

Visualizing Information Quality and Perspective on the Web

Lesia Tkacz
VU Amsterdam
dx9240@gmail.com

Robin Sharma
VU Amsterdam
robinkeralaam@gmail.com

Davide Ceolin
CWI
davide.ceolin@gmail.com

Chantal van Son
VU Amsterdam
c.m.van.son@vu.nl

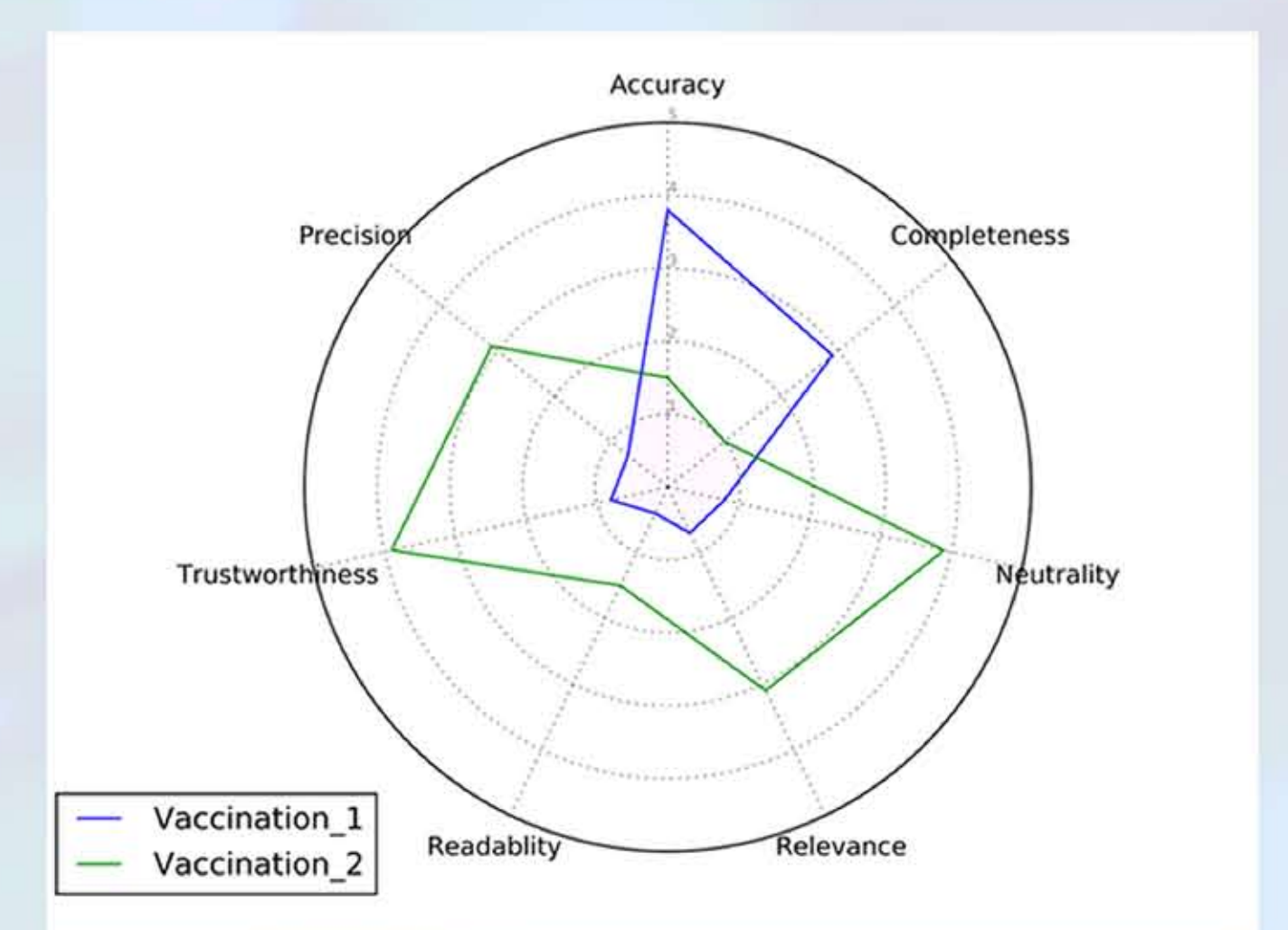
