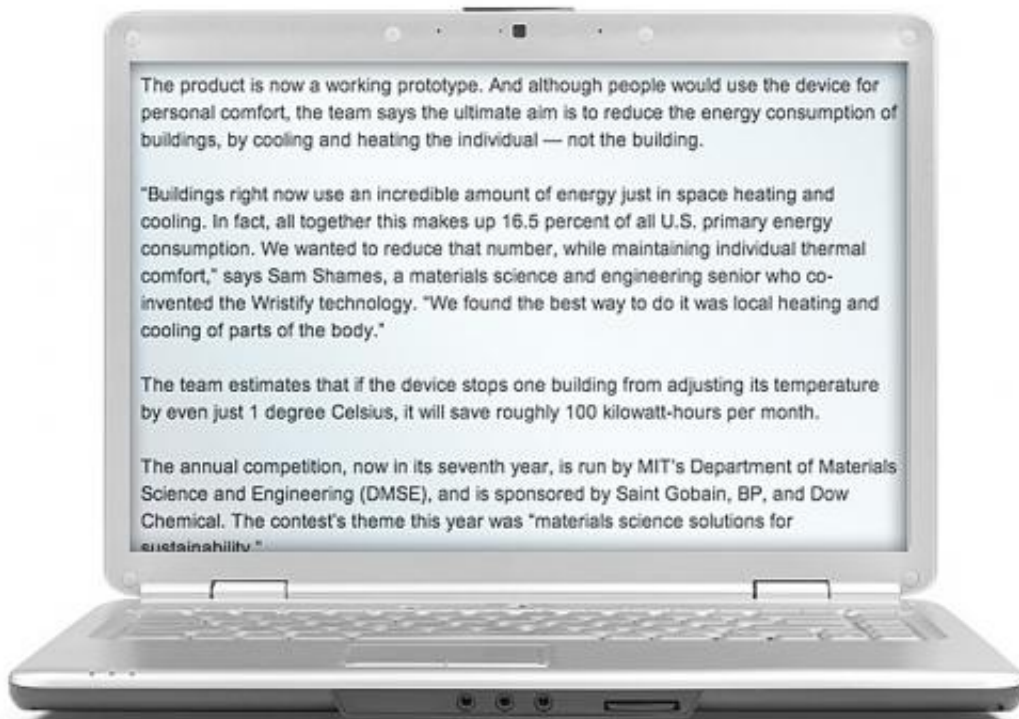


# Powering better online document viewing

October 30 2013, by Rob Matheson

---



Tech startup Crocodoc, co-founded by MIT alumni, develops and commercializes online tools that convert PDF, Microsoft Office, and other text documents into HTML, making these files easily viewed and shared across the Web, and on any device.

Viewing PDF and Microsoft Office documents on a Web browser can cause slow loading and messy formatting—and often such documents won't load at all. Most times, users will simply download the documents

to their computers to read and annotate a clean copy.

This type of thing doesn't happen with, say, videos and images, because file-sharing websites, such as YouTube and Flickr, can convert various uploaded file types into a format supported by all browsers.

Now tech [startup](#) Crocodoc, founded by MIT engineering and computer science alumni, has developed online tools that convert an array of document formats into HTML, making these files easily viewed and shared—much like videos and images—across the Web, and on any device.

"We've spent an enormous amount of time understanding documents at a very deep level so that we can reconstruct them in your Web browser or mobile device in a fast and high-quality way," says Ryan Damico '06, Crocodoc's co-founder and CEO.

Using Crocodoc's tools, clients can upload a PDF or Microsoft Office document and rapidly receive an HTML version of that same document in their browser, which can be shared and annotated in real-time. Crocodoc provides an application programming interface (API) to developers to integrate into their Web services, so users don't need to download large files or use desktop software.

Against the backdrop of our burgeoning digital-file-sharing world, Crocodoc's document-viewing solution has become a profitable endeavor. Launched in 2010, San Francisco-based Crocodoc is now powering document-viewing features for big Web companies such as LinkedIn, Yammer, Blackboard, Edmodo, and SAP.

