

# Document 3.0

How the next generation  
of dynamic documents will  
improve the way we work.

## Contents

Executive Summary.....	2
A preview of coming attractions .....	3
The 25,000-year history of Document 1.0.....	4
Document 2.0: The digital revolution.....	4
Document 3.0: The next great leap.....	5
New types of documents. And multi-author mash-ups ..	6
Document 3.0 will change the way we work .....	8
Challenges on the road to the future.....	9



# The continuing evolution of the document.

## Executive Summary

The headlines are full of visionary predictions about our cloud computing future. Virtualization is a hot topic. And people are already talking about the next phase in the evolution of the Internet... the Semantic Web or Web 3.0.

All of these developments are very exciting and suggest that we are rapidly approaching another Information Age inflection point.

What gets far less attention, however, is the changing role of documents.

The simple fact is, all of the technological advancements mentioned above will be paralleled by significant changes in the form, function and even the definition of documents. And that will inevitably have important implications for the way people work.

After all, documents are woven into our daily activities. They support virtually every organizational goal. And they have a direct impact on the efficiency of key business processes and the productivity of Knowledge Workers and other employees.

That's why it's so important for organizations to keep track of the changing nature of the document and prepare to take advantage of the dynamic capabilities of the rapidly approaching Document 3.0 era.

## A preview of coming attractions.

How will documents change and change the way we work? I'll provide a more detailed discussion later in this white paper. But here is a brief overview.

The next generation of documents will be:

- More intelligent and proactive so they can perform tasks on our behalf
- Self-structured to improve search and content mining
- Actively seeking and assembling relevant information, multimedia content and hyperlinks from a variety of sources
- Cloud-based and more dynamic
- Free from the restrictions imposed by proprietary, application-based formats

Changing documents will also:

- Virtually eliminate the barriers between paper and electronic documents
- Liberate content from the document containers so “chunks” can be easily accessed, repurposed and imported into other documents
- Give people the freedom to pursue an individualized, nonlinear information-gathering path through multiple linked documents
- Facilitate collaboration

These are some of the coming attractions of the Document 3.0 age. And they will all help us bring more speed, automation and efficiency to the business world. The Document 3.0 era will also make documents more powerful and effective.

### **Re-defining the document for the modern age.**

I'll describe the specific characteristics of 3.0 documents in more detail later. For now, however, I want to add some important context.

First of all, let's define what we mean by the word “document.” To some people, a document is a static container for information. But from my perspective, that's only half the story.

If you talk about the constitution of a particular country, for example, you may be talking about a historic artifact on display in a national archive or museum. Or you may be talking about the information contained in that artifact that still serves as a guide for legal conduct today.

In other words, the word document applies to both the container and the contents. This is a key principle for our discussion of the emerging possibilities of Document 3.0.

In addition, it's important to briefly describe the previous stages in the document's evolution to gain some valuable perspective. So let me take a moment to talk about Document 1.0 and 2.0. Then it will be easy to appreciate the remarkable changes that have taken place in documents and document-driven business processes.

## The 25,000-year history of Document 1.0.

In a sense, the Document 1.0 era began more than 25,000 years ago with the Paleolithic paintings on cave walls in my home country of France. Those are the earliest examples we have of people storing images, information or ideas on a particular medium.

If you think about it, those walls were probably fairly easy to draw or paint on—as long as you didn't make any mistakes. And they served the images well by preserving them for eons. They also exhibit one central feature of 1.0 documents: the content is permanently wedded to the container.

Of course, rock walls have serious limitations when it comes to sharing information outside your “inner circle.” So when writing systems began to develop in the Bronze Age more than 6,000 years ago, people invented other media that were much easier to use, share, distribute and store.

Stone and clay tablets, animal bones and shells, various metals... they all played a role in the early days of documents. But then paper was invented in China during the Han dynasty. And for the next 2,000 years, physical paper documents were the dominant medium for all written and recorded information. These documents exhibit the essential nature of Document 1.0:

- A physical format
- Flat and static
- Irreversible content locked into the document “container”
- Information can only be accessed by viewing the actual layout

## Document 2.0: The digital revolution.

Now let's fast forward to the late 20th century. Suddenly a new kind of document appears on the computer screen composed of bytes and bits. This digital creation emulates 1.0 documents in appearance. But it has clear-cut advantages over paper and ink.

