Content-Centered Collaboration Spaces in the Cloud

John Erickson, Susan Spence, Michael Rhodes, David Banks, James Rutherford, Edwin Simpson, Guillaume Belrose, Russell Perry

HP Laboratories
HPL-2009-11

Keyword(s):
cloud computing, collaboration, content management, communities, digital natives

Abstract:
In this article we share our vision for a new approach to content-centered collaboration. Emphasizing communication, collaborative work and community, we envision a cloud-based platform that inverts the traditional application-content relationship by placing content rather than applications at the center, enabling users to rapidly build customized solutions around their content items. We believe the future of collaboration will focus on building and sustaining communities around content, tasks and ideas, and we see hosted entities known as content spaces as a means to supporting ecosystems of users and developers around that content. To make this case we review the dominant trends in computing that motivate our exploration of new approaches for content-centered collaboration. We offer insights into how certain core problems for users and organizations are being addressed today. We identify challenges that must be conquered before solutions can be delivered, and we outline next steps for research in the area.
Content-Centered Collaboration Spaces in the Cloud

A submission to IEEE Internet Computing special issue on Cloud Computing

John Erickson, Susan Spence, Michael Rhodes, David Banks, James Rutherford, Edwin Simpson, Guillaume Belrose, Russell Perry

john.erickson@hp.com, susan.spence@hp.com, michael.rhodes@hp.com, dbanks@hp.com,
james.rutherford@hp.com, edwin.simpson@hp.com, guillaume.belrose@hp.com, russell.perry@hp.com

Abstract

In this article we share our vision for a new approach to content-centered collaboration. Emphasizing communication, collaborative work and community, we envision a cloud-based platform that inverts the traditional application-content relationship by placing content rather than applications at the center, enabling users to rapidly build customized solutions around their content items. We believe the future of collaboration will focus on building and sustaining communities around content, tasks and ideas, and we see hosted entities known as content spaces as a means to supporting ecosystems of users and developers around that content.

To make this case we review the dominant trends in computing that motivate our exploration of new approaches for content-centered collaboration. We offer insights into how certain core problems for users and organizations are being addressed today. We identify challenges that must be conquered before solutions can be delivered, and we outline next steps for research in the area.

Life within Content Spaces: A Vision

1 All are members of the FRACTAL Program, Web Services and Systems Lab (WSSL), HP Labs (Bristol, UK and Vermont, USA)
Collaboration around digital content is on the increase in the business world and beyond. Outside the enterprise, users often are drawn together by common interests to develop content—for example, to plan for a family holiday involving family members in far-flung locations. Inside the workplace, this trend is most pronounced when companies outsource tasks, those tasks often requiring close working relationships between the companies involved. These collaborations grow organically to form communities around the particular content they aim to create.

In previous generations of knowledge work both users and content had to be brought to their tools; work could not happen if the tools were not available; the sharing of tools and content across time and space was difficult; organizations bore the costs of acquiring the tools their employees needed; workers couldn't locate and use the best tools for the job at hand, as determined by them; and workers couldn't span time and space to share, discuss, collaborate or even publish content where necessary, as time and circumstance required. Instead, it was much more likely that workers had to exchange content through email or through rudimentary file sharing; more advanced workers might have been lucky enough to have had access to managed models of sharing with features like version control and logging.

Emerging content-centered work practices shatter such barriers by enabling communities to access items of interest wherever they need to get work done. Seamless, fluid collaboration within these communities is the new imperative that distinguishes the emerging generation of users known as *digital natives.*

Today's workers, using tools adopted from the social software domain, have grown accustomed to posting content of interest to sites that are set up to host conversations about that content; these conversations are the standard currency of the communities that form. The portal-based “spaces” that users create today in services like Facebook or Flickr to converse about content are mere prototypes of the content spaces we envision for tomorrow. We believe that the next generation of work will be conducted in and around content spaces with greatly enhanced capabilities over what is currently possible.

Moving beyond communities centered on content, content spaces will be complete environments where custom sets of functionality are built around content in order to facilitate the tasks that must be done with it. Part of this functionality will replicate today’s conversations and content management features, but will go beyond what is

---

available. In essence, a content space will form a custom environment for working collaboratively with specific content items to achieve a specific end.

