Google's head of research still searching for answers

KARLIN LILLINGTON

Peter Norvig has been attempting to explain the unreasonable effectiveness of data

It's strangely satisfying to find that senior Google people generally excel in the area of odd adjunct career feats. The internet giant's director of research, Peter Norvig, is no exception and can cite two.

First, there's his Gettysburg Address PowerPoint presentation, a hilarious, hugely popular rendering of Abraham Lincoln's eloquent speech into six PowerPoint slides. As a comment on the deflating effect slides and bullet points can have on the art of modern rhetoric, nothing surpasses it.

Second, he lays claim to having produced the world's longest palindrome sentence, a mind-boggling 17,826 words worked at Nasa as head scientist and done Cork.

Peter the Great: Norvig's Gettysburg Address PowerPoint presentation is a hugely popular rendering of Abraham Lincoln's eloquent speech into six slides. He also lays claim to having produced the world's longest palindrome sentence, a mind-boggling 17,826 words.

And that is how Google Translate works - by breaking a document into phrases, searching a vast database of millions of similar or identical phrases to find what one means, then going to the next phrase.

Image processing is another area that can use the same approach, says Norvig. Rather than teaching a computer what a single table looks like from every possible angle to understand the concept of a table, you offer a computer millions of images of tables, says Norvig.

Norvig notes that one person searching for the word "jaguar" may be interested in big cats, while another may be a car fanatic. Finding ways to guess the context of what its search engine users want is a major challenge and Google thinks personalised search will help.

"You will get things that you want, more of the time," says Norvig.

"But everyone agrees - some critics feel such an approach limits results and can give a too-personal view of where sites stand in Google rankings. Using someone else's computer might also skew returns. But overall Google feels this tweak will help individual searchers get a better search experience.

"Google has to continually re-evaluate and rethink what search means and take account of new technology developments, he says. Among the latest is the possibility of incorporating increasingly detailed location-based searches, especially on web browsers on mobile handsets that can pinpoint a user's exact locale. Location-based search, however, introduces more of the privacy questions that dog a company with a business model based around using and analysing user data.

But Norvig feels users will find some trade-off of information for useful services will be acceptable, and that users themselves will decide what they feel comfortable with.

Another challenge for the company is incorporating the increasing flow of information coming from the various elements of the "social web" - weblogs, Twitter, videos, profile sites and sources such as "content farms" like Ask.com that aggregate reams of content, often of mediocre quality.

Some Google users have advocated separating out blogs and other social web results from traditional webpages. But Norvig thinks Google has the best chance of providing a good array of responses to a search by offering slices of social networks as well. "Our feeling is we'd rather have you go to one spot," he says. "It's too hard to know what you want."

Many also find they didn't know they would want, say, video results until they were presented in the search response, only to find a video provides the information they want, he says.

And he's reassuring on one point. Even if the company is trying to give every user the most relevant set of search results possible.

"But everyone agrees - some critics feel such an approach limits results and can give a too-personal view of where sites stand in Google rankings. Using someone else's computer might also skew returns. But overall Google feels this tweak will help individual searchers get a better search experience.

"Google has to continually re-evaluate and rethink what search means and take account of new technology developments, he says. Among the latest is the possibility of incorporating increasingly detailed location-based searches, especially on web browsers on mobile handsets that can pinpoint a user's exact locale. Location-based search, however, introduces more of the privacy questions that dog a company with a business model based around using and analysing user data.

But Norvig feels users will find some trade-off of information for useful services will be acceptable, and that users themselves will decide what they feel comfortable with.

Another challenge for the company is incorporating the increasing flow of information coming from the various elements of the "social web" - weblogs, Twitter, videos, profile sites and sources such as "content farms" like Ask.com that aggregate reams of content, often of mediocre quality.

Some Google users have advocated separating out blogs and other social web results from traditional webpages. But Norvig thinks Google has the best chance of providing a good array of responses to a search by offering slices of social networks as well. "Our feeling is we'd rather have you go to one spot," he says. "It's too hard to know what you want."

Many also find they didn't know they would want, say, video results until they were presented in the search response, only to find a video provides the information they want, he says.

And he's reassuring on one point. Even if the company is trying to give every user the most relevant set of search results possible, the sheer breadth of what's out there on the web means serendipity will still be a key element of search.

"All the curious stuff will still be there," he says.

Gettysburg presentation: norvig.com/Gettysburg/Palindromes: norvig.com/palindrome.html