The digital document is easier to create, revise, search, store and share. And technological advancements keep making it better. Here are some of its essential characteristics:

- Its digital format mimics paper-based documents
- Dynamic instead of static content
- Easy-to-create with sophisticated design and layout applications that include some auto formatting capabilities
- Increasingly uses color, personalization, relevant one to one content and hyperlinks
- More open XML-based formats and standards
- Better access to content with metadata and logical outlines

- Helps automate document-driven business processes by replacing paper documents in manual workflows
- Reduces environmental impact

Thanks to these valuable qualities, digital documents have accelerated the flow of information and helped organizations work more efficiently and effectively. In addition, the documents themselves have become more powerful and effective.

All of these advancements have had a revolutionary impact on the workplace.

### **Document 3.0: The next great leap.**

If you take a look around, it's clear that we are still in the Document 2.0 era. At the same time, documents are continuing to evolve, keeping pace with other Information Age advancements like the Semantic Web. So what will our "next generation" documents be like?

It's impossible to predict the future. But I believe the 3.0 era will represent another giant leap forward. Here are some of the reasons why.

#### **Easier access with cloud-based documents.**

Since documents are an integral part of the Information Age, any game-changing innovation that affects IT and the Internet will change them as well.

Take cloud computing, for instance. It has the potential to totally change the way organizations manage information technology by turning applications, infrastructure, and storage for documents and data into services accessed by a Web browser from any computing platform anywhere in the world.

The implications of this technological sea change are enormous. Documents and the information they contain will become more accessible to an increasingly mobile work force. It will also be easier for people to collaborate, which will improve the quality and effectiveness of their documents.

This cloud-computing environment is a distinctive feature of Document 3.0.

#### **Freedom from applications and proprietary document formats.**

With the development of cloud computing, Knowledge Workers will gradually become less conscious of the specific software applications they are using to create and manage documents. In fact, those cloud-based applications will eventually work behind the scenes much like most operating systems do today.

As a result, you will be able to quickly create any type of document on any computing device without having to open a particular application or office suite. These changes will make it easier and more intuitive to work with documents.

## **New types of documents. And multi-author mash-ups.**

Web pages, blogs, dashboards and social networking sites have already changed the modern definition of the document. In the future, that definition will expand as new formats for presenting and sharing text, data, images and multimedia content are developed.

In addition, the content of documents will continue to change as they automatically aggregate up-to-date information from multiple authors and sources. (I'll talk more about "evergreen" documents later.) This powerful trend, which is already underway, will literally transform the way people use documents and gather information.

It's another example of how changes in the nature of the document change the way people work every day.

### **More standardization and structure.**

In the Document 3.0 era, proprietary, application-based document formats will gradually disappear, replaced by open standards like Extensible Mark-up Language (XML) which will bring more structure to unstructured content.

XML will simplify the exchange of documents, increase access to the information contained in them and facilitate automation in document-driven business processes.

The development of XML-based standards like HR-XML in Human Resources and XBRL for financial services will also improve information access and efficiency in specialized back office functions and vertical industry applications.

Technological advancements will also help companies quickly convert legacy paper and digital documents into XML formats to facilitate machine processing, which will be a defining characteristic of the Semantic Web.

### **Remarkable advances in content mining.**

As everyone knows, the amount of information available on the Internet is growing at an exponential rate. In fact, the market research firm IDC estimates that more than 2,400 exabytes will be added to the cloud by the year 2012. Which brings up the big question: With that much information floating in the digital universe, how do you find the particular molecule that you need?

In the world of Document 3.0, search, retrieval and content mining—which essentially means maximizing the use of available information—all become easier than ever due to the development of innovative technologies like Factspotter™ and Hybrid Categorization.

Factspotter™ is a natural language processing technology from the Xerox Innovation Group that helps computers understand languages and turn unstructured text into coded data. Once that challenging task is accomplished, you can mine unstructured documents like a structured digital database. This 3.0 innovation will undoubtedly have powerful applications in litigation, medical research, risk management and many other fields.

The **Hybrid Categorizer** is another Xerox invention. It uses visual and textual clues to automatically sort and route documents and extract key information. The categorizer will make a huge contribution in the effort to fight Information Overload, since research suggests it can take up to 15 days for a document to reach the right person in the age of overflowing email in-boxes.

These and other technologies will use logic, natural language processing and semantics to break down the barriers between structured data and unstructured documents so we can take full advantage of all the information available to us.

### **Dynamic paper documents.**

Most of the advancements I've talked about so far relate to digital documents. But in the Document 3.0 world, paper documents will have remarkable capabilities, too.