In important ways, content spaces will be the antithesis of today’s content management systems and repositories. The isolated operation of current systems makes cross-organization, content-centered collaboration inherently painful, requiring significant overhead: for example, when workers from different organizations need shared access to items of work, they must establish *ad hoc* channels such as email, set up extra-organizational privileges on one company's systems, or arrange for a secure, inter-organizational shared repository. It is rare today to share content in a working environment that seamlessly integrates discussions of the work. And it is even more rare for collaborators to set up such collaborative environments on their own terms, incorporating whatever content and functionality they require.

Today we often need to manually remove a content item from its storage location, perform some function on it, and store the result back to the original location. While some of these processes will involve teams of people, they may also include steps that can be automated. What is needed is a way to coordinate human tasks with a form of simple automation that can be attached to a content space and instructed to act on content within it; we call this an *active behavior*. The functionality offered by an active behavior may be as simple as automatically creating (e.g.) Microsoft Word versions of all imported text documents or as complex as running workflows to collate information from several different organizations working in collaboration. In this way functionality is brought to a content space rather than forcing content to be exported before applications can work on it.

Future workers will build content spaces around content, molding them to fit the job at hand by adding and removing functionality as needed. Functionality will be bundled into active behaviors found by searching the greater community. Sometimes the desired function will be available at zero cost; at other times, a user will subscribe to an advanced behavior, fueling an ecosystem of professional development driven by community needs. Following from this, we see an inversion of the traditional application-content hierarchy through a move to content-centered working—applications are brought to content rather than content to applications.

Each content space can be unique, created by composing a set of behaviors acting on content items. Users will invite internal and external collaborators into this customized environment, assigning appropriate roles and
responsibilities—and thus authority—as easily as adding names to a mailing list. After the group’s work is “complete,” their content space might be destroyed completely—its configuration, including associated active behaviors and role models, being saved to serve as a template for future work or sharing with others—or it might be archived and maintained according to applicable legal requirements. The ability of users and groups to adapt content spaces to their immediate requirements allows shared working environments to be created in a fraction of the time required today, as needed; used, then destroyed, cheaply and simply, fundamentally changing the mechanics of collaborative work.

**Evolving to Hosted Content-Centric Communities**

For many organizations today, content is their end product which they manifest as magazines, market research, advertising and graphics art. For others, content is generated as a side-effect of doing business in the form of invoices, statements, RFPs, images and data sets. In all cases, organizations seek more efficient ways to create, manage, distribute, archive and re-purpose these artifacts.

Several ingredients are necessary to achieve these efficiencies: high quality domain-specific metadata, ideally extracted automatically from the content; tools to effectively model content-centric processes, enabling automation of repetitive steps, while retaining the agility to allow process changes on the fly; and policy-driven storage, providing control over how content is preserved, expressed in terms of performance, accessibility, consistency, encryption and levels of replication. Flexibility will be the key to giving businesses control over all aspects of their content throughout its lifecycle.

Businesses of all sizes are increasingly looking toward the web and cloud-based services to offer facilities previously available only to enterprises with the IT infrastructure available for hosting large on-site applications. The use of these services frees businesses from having to manage their own infrastructure and allows them to

---

concentrate on their key business strengths. Offsite management of their content is the next step along this path, with companies now starting to appear to address this need.\(^4\)

The Internet platforms that power these new delivery models must embrace social software principles, especially the concept of enabling instant community formation around common interests such as specific content items, tasks that they must accomplish or simply ideas that they share. Tomorrow’s users will be born and educated in an era of total connectivity and will possess highly-developed sensibilities about working in the context of community. New applications and services must support this style of working, provide interfaces which allow end users to expose different combinations of content and functionality to suit their needs and provide simple means to integrate themselves with other services.

Motivated by these imperatives, the FRACTAL program at HP Labs has set out to address a number of core challenges that cloud-based services must overcome in order to effectively support content-centric collaborative work, especially collaborations demanding rich content management capabilities. Some of these challenges include:

*How can users collaborate with their peers in new, agile, content-centered ways, unfettered by the limitations of on-premises infrastructure; in ways that leverage an ecosystem of experts within their business domain; and which are consistent with emerging new norms for network-based social interactions?*

*How can developers and service providers be empowered to more rapidly create domain-focused, content-centered applications “in the cloud” that minimize their investment in infrastructure; that leverage the collective skills of an ecosystem of contributors; and thus enable them to focus more of their energy on delivering their unique solutions?*

**Key Trends**

1. **Outsourcing Applications and Services to the Cloud**

\(^4\) For example, NetDocuments ([http://www.netdocuments.com/](http://www.netdocuments.com/)) and SpringCM ([http://www.springcm.com](http://www.springcm.com)) provide these services for the legal and contract management domains, respectively.
A growing number of small- to mid-sized companies are looking to improve their operational efficiency and to access business applications previously only available to large enterprises by using cloud-based application outsourcing. Such incentives must be weighed against today’s uncertainties over outsourced solutions, especially related to shared infrastructure and multi-tenanted applications; concerns range from the obvious “teething pains,” including the reliability and scalability of early providers, to hard technical problems like isolation and security, to unprecedented questions such as the legal status of cloud service providers, application providers and end users.

