

NEW  [Join the conversation by adding Mashable's Pete Cashmore to your Circles on Google+](#) [Close](#)

Welcome to Mashable! Sign in with [Facebook](#) or [Twitter](#) to get started!

[Semantic Text Startup Reads and Summarizes Documents for You](#)



February 8, 2011 by [Jennifer Van Grove](#) 74



The [Spark of Genius Series](#) highlights a unique feature of startups and is made possible by [Microsoft BizSpark](#). If you would like to have your startup considered for inclusion, please see the details [here](#).

Name: [Topicmarks](#)

Quick Pitch: Topicmarks summarizes your cloud documents and extracts the most important points for you.

Genius Idea: If you've ever come across a document too large to digest in a few minutes, you're likely to find a new friend in Topicmarks.



The startup's beta product represents more than four years of machine learning and artificial intelligence research that can statistically analyze cloud documents, or those you upload to service, to quickly distill them down to six sections: Overview, Facts, Summary, Keywords, Index and Properties.

The technology is so adept at reading text and understanding context that Topicmarks can cut up the full text within documents into sentences, assign meaning to words and pluck out relevant facts, all within minutes. Think of it as a better, machine-enabled version of CliffsNotes — but for all text documents.

Topicmarks users can upload documents, use the bookmarklet to grab web content, e-mail in text, or copy and paste text for nearly instant analysis. Key points are extracted and personalized to the user.

The company is also working on an [Evernote](#) plugin capable of automatically tagging new notes, applying tags to old notes and adding-in quick links to condensed summaries. Currently, users can already import their Evernote notes to view summaries.

Topicmarks technology could also be applied to distilling key points from social media streams or RSS feeds, avenues that the startup may choose to pursue at a later date. For now, the immediate focus is integrating with popular cloud services such as Dropbox or Scribd and improving the design of the site to make it more appealing to end users.

The product, which has an infinite number of use cases, is primarily targeted at the more than 50 million information workers in the world producing or reproducing information, says CEO Roland Siebelink.

Interestingly enough, the technology is already proving popular with students and teachers, though a few in the latter group have e-mailed to complain that their students are "cheating" with the help of Topicmarks. Siebelink tells the story of a disgruntled teacher who awarded one sneaky student with the highest mark on a class assignment, only to later discover that the summary had been copied word for word from Topicmarks.

This real-life example proves the quality of Topicmarks' results, says Siebelink, who believes the tool ultimately helps both students and education workers.

Topicmarks is gaining interest from both end users and investors after winning two startup competitions — the winning [Founder Showcase](#) pitch is embedded below. The startup is reviewing term sheets from institutional

investors and is said to be going after \$1 million to \$2 million in Series A funding, according to a source familiar with the matter.

The product is free for the time being, but the company will enact a freemium pricing structure in the future that may charge to process larger documents, says Siebelink.

Image courtesy of [iStockphoto, thesuperph](#).

Series Supported by Microsoft BizSpark



The [Spark of Genius Series](#) highlights a unique feature of startups and is made possible by [Microsoft BizSpark](#), a startup program that gives you three-year access to the latest Microsoft development tools, as well as connecting you to a nationwide network of investors and incubators. There are no upfront costs, so if your business is privately owned, less than three years old, and generates less than U.S.\$1 million in annual revenue, you can [sign up today](#).

[Email Story Reprints](#)

We recommend

[Startup Seeks to Be Online Destination for World's Best Photography](#) *Startups*

[HOW TO: Land a Job at LivingSocial](#) *Startups*

From around the web

[50% unemployment, 90% stock market drop, 100% inflation. See the Evidence](#) *Newsmax.com*

[A golden rule: Discipline trumps conviction](#) *MarketWatch.com*

[5 Big Mistakes To Avoid in Your QR Code Marketing Campaign](#) *Business & Marketing*

[SFPark – The iPhone App for Smart Parking in San Francisco](#) *Reset San Francisco*

[IE Users Have Lower IQ Than Users of Other Web Browsers \[STUDY\]](#) *Tech & Gadgets*


[SES San Francisco 2011 Fast Approaching Register Today](#) *ReelSEO*


[Apple Launches iCloud Beta \[PICS\]](#) *Dev & Design*


[Guinness World Record for Babylon.com](#)
Babylon.com

[\[what's this\]](#)

