DocCraft

Andrei Plugaru, Cosmin Turtureanu, Florin Rotaru, Radu Alexandru

May 2024

1 Introduction

This case study explores the development and potential impact of **DocCraft**, a collaborative document platform designed without LaTeX support. DocCraft aims to provide a user-friendly interface and robust collaboration tools, targeting non-STEM academics, business professionals and students.

2 Market and Context Analysis

2.1 Global Trends and Needs

Digital Collaboration Growth: The global increase in remote work and online collaboration highlights the need for efficient collaborative tools. As businesses and educational institutions shift towards more flexible working and learning environments, the demand for tools like DocCraft that facilitate real-time collaboration and seamless document sharing has surged.

Education and Research: The rise of online education and the demand for collaborative tools in academic research, especially in non-STEM fields, underscores DocCraft's relevance. Students and educators need platforms that allow them to co-author papers, share resources and provide feedback in a centralized and accessible manner.

Professional Use Cases: Businesses worldwide are increasingly adopting cloud-based solutions for document management and collaboration. Professionals require tools that enable them to work together on projects, manage versions and maintain a clear record of document changes and contributions.

2.2 National Context

Adoption of Technology: In Romania, the adoption of collaborative tools within the education and professional sectors has been growing steadily. Local universities and businesses are recognizing the benefits of digital transformation and the need for tools that support remote and hybrid working models.

Local Needs: Specific local challenges include language preferences and regulatory considerations. DocCraft addresses these by offering multilingual support and ensuring compliance with local data protection regulations. Integration with national digital infrastructures, such as government databases and educational portals, is also a priority.

Government and Institutional Support: Government initiatives promoting digital literacy and online collaboration tools are driving the adoption of platforms like DocCraft. Institutional support from universities and professional organizations further boosts its relevance and uptake.

3 Implementation and Impact

3.1 Development and Deployment

Phased Development:

1. **Phase 1**: Initial development and beta testing, focusing on core features such as real-time collaboration, document editing and user authentication.
2. **Phase 2**: Feature enhancement based on user feedback, including advanced formatting options, integration with third-party tools and mobile app development.
3. **Phase 3**: Full deployment and scaling, targeting educational institutions and professional organizations for broader adoption.

User Testing and Feedback: During beta testing, feedback from educators, students and professionals is collected to refine features and improve the user experience. Regular updates and feature rollouts are planned based on this feedback.

3.2 National Impact

Education Sector: DocCraft enhances collaborative learning and academic writing in local educational institutions by providing a platform where students and educators can co-author documents, share resources and provide real-time feedback. This fosters a more interactive and engaging learning environment.

Professional Sector: In the professional sector, DocCraft improves productivity and document management by allowing teams to collaborate seamlessly on reports, proposals and other documents. Version control and commenting features ensure that all contributions are tracked and managed efficiently.

3.3 International Impact

Global Reach: DocCraft aims to attract an international user base with features catering to global users, such as multilingual support and compliance with international data protection standards.

Cross-Border Collaboration: The platform facilitates international research collaborations and professional projects by providing a centralized space for document creation and management. This is particularly beneficial for multinational companies and research teams working across different time zones and locations.

4 Comparative Analysis

Comparison to Other WYSIWYG Editors

Strengths and Weaknesses: Compared to other WYSIWYG editors like Google Docs, Microsoft Word Online and Notion, DocCraft offers several unique advantages:

- **Simplicity and Accessibility**: DocCraft is designed to be intuitive and easy to use, with a focus on non-technical users.
- **Real-Time Collaboration**: Like Google Docs and Microsoft Word Online, DocCraft supports real-time editing and collaboration but also includes features tailored specifically for academic and professional use.
- **Multilingual Support**: DocCraft offers extensive language support, making it suitable for international users and cross-border projects.
- **Customizable Workflows**: Unlike more rigid platforms, DocCraft allows users to customize their workflows and document management processes to fit their specific needs.

User Demographics: DocCraft serves non-STEM academics, business professionals and students who require a collaborative document platform that is simple to use and does not rely on LaTeX. This contrasts with platforms like Overleaf, which are more suited to STEM fields and users familiar with LaTeX.

5 Case Studies and Testimonials

5.1 National Case Study

Education: DocCraft was implemented in a national university, where it significantly improved collaboration and document creation for faculty and students. Feedback highlighted its ease of use and the effectiveness of real-time editing and commenting features in enhancing academic writing and group projects.

Business: A local business deployed DocCraft to enhance document workflows and team collaboration. The platform streamlined the process of creating, editing and reviewing documents, leading to increased productivity and better project management.

5.2 International Case Study

Academic Research: An international research project utilized DocCraft to facilitate cross-border collaboration. Researchers from multiple countries were able to co-author papers and share resources in real-time, overcoming geographical barriers and improving the efficiency of their collaborative efforts.

Professional Network: A multinational company adopted DocCraft to streamline document management across different offices. The platform's multilingual support and customizable workflows allowed teams to work together seamlessly, regardless of location or language.