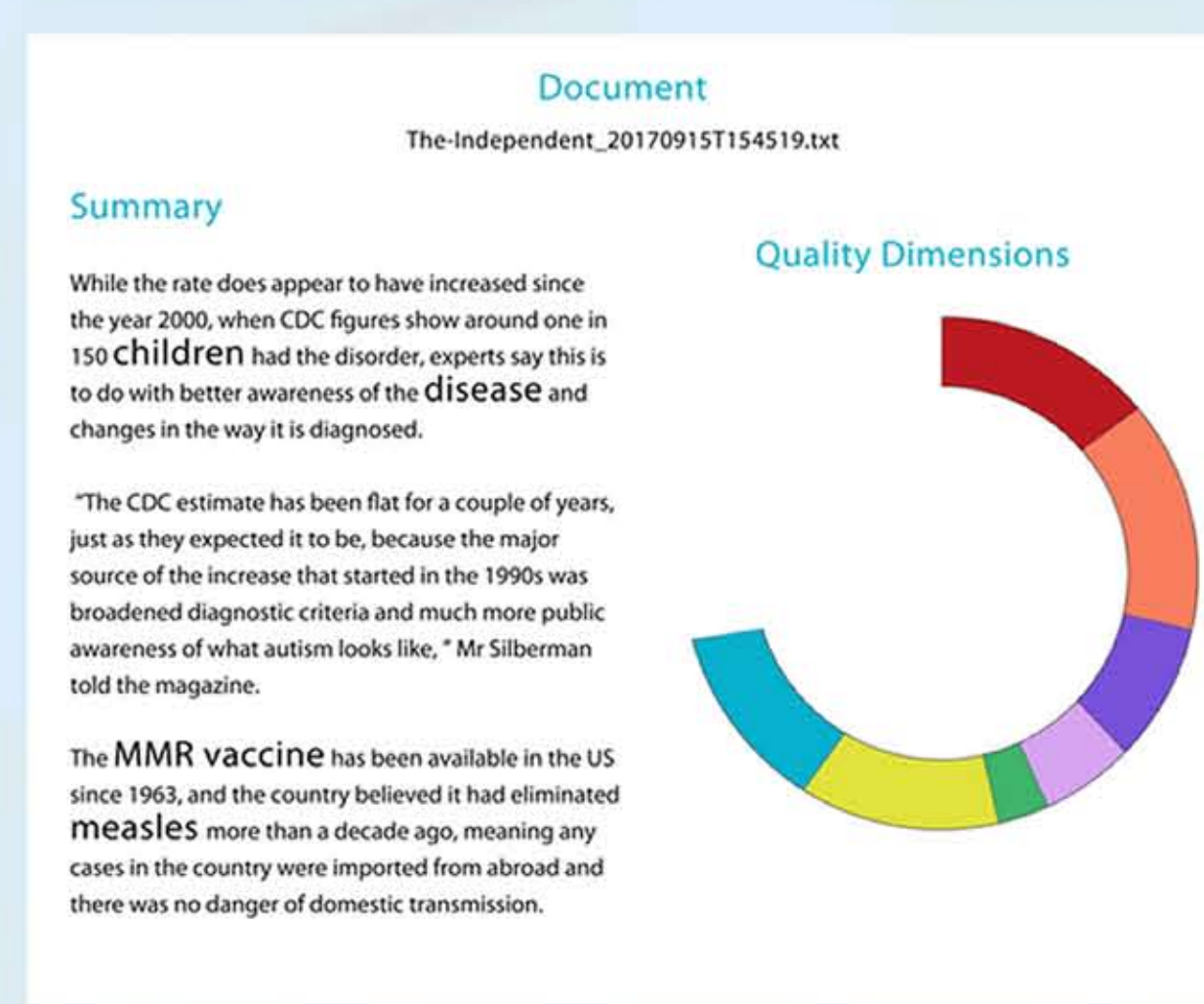
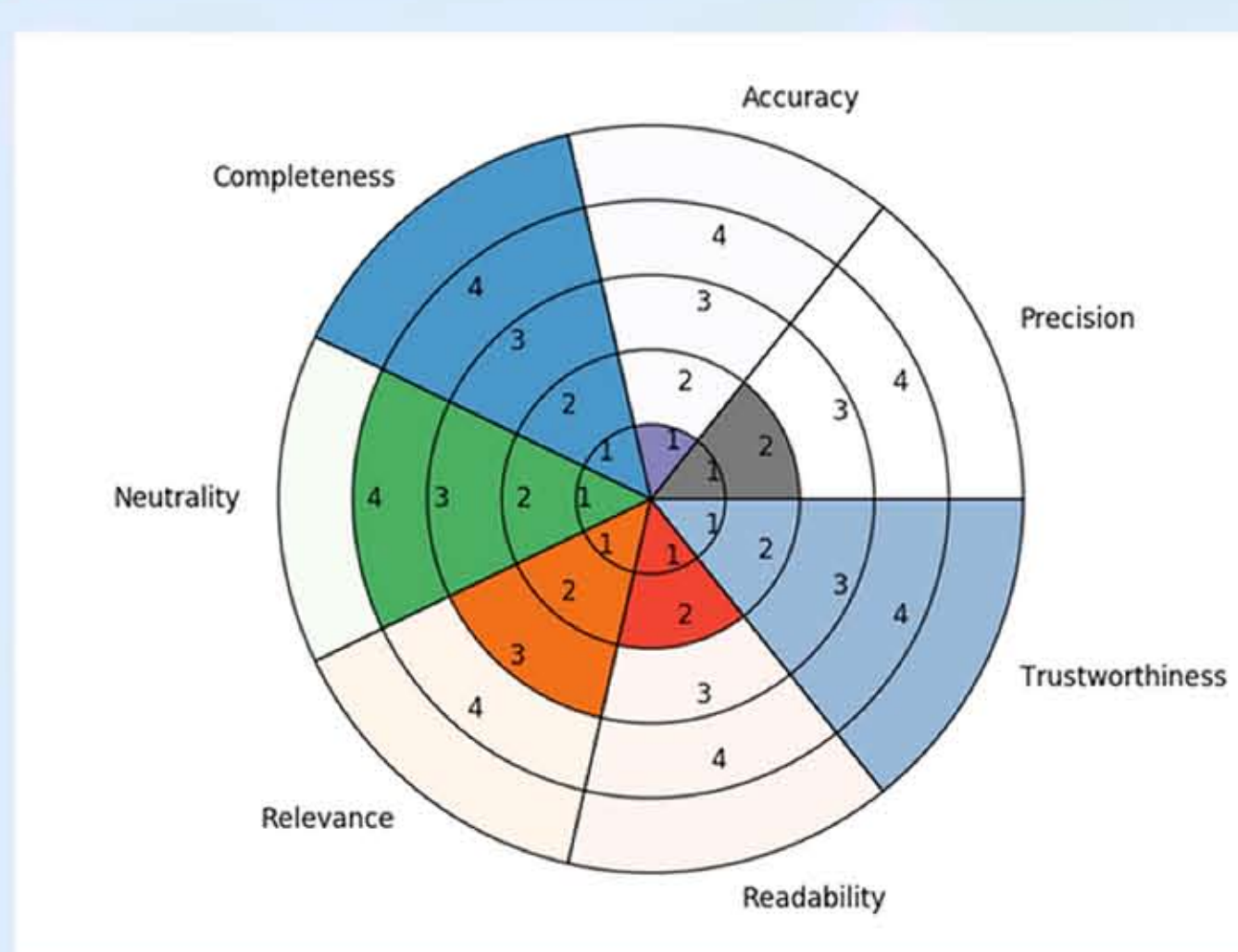
Problem: Information Overload & Information Bubbles