Not just for mid-markets: One might regard this trend as being restricted to mid-market companies, but a September 2008 IDC survey shows that of the 244 major IT customers surveyed, the current level of cloud services adoption is significant (ranging from 15-26% across seven factors) and is expected to nearly double by 2011. Frank Gens argues that the industry is at a critical time and invokes Geoffrey Moore’s Crossing the Chasm:

...[T]he implication for IT suppliers is very clear: the next three years, as IT cloud services expand from "Early Adopters" to the "Early Majority," is the critical time to develop strong cloud offerings, and play a leadership role in bringing customers, your own organization and your partner ecosystem across the chasm. ...[S]uppliers who fail to seriously contend for a leadership role will be left with decidedly minority share as their reward...  

Increasingly Sophisticated Web-Based Applications: Many of the traditional obstacles to browser-based application development, including a lack of interoperability between different browsers and slow, unwieldy interfaces are vanishing as browsers continue to evolve. Starting with the push toward standardization by WaSP,
and continuing with browser-independent library stacks, browsers have become steadily more interoperable. New Javascript engines are removing speed gaps between browser and desktop programs. Moreover, efforts such as Google Native Code browser extensions, Microsoft Silverlight and Adobe Flex/AIR which allow browsers to run the same code as native applications mean that complex operations such as image manipulations are within reach. This trend will lead to the introduction of rich, collaborative content-editing environments in the browser, with capabilities ranging from simple word-processing to complex CAD.

**Beyond the browser:** Desktop applications are steadily becoming integrated with web-based services. A widely-adopted example in today's enterprise is Microsoft Sharepoint, which users may access from Microsoft Word, Outlook, Windows Explorer, as well as a wide variety of web browsers. This architectural style forms a core of Microsoft’s “Software + Services” initiative. It is clear that, for many hosted applications, enabling a range of interface points beyond the traditional web browser will be vital.

### 2. Collaboration across Organizational Boundaries

Businesses increasingly rely on external partners to help meet their objectives. The trend toward outsourcing non-core competencies means that businesses have increasingly aggregated the efforts of others into their products. IT historically has not supported this type of interaction; in particular, content hidden behind company firewalls has made cross-organizational collaboration a difficult process with high overheads.

An infrastructure that supports content spaces will empower individuals and organizations to organize and manage content distributed across multiple administrative domains, resulting in a view of content managed independently of

---

9 Development has been led by Prototype (http://www.prototypejs.org/) and JQuery (http://jquery.com/).


the original storage repositories. As for content contained within a hosted environment, controlled sharing can occur without requiring movement of content, making sharing both instantaneous and more secure. Content spaces provide building blocks that enable users to build logical collections that control access and support discovery, content retrieval, and preservation. Such functions may be implemented as services that users attach to their content spaces.

**Shareable spaces:** In our vision, originators of content spaces who must partner with colleagues spread across different companies and geographies will invite collaborators “into” their hosted content spaces as easily as they would their colleagues sitting next to them. They will discuss the work which needs to be done, decide on how to carry it out and do the work within the content space, ensuring that the needed context for the work is complete, readily accessible when they need it. They will use third-party applications to manipulate and interact with the data in the content space and when their work is complete the originator of the content space might elect to bring the space back under their sole control. Such ad hoc work patterns are becoming more common as organizations themselves are virtualized, many composed entirely of teleworkers and temporary arrangements with third-party partners.

**Shareable workflows:** A key aspect of this vision is the coordination of work between different organizations. We see shared workflows as enabling this ad hoc collaboration; these are workflows in which participants are spread between organizations. In shared workflows each team member can execute tasks, receive alerts and otherwise monitor progress, but can do so in a highly individualized way that is consistent with the current generation of social applications.