6 Financial Analysis

6.1 Budget and Funding

Development Costs: The initial investment required for development, testing and deployment includes costs for software development, server infrastructure and marketing.

Operational Costs: Ongoing expenses include maintenance, updates, customer support and server hosting.

6.2 Revenue Model

Subscription Plans: DocCraft offers different pricing tiers based on features and user needs, from free basic accounts to premium subscriptions with advanced features.

Enterprise Solutions: Custom packages are available for large organizations and institutions, providing tailored solutions and dedicated support.

7 Conclusion for Market and Context Analysis

7.1 Summary of Benefits

Innovative Solution: DocCraft fills a gap in the market by offering a non-LaTeX collaborative document tool that is simple, accessible and suited for a wide range of users, from academics to professionals.

National and International Impact: The platform benefits local and global users by enhancing collaboration, improving productivity and supporting diverse workflows and languages.

7.2 Future Prospects

Expansion Plans: Future developments include adding new features, expanding mobile app capabilities and enhancing integration with other digital tools and platforms.

Call to Action: Stakeholders are encouraged to engage with DocCraft through investment, pilot programs, or partnership opportunities to support its growth and further development.

8 Technical details about existing solutions

Other solutions similar to ours are **Google Docs**, **Microsoft OneDrive** and **Overleaf**.

8.1 Google Docs

Architecture:

- *back-end*: Built on top of Google's cloud infrastructure, leveraging services like Google Cloud Storage, Google Compute Engine and Google App Engine.
- *front-end*: Uses a combination of JavaScript, AngularJS and other proprietary Google libraries for real-time collaboration.
- *database*: Google's proprietary Spanner and Bigtable for handling vast amounts of data with high availability and low latency.
- *real-time collaboration*: Utilizes WebRTC and Firebase for synchronization and real-time updates.

Technologies:

- *languages*: Primarily JavaScript for front-end; Java, Python and Go for back-end services.
- *frameworks*: AngularJS for the web interface; gRPC for inter-service communication.
- *APIs*: Google Drive API for file storage and access; Google Docs API for document manipulation.
- *data storage*: Google Cloud Storage for file storage; Spanner for transactional data.

Marketing Approaches:

- *integration*: Deep integration with Google Workspace (formerly G Suite), offering a seamless experience with Gmail, Google Drive and other Google services.
- *freemium model*: Basic features are free, while advanced features are part of a paid subscription.
- *enterprise solutions*: Targeting businesses and educational institutions with Google Workspace plans that offer enhanced security, admin controls and additional storage.

8.2 Microsoft OneDrive

Architecture:

- *back-end*: Hosted on Microsoft Azure, leveraging services like Azure Blob Storage, Azure Functions and Azure SQL Database.
- *front-end*: Uses React.js for dynamic user interfaces and SignalR for real-time collaboration.
- *database*: Azure Cosmos DB for distributed data storage and Azure SQL for relational data.
- *real-time collaboration*: SignalR and WebSockets for real-time updates and co-authoring.

Technologies:

- *languages*: JavaScript/TypeScript for front-end; C#, .NET and Azure services for back-end.
- *frameworks*: React.js for the web client; .NET Core for back-end services.
- *APIs*: Microsoft Graph API for accessing Microsoft 365 services; OneDrive API for file storage and management.
- *data storage*: Azure Blob Storage for document storage; Cosmos DB and SQL Database for other data needs.

Marketing Approaches:

- *bundling*: Bundled with Microsoft 365 subscriptions, providing access to Office Online along with desktop versions of Office apps.
- *enterprise focus*: Strong focus on enterprise customers with comprehensive IT admin tools, security features and compliance solutions.
- *cross-platform availability*: Ensuring availability across multiple platforms (Windows, macOS, iOS android) to cater to a wide user base.

8.3 Overleaf

Architecture:

- *back-end*: Overleaf is built on a scalable cloud infrastructure, primarily using AWS for hosting and storage. It uses a microservices architecture to manage different functionalities.
- *front-end*: The front-end is built using modern web technologies, including React.js, to provide a responsive and interactive user experience.
- *database*: MongoDB is used for handling document storage and metadata, while Redis is employed for caching and real-time data synchronization.

- *real-time collaboration*: WebSockets are used to facilitate real-time collaboration, allowing multiple users to edit the same document simultaneously with instant updates.

Technologies:

- *languages*: JavaScript (React.js) for the front-end; Node.js for the back-end services.
- *frameworks*: React.js for the front-end UI; Express.js for back-end services.
- *APIs*: Overleaf API for document creation, collaboration and integration with other services.
- *data storage*: AWS S3 for storing document files; MongoDB for database management; Redis for caching and real-time operations.