Web users are typically faced with an 'information overload' problem when they research controversial topics on the web, and the perspectives which can be expressed about the topics. This problem is amplified by the glut of web documents which can be retrieved, and is further complicated by both the varying quality of information across web documents, and by the absence of a variety of perspectives in information bubbles.

Solution: Visualization Design

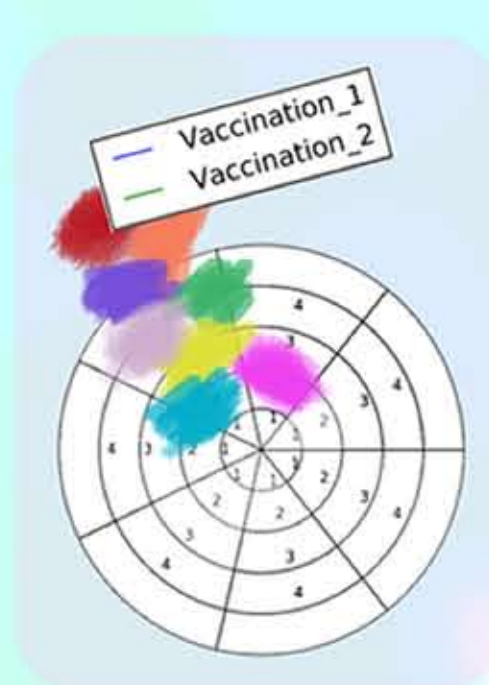
- We develop a web browser visualization tool aiding web users to navigate and informedly consume large collections of web documents
- The output of a quality assessment tool*, which scores a web document based on eight quality dimensions, is modelled

OverallQuality 3.5
Accuracy 3.7
Completeness 2.8
Neutrality 3.8
Relevance 3.2
Readability 1.6
Trustworthiness 3.8
Precision 3.2



An iterative design process is used to approach challenges and learn how to indicate documents' quality and perspective towards a topic:

- compare multiple documents
- represent a text's perspective and subtopics
- simultaneously combine and show several information quality scores
- maximize readability but avoid suggesting of false relationships between data



While the rate does appear to have increased since the year 2000, when CDC figures show around one in 150 children had the disorder, experts say this is to do with better awareness of the disease and changes in the way it is diagnosed.

"The CDC estimate has been flat for a couple of years, just as they expected it to be, because the major source of the increase that started in the 1990s was broadened diagnostic criteria and much more public awareness of what autism looks like," Mr Silberman told the magazine.

The MMR vaccine has been available in the US since 1963, and the country believed it had eliminated measles more than a decade ago, meaning any cases in the country were imported from abroad and there was no danger of domestic transmission.

- Natural Language Processing methods extract content from documents (text summarization, TF-IDF, word vector similarity measures)
- These methods point to expressed perspectives and thus contextualize the document quality dimensions
- The fourth design finally enables users to benefit from a fast but informative overview of web documents