- Document 2.0 already gave us innovations like optical character recognition (OCR), glyphs and bar codes that help us bring paper documents into a digital workflow. Document 3.0 will build on this foundation with Radio Frequency ID (RFID) tags that can be printed on paper documents to facilitate tracking. The tags can even be designed to set off alarms at building entrances and exits. That's a big step forward for organizations trying to maintain tight control over proprietary information and paper documents.
- Technological advancements will enable us to add computer processing power to paper documents by printing micro-batteries, circuitry and antennas right on the page. As a result, these documents will become intelligent, self-powered devices with remarkable sensing and communication capabilities.
- A few years ago, Xerox researchers noted that up to 40% of printed office documents were used for less than a day before being recycled or discarded. That insight helped inspire the development of reusable paper that automatically erases content after 24 hours. This Document 3.0 innovation will definitely help us reduce office waste, save on paper costs and advance the cause of environmental sustainability.
- Electronic paper and e-readers are other exciting developments that will reach fruition in the Document 3.0 era. It will have applications in everything from advanced electronic readers to merchandising displays.

These innovations will help bring one of the world's most successful technologies—the paper document—into the dynamic Document 3.0 world.

### **Passive documents become proactive agents.**

I've saved the most exciting development for last. In the 3.0 world, documents will not only become more dynamic and intelligent, they will become proactive agents, working on our behalf without our conscious involvement or direction.

They will learn our viewing preferences and present information in an appealing way, no matter which viewing platform or device we are using at the time.

- They will be evergreen, which means they will automatically search for information that will keep the content relevant and up-to-date. They will also retire themselves to an archive when they have reached the end of their useful lifecycle.

- They will help us automate time-consuming steps in document-driven business processes, which will help us increase efficiency and reduce costs.
- In marketing and customer communication applications, they will search databases, social networks and the Internet for information and insights that will make their content more personalized, relevant and effective on a one-to-one basis.
- And even though these 3.0 documents will be constantly changing in terms of both form and content, they will maintain a comprehensive audit trail so Knowledge Workers can easily track their evolution.

All of these predictions are based on emerging technologies and innovations in development today. But by the time the Document 3.0 era arrives, there will undoubtedly be other exciting chapters to add to the story.

## **Document 3.0 will change the way we work.**

Throughout history, changes in the form and content of documents have had a profound impact on the way people work. That will certainly be the case in the next stage of the document's evolution—Document 3.0.

With cloud-based documents and applications, it will be easier and more economical for people to create powerful documents, access information and collaborate. Technological advancements will ensure information security and provide a comprehensive audit trail for documents, which will help people track every stage in their development. The cloud computing environment will also help level the playing field between small- and-medium-sized businesses (SMBs) and large global companies, since every enterprise will have access to the same dynamic services.

The disappearance of proprietary formats will make it easy to exchange documents and incorporate information from virtually any source without tedious reformatting.

Technological advancements will bring coded structure to the content of all paper and digital documents, dramatically expanding the amount of proprietary and Internet-based information that's available for our use.

Breakthroughs like reusable and electronic paper documents will help organizations reduce costs and reduce their environmental impact by continuing the remarkable advancements in sustainability made possible by digital documents.

Other advancements will help us manage the very real problem of Information Overload by automatically sorting and prioritizing email and other documents for us.

Efforts to standardize document formats and add structure to content will unlock information from their original document containers and turn all digital information into a structured, searchable database.

Proactive documents will save us time by automatically updating information, making the content more relevant to recipients and by shaping their visual representation to suit our preferences.

In addition, the very concept of documents will radically change as they become more fluid and fast-changing.



### **Specific applications of 3.0 advancements.**

All of these breakthroughs will have a sweeping impact on Knowledge Workers and document-driven workflows and business processes.

- Organizational mailrooms will become models of efficiency by leveraging the automatic routing, filing and sorting capabilities of innovative technology and intelligent, proactive documents.
- Standard back office processes like Accounts Payable that still rely on manual, paper-based workflows will make the leap into the digital age, improving cash management strategies and vendor loyalty.
- Complex step-by-step approval processes for documents like mortgages will be re-engineered to provide simultaneous access to multiple reviewers inside and outside the enterprise, dramatically reducing cycle times and boosting customer loyalty.
- The widespread use of electronic medical records will improve the quality and efficiency of health care organizations.
- Processes for managing the huge volume of documents involved in litigation and regulatory approvals will become optimized and digitized, saving millions of dollars and reducing the risk of fines for noncompliance.
- And last but not least, organizations that make a strategic commitment to employ Document 3.0 best practices and technologies will move ahead of their competitors, thanks to document-driven improvements in efficiency and effectiveness.

