



# User-oriented Keyword Visualizer

SANGHO SUH  
MOONSHIK KANG  
HOONHEE CHO

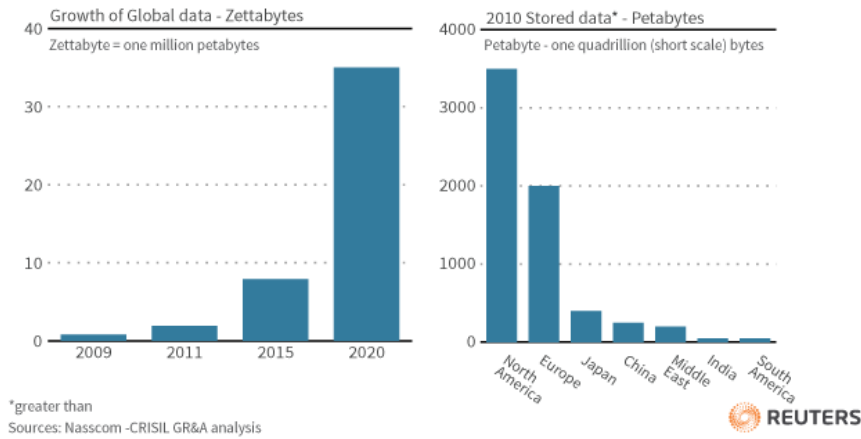
10/11/2015

# Introduction

## ► Problem

### Big data growth

Big data market is estimated to grow 45% annually to reach \$25 billion by 2015



**Find Keywords Analysis Tool**

This tool helps to analyse text in order to find keywords. It also generates general text statistics. Please paste the text for keyword analysis. The found keywords will be listed with total count and percentage (keyword density).

Takaaki Kajita of the University of Tokyo and Arthur B. McDonald of Queen's University in Canada were awarded the Nobel Prize in Physics on Tuesday for discovering that the ubiquitous but elusive subatomic particles known as neutrinos have mass.

Neutrinos are the second most abundant subatomic particles in the universe after photons, which carry light. Their existence was predicted in 1930, but decades they remained some of the most enigmatic elements of astrophysics.

The ghostly neutrino — it means "small neutral one" in Italian — carries no electric charge and is so light that it had been assumed for many years to have no mass at all. The scientists showed that neutrinos, which are found in three "flavors," could oscillate from one flavor to another — changing as they traveled through the atmosphere or water, thus demonstrating that they have mass.

This is the text analysis result. The calculation took about 112.80 ms.

Characters: 5099  
Characters (without spaces or punctuation): 4001  
Most counted: 65  
Total word count: 1033

Count	Percentage	Found keyword
65	6.29 %	the
16	1.55 %	and
15	1.45 %	that
14	1.36 %	neutrinos
11	1.06 %	prize
9	0.87 %	for
7	0.68 %	they
7	0.68 %	universe
7	0.68 %	from
6	0.58 %	was
6	0.58 %	kajita
6	0.58 %	university
6	0.58 %	have
5	0.48 %	through

**Takaaki Kajita and Arthur McDonald Share Nobel in Physics for Work on Neutrinos**

By DENNIS OVERBYE, OCT. 6, 2015

Takaaki Kajita of the University of Tokyo and Arthur B. McDonald of Queen's University in Canada were awarded the Nobel Prize in Physics on Tuesday for discovering that the ubiquitous but elusive subatomic particles known as neutrinos have mass.

Neutrinos are the second most abundant subatomic particles in the universe, after photons, which carry light. Their existence was predicted in 1930, but for decades they remained some of the most enigmatic elements of astrophysics.

The ghostly neutrino — it means "small neutral one" in Italian — carries no electric charge and is so light that it had been assumed for many years to have no mass at all. The scientists showed that neutrinos, which are found in three "flavors," could oscillate from one flavor to another — changing direction like a zig-zag as they traveled through the atmosphere or through water, thus demonstrating that they have mass.

Dr. Kajita was part of a team that discovered neutrinos from the atmosphere in 2002, but for decades they remained some of the most enigmatic elements of astrophysics.

Arthur B. McDonald in Kingston, Ontario, on Tuesday. Lars Hagberg, via Reuters

1. Inconvenient or Hard to Understand (Not Intuitive)
2. 'Control' element unavailable for users (No Interactivity)

# Solution modeling (keyword-visualizer)

The image shows a screenshot of a New York Times article titled "Farhad and Mike's Week in Tech: Figuring Out Twitter's Future". A blue rounded-rectangle overlay is placed over the main text of the article, containing a keyword-visualizer tool. The tool features two vertical sliders on either side, and a central word cloud. The word cloud is composed of various words, with "Farhad" and "Mike" being the most prominent. Other visible words include "tech", "Google", "Amazon", "Microsoft", "Twitter", "industry", "news", "things", "the", "a", "of", "and", "but", "to", "is", "was", "sort", "of", "checked", "out", "of", "the", "news", "flow", "but", "came", "back", "to", "a", "deluge", "of", "tech", "craziness", "Google", "is", "doing", "this", "thing", "with", "a", "bunch", "of", "other", "tech", "companies", "to", "make", "web", "pages", "load", "faster", "on", "our", "phones", "which", "is", "theoretically", "a", "good", "thing", "Jet", "the", "online", "shopping", "Amazon", "competitor", "killed", "one", "of", "its", "main", "business", "model", "decisions", "to", "charge", "for", "a", "membership", "which", "seems", "theoretically", "like", "a", "bad", "thing", "for", "Jet", "at", "least".

Each Saturday, Farhad Manjoo and Mike Isaac, technology reporters at *The New York Times*, review *the week's news*, offering analysis and maybe a joke or two about the most important developments in the **tech** industry.

**Mike**: Hello, **Farhad**! How are you? I just got back from a week of "new hire orientation" for **The Times**. I've been an employee here for 18 months. I guess we run on a slightly different timetable.

**Farhad**: To be fair, 18 months doesn't seem like nearly enough time to complete a thorough background check on you.

**Mike**: Fair. So, I was sort of checked out of the news flow, but came back to a deluge of **tech** craziness.

Google is doing this thing with a bunch of other **tech** companies to make web pages load faster on our phones, which is theoretically a good thing. Jet, the online shopping Amazon competitor, killed one of its main business model decisions — to charge for a membership — which seems theoretically like a bad thing (for Jet, at least).

**Farhad**: To be fair, 18 months doesn't seem like nearly enough time to complete a thorough background check on you.

**Mike**: Fair. So, I was sort of checked out of the news flow, but came back to a deluge of tech craziness.

Google is doing this thing with a bunch of other tech companies to make web pages load faster on our phones, which is theoretically a good thing. Jet, the online shopping Amazon competitor, killed one of its main business model decisions — to charge for a membership — which seems theoretically like a bad thing (for Jet, at least). And Dell and EMC are considering some sort of giant merger or takeover, which is still in its theoretical stage. Also, I have no idea how to speak intelligently on the cloud — the cloud! — so let's just tiptoe past that.

**Farhad**: You forgot about the Microsoft event. They made a laptop! With a screen that's actually a tablet! I thought it was pretty cool. Maybe PCs are cool again?

# Solution modeling (keyword-visualizer)

## ► Idea

1. Implementation in web browser as a 'Plug-in' (ex. Chrome plug-in)
2. Variations in Size, Color, Shading of terms depending on its importance in document (e.g. WordCloud)
3. Application of various text retrieval algorithms (Radio button for selection)
4. Evaluation and Improvement of algorithms through feedback button
5. Recommendation for a particular text retrieval algorithm based on user or genre of document
6. Slider bar to control the number of keywords shown