Result: Web Document Quality Visualization & Perspective Tool

Document
The-Independent_20170915T154519.txt

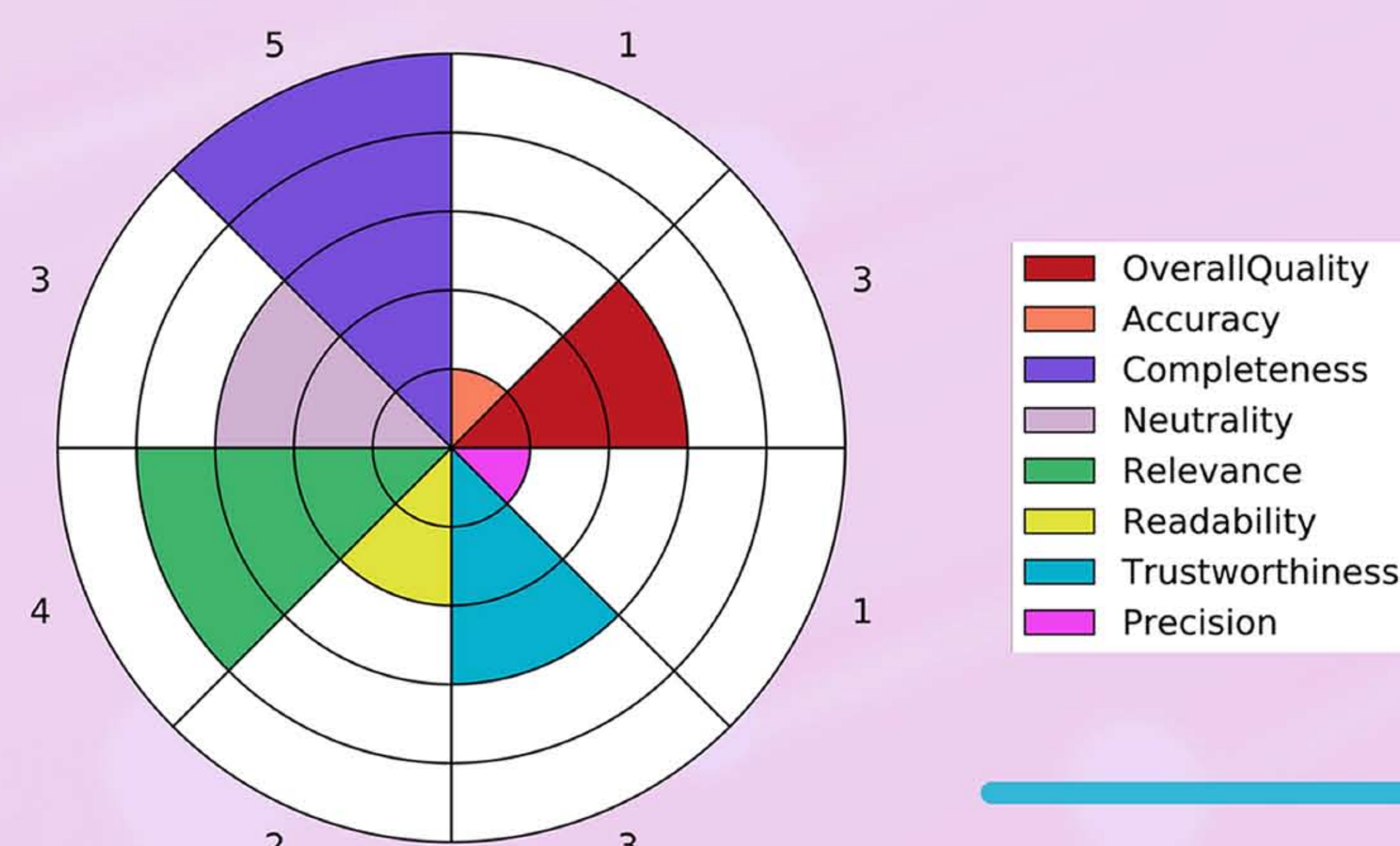
Summary

While the rate does appear to have increased since the year 2000, when CDC figures show around one in 150 **children** had the disorder, experts say this is to do with better awareness of the **disease** and changes in the way it is diagnosed.

"The CDC estimate has been flat for a couple of years, just as they expected it to be, because the major source of the increase that started in the 1990s was broadened diagnostic criteria and much more public awareness of what autism looks like," Mr Silberman told the magazine.

The **MMR vaccine** has been available in the US since 1963, and the country believed it had eliminated **measles** more than a decade ago, meaning any cases in the country were imported from abroad and there was no danger of domestic transmission.

Quality Dimensions



Future designs will contribute to sharper representations of perspective, by using a Neural Network model to evaluate documents for general agreement or contradiction with each other, based on Textual Entailment. This allows document grouping according to their perspective on a topic or subtopics, and their information quality.