12 Comments

-
1.  [Ashish Gurung](#)
Four years of research went into this? Wow. Hopefully it'll pay off for them!
<http://www.QuikPig.com>
[February 8, 2011 06:59:00 Reply](#)

-
2.  M9R
This very article, summarized by Topicmarks:
1. Topicmarks is gaining interest from both end users and investors after winning two startup competitions
 2. Topicmarks users can upload documents, use the bookmarklet to grab web content, email in text, or copy and paste text for nearly instant analysis
 3. startup's beta product represents more than four years of machine learning and artificial intelligence research that can statistically analyze cloud documents, or those
 4. users import Evernote notes to view summaries
 5. Siebelink tells story of a disgruntled teacher who awarded one sneaky student with the highest mark on a class assignment, only to later discover that the summary had been copied word for word from Topicmarks
 6. Topicmarks summarizes cloud documents and extracts the most important points for you
 7. technology is so adept at reading text and understanding context that Topicmarks can cut up the full text within documents into sentences, assign meaning to words and pluck out relevant facts, all within minutes
 8. company working on an Evernote plugin capable of automatically tagging new notes, applying tags to old notes and adding-in quick links to condensed summaries
 9. you upload six sections: Overview, Facts, Summary, Keywords, Index and Properties
 10. Topicmarks is delicious with a white wine sauce
- (okay so originally #10 was basically a repeat of #7, otherwise not bad)
- [February 8, 2011 07:18:00 Reply](#)

-
3.  [donfelipe](#)
This is a good technology but my documents are too small so I don't have a need for this technology right now. But I wish them all the luck with this new venture.

[February 8, 2011 07:18:00 Reply](#)

4.



[znmeb](#)

Freemium? For four years of cutting-edge NLP research? They're insane! Ask for the *sale*!

[February 8, 2011 07:19:00 Reply](#)

5.



[Matt Ormrod](#)

I just tried this with a few documents and everything that it returned was broken biscuits, it just didn't make sense. I think it's a really good idea that would be really useful, maybe I was unlucky with my document choice, I'll give it another go!

[February 8, 2011 08:49:00 Reply](#)

o



[Roland Siebelink](#)

Hi Matt, you don't say what documents you tried to upload but Topicmarks usually works best with long and dense documents, especially for the first ones you upload. Looking forward to your feedback!

[February 8, 2011 14:52:00 Reply](#)

▪



[Matt Ormrod](#)

Hi Roland,

The documents were around 2,000 words in length. Maybe I'll try again with longer documents...

[February 8, 2011 14:56:00](#)

6.



[Personal Version Control](#)

Nice app. It would be nice if they could make it downloadable. Some documents can not be uploaded to the cloud

[February 8, 2011 14:26:00 Reply](#)

7.




[Free Enterprise](#)

This software is very innovative! Topicmarks is a great example of the power and creativity of American free enterprise. I work with the Campaign for Free Enterprise, a project of the U.S. Chamber of Commerce, and we believe that startups like Topicmarks are the backbone of our economy. Free enterprise means that companies are able to compete in a relatively free market with minimal road blocks to entry. Under this system, we've seen technology startups, like Topicmarks, thrive.


We hope you'll check out our content on startups and entrepreneurship at <http://www.freeenterprise.com/entrepreneurs>. Give us your opinion and join the conversation about free enterprise on Facebook at <http://www.facebook.com/AmericanFreeEnterprise>.

-Hilary, freeenterprise@uschamber.com


[February 9, 2011 03:05:00 Reply](#)

8.  [Cornelius Wilson](#)
- My issue with these kind of text to speech sort of thing is that the words are not pronounce properly and it is just terrible to listen to. I just don't think technology is ready for this kind of thing yet, as it has been attempted many times in the past and I have yet to come across one that is worth using.

[February 11, 2011 17:14:00 Reply](#)

-  [Roland Siebelink](#)
- Hi Cornelius, we don't do text-to-speech. Maybe you confused us with another company?

[February 15, 2011 22:39:00 Reply](#)

9.  [Luke Metcalfe](#)
- Summarisation is a huge challenge and I wish them all the best!

We have a different approach at <http://www.thefullwiki.org> where we view the summary of the topic as the Wikipedia article and provide related sentences to each. So we start where the students start – at Wikipedia, and help them make the next step. It makes it very easy for students to jump to relevant sources and cite them.

[March 8, 2011 06:00:00 Reply](#)