Shared workflows present several challenges that make them difficult to implement in today's on-premises computing environments and not much easier in the current generation of hosted environments, where many of the characteristics (and limitations) of the enterprise have been preserved. Current workflow models invariably require the replication and exchange of content, an inefficient solution that in the best cases can lead to consistency problems and in the worst cases results in lost or leaked content. We therefore believe that content spaces will also be critical building blocks to implementing shared workflows for content manipulation and management across organizational boundaries.

3. **Digital Natives: A New Generation of Users Enters the Workforce**
A new generation of knowledge workers, branded digital natives by commentators, is entering business. These workers are characterized as being completely at ease with technology due to their immersion since birth in a world of digital tools and unrestricted connectivity. Digital natives have already had an impact on collaboration; in particular, their comfort with applying new technologies in unexpected ways has helped popularize the leading social network sites.

The intuition digital natives have for new, malleable, communications-centric tools causes them to bristle at the rigid, constrained domain of “classic” Enterprise IT, which in most cases prefers to dictate the tools employees may use and control access to others—especially those available on the public Internet! Digital natives are willing to rebel against these (perceived) antiquated restrictions to get the job done. Their failure to respect traditional boundaries can have potentially serious consequences as (for example) some use public sites such as Facebook to share content and even collaborate, not appreciating the many conflicts with corporate polices this use generates. New tools that respect business sensibilities while supporting the innovative work practices of digital natives are required.

**Give them spaces of their own:** Content spaces will provide individuals with ways to easily establish a shared context for their work, allowing them to create workspaces that define the content of interest, the people who will gather to do the work and the applications used to work with the content. These content spaces will be both private and public, will support teams spanning organizations as well as working within them, and will integrate existing business processes while serving as catalysts for the development of totally new ones.

### 4. The Social Web

In “Harnessing the Power of Social Applications,” Josh Bernoff and Charlene Li highlight the opportunities inherent in infusing vertically-oriented applications (or spaces) with social software capabilities. The authors argue that with increasing levels of social participation among consumers and the growing sophistication of the underlying technologies, it is now possible to put social applications on an equal footing with other business projects; social

---

applications can deliver measurable progress toward significant, strategic business goals. In their paper Bernoff and Li give examples of ways that companies can harness the power of the *groundswell*\(^\text{14}\) for a variety of objectives including product research and development; marketing; sales; customer support; and operations.

The web consultancy Opteros has argued that in order to meet the heightened expectations of tomorrow’s worker, the enterprise must move toward so-called next-generation Internet applications and services that embody three key principles:\(^\text{15}\):

- **The Culture of Online Communities:** The “participatory nature of Web 2.0” has become entrenched in the minds of many users, and organizations should take advantage of the tools of that culture such as blogs, wikis, tagging, social networking, forums, collaborative filtering in order to establish and create personalized one-to-one relationships with customers as well as to create internal communities for employees to share, collaborate, and “increase the velocity of knowledge within the company”\(^\text{16}\).

- **Rich Interfaces:** Next-generation Internet applications provide a higher level of user experience than the simple HTML pages of old, giving users instant responses to their selections and the ability to configure their own solutions graphically in real time. AJAX and related technologies are being applied to provide richer, more intuitive and more engaging interfaces that create compelling experiences and interactions for users.

- **Service-Oriented Architectures:** User expectations and competition are driving adopters away from single-platform solutions, with their inherent limited ability to quickly roll out new functionality. Most adopters are taking a service-oriented approach to their next-generation Internet applications, making component decisions that can be swapped in and out rather than committing to restrictive, long-term single-platforms.

### Research Challenges and Hard Problems


Researchers must overcome a number of technical challenges before this vision of collaborative content spaces can be fully realized. Our team is particularly interested in the problems of defining, provisioning and interacting with content spaces “in the cloud”; defining, deploying and implementing active behaviors that are bound to content spaces; and the construction of new application frameworks to allow the tech savvy to rapidly create, use and share their own new content-centric applications.

**Content Spaces in the Cloud**

We think content spaces will require a core repository architecture that can capitalize on the virtues of cloud implementation while giving responsible parties sufficient control over and access to “their” content. An important difference from today’s cloud-based persistence models will be the introduction of policy-based automation, which in turn will present challenges in policy expression, management and enforcement in hosted environments. Related to this are the problems inherent in multi-tenanted content management, including isolation, sandboxing, scaling, and usage tracking, which will continue to be an on-going problem for cloud infrastructure service providers.