### **Challenges on the road to the future.**

Document 3.0 will bring major improvements to the workplace. But we will have to overcome some significant obstacles to reach the full potential of the next stage in the document's evolution.

The rapid development of cloud computing raises important questions about privacy, information security and business continuity, since documents and data will be stored outside enterprise firewalls. We have to find ways to address these issues.

Knowledge Workers will have to adjust to a new paradigm based on virtualization, since they will no longer have physical document originals in their possession and control.

Organizations will have to eliminate the bottlenecks in their business processes caused by the reliance on paper documents and manual workflows.

We will have to learn to think outside the "silo" box. At the present time, document-driven business processes often run through different organizational silos. As a result, improvement efforts are often silo-based and fragmented, falling short of their true potential. If we break down the barriers between silos and get all stakeholders to embrace a common vision, we can launch holistic business process optimization programs that will fully utilize Document 3.0 innovations.

These issues are significant. But I believe that will all be solved with technology, the development of best practices, and by the powerful logic that compels us to take full

advantage of better ways to work.

**A final word.**

In the past quarter century, documents have gone through dramatic changes. From fixed, physical containers, they have become dynamic mosaics of information, including text, images, multimedia content and hyperlinks.

In the future, the very idea of what constitutes a document will continue to evolve. Instead of fixed containers of static information that have to be read start to finish, documents will become dynamic sources of content that can be quickly accessed, analyzed and imported into other documents, some of which may exist only for a single viewing.

In addition, the passive documents that have served us so well through the course of human history will be replaced by a new breed of documents that possess the intelligence to act as agents on our behalf, simplifying the search for information and automating time-consuming tasks.

One thing that won't change, however, is the very important role that documents play in our daily lives. Whether they are fixed in form or constantly changing... whether they are inanimate tools or active agents... whether they last for hundreds of years or a single moment in time, documents are essential for organizational progress.

They sell products and services. They build loyalty. They keep track of records and transactions. They establish strategy. They provide direction and inspiration. They turn raw data into business intelligence. And they give birth to innovation. As a result, they have a direct impact on virtually every activity every single day.

In the next phase of their evolution—Document 3.0—they will still perform all of those vital duties. They'll just do it faster, easier and better.

You've got my word on that.

## About the Author

As Managing Principal, Technology Innovation, within Xerox Global Services, François Ragnet leads a team charged with transferring novel technologies into mainstream Xerox solutions offerings. Current initiatives focus on text-, image-, or feature-based categorization of documents, as well as identifying deeper semantic analyses which will enable “smart document” generation from traditional legacy formats or paper. His team also focuses on enhancing current offerings within the office environment to improve the efficiency of current products and streamline support processes.

Previously, Ragnet served as program manager and senior project leader for the Xerox Research Centre in Europe where his team provided innovative technologies in support of next-generation Xerox offerings. Their goal was to provide solutions that allow users to filter, configure and extract information from documents across advanced platforms that bridge production, printing and scanning. François was also a project leader in wireless technologies—specifically mobility and wireless (Bluetooth), content management, security, print and infrastructure management.

While at the Research Centre of Europe in Grenoble, France, Ragnet was a founding member of the Technology Showroom—a showcase of experimental technologies that hosts international events with customers from all over Europe.

Prior to joining Xerox, Ragnet was a researcher at the National Institute of Standards and Technologies, in Gaithersburg, MD, where he focused on creation of a demonstration platform for state-of-the-art collaborative work technologies. He holds a master’s degree in telecommunications, from the Institut National des Telecommunications, Paris, France.

Visit Francois’ page on the Xerox Thought Leadership Site for more information about Francois and to find the following:

- Francois’ Full Biography PDF
- Ask the Expert
- Request a Speaker
- Sign up for Francois’ eNewsletter
- Hear Francois talk about The Future of Documents

## Xerox Global Services

Xerox Global Services is a world leader in document outsourcing services with industry expertise that helps you reduce costs, achieve operational excellence and grow revenue. Our unique combination of experience, technology and delivery capabilities enable integration, transformation and continuous innovation of your office environment, centralized print production, and communication and business processes. We deliver excellence in every corner of the world—locally, nationally, globally.