According to Crocodoc, its tools have powered more than 200 million document conversions and 14 million document annotations.

In May, cloud-storage giant Box, which focuses on file sharing among businesses, acquired Crocodoc and its [technology](#) for an undisclosed amount. There, the startup is positioned to expand, Damico says. "With Box, we're staying true to our vision at Crocodoc, but have 10 times more resources at our disposal," he says. Box aims to soon swap out its current document-viewing mechanism with Crocodoc's, as well as release a new API that will allow third-party businesses to use the latest version of Crocodoc's technology.

## WebNotes to Crocodoc

Crocodoc is actually an offshoot of WebNotes, a startup launched out of an MIT dorm room by Damico and his Crocodoc co-founders—Bennet Rogers SM '07, Matt Long '08, and Peter Lai '08, SM '09—that allowed users to highlight and annotate text on Web pages.

Shortly after graduating from MIT, the team spent nights and weekends growing the startup. But for a number of business- and technology-related reasons, Damico says, WebNotes became a "spectacular failure," and the team members found themselves about to go under. (Primarily, they couldn't find customers to buy their product.)

After running out of capital and nearly folding their company, they entered California's startup accelerator, Y Combinator, where they soon had an epiphany: The document-viewing technology used for WebNotes actually functioned better than any similar technology, and was far more marketable.

Most technology that allowed online document viewing, Damico explains, had to generate an image of each page, which was slow, low-quality, and plagued by formatting issues. Instead, Crocodoc strips the contents of a document and reconstructs them to meet Internet standards: It converts the text to HTML and the images to Scalable Vector

Graphics, and formats the page using Cascading Style Sheets.

This represents a novel approach to online document viewing, and a tough technological challenge, Damico says. "What we're doing is taking documents and recreating them flawlessly, treating text, lines, and shapes as native objects in your browser so that documents look just the way you'd expect them to when opened on your computer," he says. "And all this has to be fast and responsive, so it works on your mobile device. It's really difficult to meet both of those standards at the same time."

Seeing commercial value in this technology for larger file-sharing services, WebNotes pivoted to Crocodoc over the course of a single weekend. The team designed a completely new Web page and focused on licensing their document-viewing product to enterprises, instead of selling it directly to individual customers. "Once companies saw what we had to offer, we had big clients knocking at our door," Damico says.

## **Finding a problem to solve**

Although it was a meandering road to Crocodoc, the team's early entrepreneurial roots trace back to MIT, where, Damico says, "a lot of ideas were fleshed out. MIT was a great place to start, because there's such a vibrant entrepreneurship community there."

As WebNotes, the team found guidance from the Venture Mentoring Service (VMS), "which was a fantastic organization that helped us think through our ideas."

VMS mentors, for instance, put focus on identifying markets, and on finding customers and potential partners. "It really challenged us, and if it wasn't for them, we wouldn't have even gotten to a point where we'd apply for Y Combinator or think more broadly about business plans," Damico says.

Today, after years of struggling with WebNotes, and then running a successful operation with Crocodoc, Damico says he has learned two key lessons for entrepreneurship: Always speak with, and find, potential customers, and—first and foremost—focus on solving a real problem.

"Crocodoc's success came down to being persistent and having a good nose for finding a real problem to solve," he says. "We saw a larger problem to be solved, so we focused on what we could do best: developing the world's best online document-viewing technology. That's how we took off."

In the future, Damico says, Crocodoc's technology could also have broader, societal implications: For instance, Crocodoc's technology could be used in the health-care industry, making patient records and medical documents easily accessible as digital HTML [documents](#) that could be accessed from browsers and [mobile devices](#).

*This story is republished courtesy of MIT News ([web.mit.edu/newsoffice/](http://web.mit.edu/newsoffice/)), a popular site that covers news about MIT research, innovation and teaching.*

Provided by Massachusetts Institute of Technology

Citation: Powering better online document viewing (2013, October 30) retrieved 4 May 2024 from <https://phys.org/news/2013-10-powering-online-document-viewing.html>

<p>This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.</p>
--