**Component Services**

A virtue of content spaces will be their potential to foster ecosystems of services bound to the content they contain. The underlying platform must be built on architectural principles that encourage independent providers to extend content spaces with new behaviors and advanced component services. Driven by the need to accommodate accessible, community-driven development models, architects must define appropriate levels of API that encourage development and adoption, with an eye toward base service types and capabilities that define a minimum acceptable level of service that will win over early adopters.\(^{17}\)

**Active behaviors and content spaces**

---

Active behaviors will provide a flexible way to define collaborative processes involving multiple individuals, organizations and component services. We see active behaviors as a form of readily-accessible automation that will enable end users to collaboratively develop, share and evolve shared work patterns over distributed teams. They will include both manual (human) and automatic (service) process steps and will accommodate both structured and *ad hoc* processes. Importantly, active behaviors will enable “off-premises” or third-party process implementations to connect back to particular content and services within organizations where needed.

Active behaviors must be easy to create, customize and share in order to encourage the independent development of a variety of domain-specific applications. Platforms designed to support content spaces must make it easy for adopters and end users to adapt virtually any customization or extension to meet their specific needs, and then to share those customizations with others in the greater community. In general, the platform must provide a way to modularize customizations, aggregate them, and track dependencies in a systematic way across all levels of customization in order to accommodate “trivial” discovery, customization and sharing.

The declaration and activation of active behaviors must be more accessible than approaches based on methodologies such as BPM. Our experiences and the experiences of many others have shown that current approaches to the definition of workflows and workflow-like processes are too difficult for users. The successes of the myExperiment research portal\(^\text{18}\) have helped us to see the potential for easy creation, adoption and invocation of active behaviors, with a minimum of developer intervention, in the context of community.

**Application frameworks for content-centered collaboration**

Our vision of content spaces fundamentally inverts the common definition of an “application.” More precisely, our vision takes the view that applications in the future will be aggregations of behaviors that are bound to and augment content, rather than content being the by-product of applications. In important ways this can be seen as an advancement of the document-centric and “active document” computing models proposed by researchers over the

last decade.\textsuperscript{19} We see this as an exciting and intuitively attractive view of computing, but it raises a number of questions, specifically: what it means to develop and deploy “applications” at various scales; what languages and tools will be required; and what provisions within the infrastructure are required for these applications to scale?

**Realizing the Content Spaces Vision**

During 2008 HP committed itself to a comprehensive cloud strategy known as "Everything as a Service" (EaaS). HP chief technology officer Shane Robison describes this as a future state where everything will be delivered to the end user as a service, from their work life through entertainment to various communities.\textsuperscript{20} HP Labs is making major investments in high-impact research in support of HP's strategic shift to cloud computing and EaaS,\textsuperscript{21} and in support of this strategy the FRACTAL program is engaged in creating a practical implementation of the content spaces vision.

To date, our emphasis has been on understanding the problem; identifying illustrative use cases that highlight the capabilities of content spaces and are compelling to key stakeholders; and developing a functioning conceptual prototype based on an existing open source content management platform to communicate FRACTAL’s content spaces vision.

In the coming months FRACTAL will turn its attention toward experimental implementations of cloud-based services that more directly realize this vision. In particular we will address the “hard problems” outlined in the previous section, with a focus on core repository issues, active behaviors and application frameworks.


We believe content spaces are an exciting new way to think about collaboration that emphasizes building and sustaining communities around content, task and ideas and which will have fundamental implications on society, impacting the way individuals and organizations work and collaborate. We are interested in engaging with others to develop and extend these ideas with us and to experiment with prototyping, applying and commercially exploiting this approach. In the coming months our team will announce a public web presence to support a community of interest centered on the content spaces model. Please contact the authors for more information.

Conclusions

The future of collaboration will be focused on building and sustaining communities around content, and we envision scalable cloud-hosted objects called content spaces that will support ecosystems of users and developers growing around content and content-related tasks and ideas. Content spaces will enable seamless collaboration around items of content, in a manner far beyond the capabilities of today’s tools. A new generation of platform, optimized for the creation and composition of cloud-based services to deliver content spaces to a wide range of users across many domains, will give users, developers and service providers the freedom to build content- and user-centered solutions driven by individual requirements and preferences. Critically, future applications based on content spaces will be aggregations of behaviors that are bound to and augment content, rather than content being the by-product of applications.

Further Reading


