice from friends and family about what, where and when to buy something. We read newspapers and magazines to form opinions about current events, movies to watch, new books worth reading, who or what to vote for, and why. We went to libraries to check out how to solve problems that concerned us. In business, we relied on surveys, focus groups and consultants.

That's all begun to change. Thanks to the internet, we are now inundated with advice from millions upon millions of opinionated folk we've never met—and frankly haven't a clue how to assess. The fashionable term for trying to glean useful insights from it all is crowd sourcing. But read any selection of blogs (there are over 112m of them in the English language alone) and you quickly learn that meaningful information is in short supply.

Ironically, that may not matter much. As a fledgling investment banker, your correspondent learned many years ago that, as far as markets were concerned, emotions trump facts any day. It isn't events that move markets, but reactions to them, so long as they are shared by a big enough bunch of traders. "What's the sentiment on the street?" is the cry heard on trading floors around the world. Sure, in their Mr Spock mode traders scour the horizon for hard-nosed data, read their analysts' reports and digest breaking news. But then Captain Kirk kicks in and they interpret these through a prism of personal experience, with its predilections, prejudices, doubts, and fears.

That is equally true of politics and public opinion. At Oxford University, Sandra González-Bailón and colleagues at Barcelona Media Innovation Centre have been using the emotional content of online discussions to predict how American presidents fare at election time. The technique provides an alternative to approval ratings, which gauge support based on a wide range of issues over the short-term, or opinion surveys, which collect responses to a narrow, pre-selected set of issues over the long-term. The attraction of the emotion-based approach is that it hones in on issues that people actually find important and want to discuss, rather than on topics predetermined by pollsters. It also offers clues to the psychological mechanisms that lie behind shifts in public mood—as happened most noticeably in America after the terrorist attacks of September 11th, 2001.

But it is in the economic sphere that emotionally charged opinions matter most. They make the world go round by informing our purchasing decisions about houses, motor cars, mobile phones and many other bits of merchandise. So, we need to take opinions—whether level-headed or misguided—very seriously indeed. Above all, we need to find

Comment (3)

Recommend (16)

E-mail

Share

Print

Permalink

### About Babbage

In this blog, our correspondents tinker with data, machines and their own lives so as to understand technology and its consequences.

[Follow Babbage on Twitter »](#)

[RSS feed](#)

Advertisement

### Economist blogs

[Babbage](#)[Bagehot's notebook](#)[Banyan's notebook](#)[Baobab](#)[Buttonwood's notebook](#)[Charlemagne's notebook](#)[Democracy in America](#)[Eastern approaches](#)[Free exchange](#)[Gulliver](#)[Johnson](#)[Lexington's notebook](#)[Newsbook](#)[Schumpeter's notebook](#)

### Products & events

#### Stay informed today and every day

Subscribe to *The Economist's* free e-mail newsletters and alerts.

[Get e-mail newsletters](#)

better ways of tapping the wisdom of the jabbering online masses while dispensing with the drivel.

Your correspondent isn't the first to ponder how to exploit the wonders of word-of-mouth. Social scientists have been asking themselves similar questions for years. More recently, academics in natural-language processing have embraced the topic. Now, entrepreneurs are getting in on the act. Over the past few years, 60-odd companies have set up shop to develop tools for clients needing a better grip on what, deep down, their customers or constituents really think.

Most of the work to date has used semantic search engines to parse text retrieved from the web for meaning, disambiguating words with similar spellings by taking their context into account. So far, however, such natural-language processing has favoured narrow fields like medicine or law where the terminology is limited. The computational burden would be too much if used with conventional search engines like Google or Bing, which continuously index the entire web rather than merely a slice of it.

A better understanding of what's actually being said on the web has come from an approach called "deep content analysis". This goes way beyond the realm of simple semantic search, allowing computers to understand the complete and unambiguous meaning of sentences. Still, it doesn't help distinguish the relatively clear, objective statements of fact from the invariably subjective and shifting opinions that give voice to a person's inner feelings, sentiments and attitudes to various things.

It is precisely the difficulty of extracting this emotionally charged content from the detached, hard-boiled sort that makes sentiment analysis such a tough nut to crack. Often, the relevant sentences or clauses are buried in long forum posts, blogs, or open-ended replies in stacks of questionnaires. Just finding them can take armies of analysts equipped with marker pens and printouts weeks on end. And then one has to decide whether the sentiment concerned is positive, neutral or negative, assigning some numerical ranking to it (say, +5, 0 or -2), so the overall results can be digitised and processed as raw data.

The problem doesn't end there. In mathematical terms, an opinion is what Bing Liu, a computer scientist at the University of Illinois in Chicago, calls a "quintuple" or five-variable expression. The variables include the object being evaluated, its various features, the strength of the sentiment (in numerical terms), the person who expressed it, and when. The analysts' job is to mine the text for all five pieces of information in order to identify distinct quintuples. Next, any pseudo-sentiments (spam) must be filtered out. Only then can the data be processed—and an averaged-out quintuple generated.

All of which sounds like a lot of hard work. No surprise, then, that so many start-ups have rushed to fill the need. Providing smart software that takes the grunt work out of mining text for opinions has helped Clarabridge of Reston, Virginia, grow at over 50% annually for the past few years. The company's automated sentiment tools are used by AOL, Marriott, Nissan, Wal-Mart, Wendy's, United Airlines and a dozen other Fortune 1,000 firms.

Clarabridge's software lets firms process all the customer feedback that normally goes to waste (typically 80%) because it's trapped in some unstructured form. Equally important, such software allows sentiment analysis, which would take weeks to do manually, to be carried out in real-time—and on an 11-point scale instead of the basic three-value sort (positive, neutral or negative). This gives firms a deeper understanding of their customers' needs, and helps them respond more rapidly to changes in demand.

Lately, your correspondent has seen a sentiment engine based on ideas derived from decoding the human genome that spits out real-time opinions about the stockmarket's behaviour almost as quickly as the index can react. He wouldn't be at all surprised if in a year or two such an opinion-harvester were bundled with a program-trading system to create a money-spinning killer app. If only he had got his hands on one before bidding farewell to the trading floor to become an impoverished inky-fingered wretch instead. An opinionated one, mind you.

Subscribe to *The Economist's* latest article postings on Twitter

[Follow \*The Economist\* on Twitter](#)

See a selection of *The Economist's* articles, events, topical videos and debates on Facebook.

[Follow \*The Economist\* on Facebook](#)

Advertisement

**We are making continuous improvements to *The Economist* website and are interested in your thoughts.**

[Please leave your feedback.](#)