Marketing Approaches:

- *target audience*: Overleaf primarily targets academic and research communities, including students, researchers and institutions that require collaborative LaTeX editing.
- *freemium model*: Offers a free tier with basic features suitable for individual users and small collaborations. Paid plans provide additional features such as increased storage, more collaborators per project and advanced version control.
- *educational partnerships*: Overleaf has partnerships with many universities and research institutions, providing institutional licenses that offer premium features to students and faculty.
- *integration*: Integrates with various reference management tools like Zotero, Mendeley and reference libraries, as well as GitHub for version control, making it a comprehensive tool for researchers.
- *community engagement*: Active engagement with the LaTeX community through blogs, tutorials, webinars and direct support to promote usage and address user needs.

9 Overview of technologies

The application is structured in 2 components: the front-end and the back-end, which exposes a REST API.

9.1 Front-end

For the front-end component, **React.js** was chosen.
The reasons are further presented. Firstly, its component-based architecture promotes modular and reusable code, facilitating easier maintenance and scalability of large-scale applications. Additionally, React's virtual DOM efficiently updates only the necessary components when data changes, enhancing performance by minimizing DOM manipulation. Furthermore, React's extensive ecosystem, supported by a vast community and numerous third-party libraries, offers developers a plethora of tools and resources to streamline development processes and address various needs. Lastly, its adoption by major companies and widespread usage in industry projects ensure robust support and continual improvement, making React a reliable choice for building modern web applications.

9.2 Back-end

For the back-end component **Node.js with Express**. It was chosen for a few reasons. Firstly, Node.js itself is known for its asynchronous, event-driven architecture, which makes it highly efficient for handling concurrent requests, making it suitable for building scalable applications. Express.js, being a minimalist and flexible web application framework for Node.js, provides a lightweight and unopinionated structure that allows developers to customize their application's architecture according to their specific requirements. Furthermore, the active community surrounding Node.js and Express.js ensures continuous updates, support and a plethora of resources for developers, contributing to the frameworks' stability and longevity.

9.3 Cloud technologies

A few cloud technologies have also been used in order to facilitate the implementation of the requirements. The main cloud provider is Google. It was chosen as it is one of the biggest cloud providers and it offers both professional and beginner friendly services.
Also, some specific services from Microsoft Azure were used as found better support there.

1. **App engine** It is used for the deployment of both the front-end and back-end components.

2. **Cloud Functions** They are used as triggers for async operations. Specifically, we used for pushing some notification when new file permissions are added.

3. **Cloud SQL** It is used for a SQL Server database.

4. **Firebase** It is used for authenticating users.

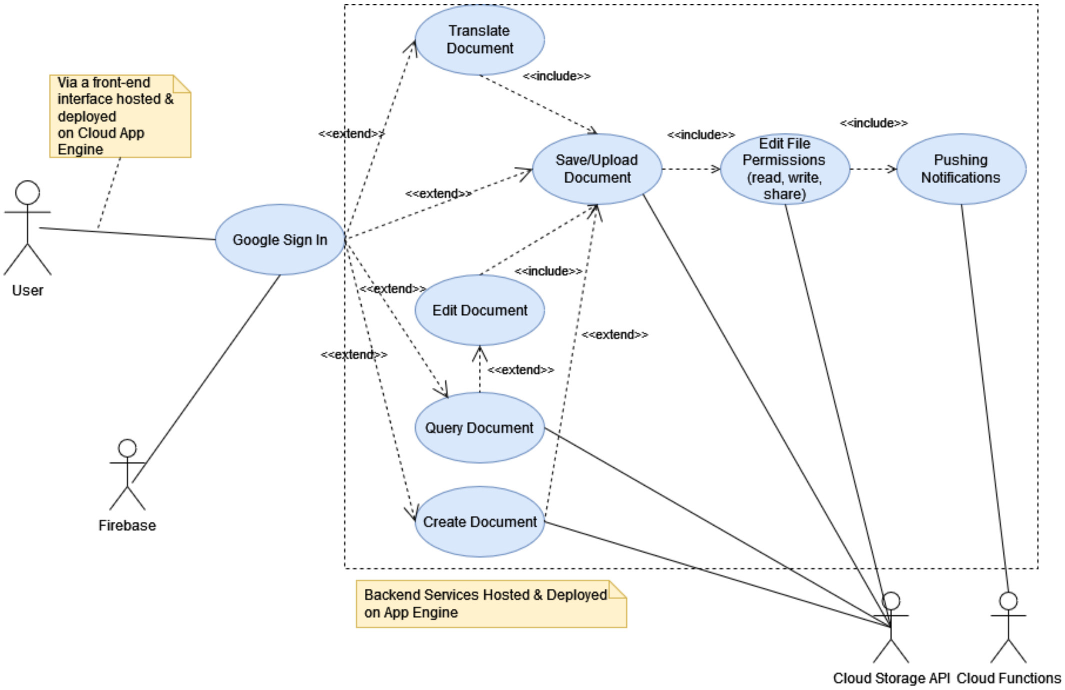
5. **Cloud Storage** It is used for storing the actual file contents.

Other used services are Google Pub/Sub, Azure AI Translator, Autoscaling, Monitoring

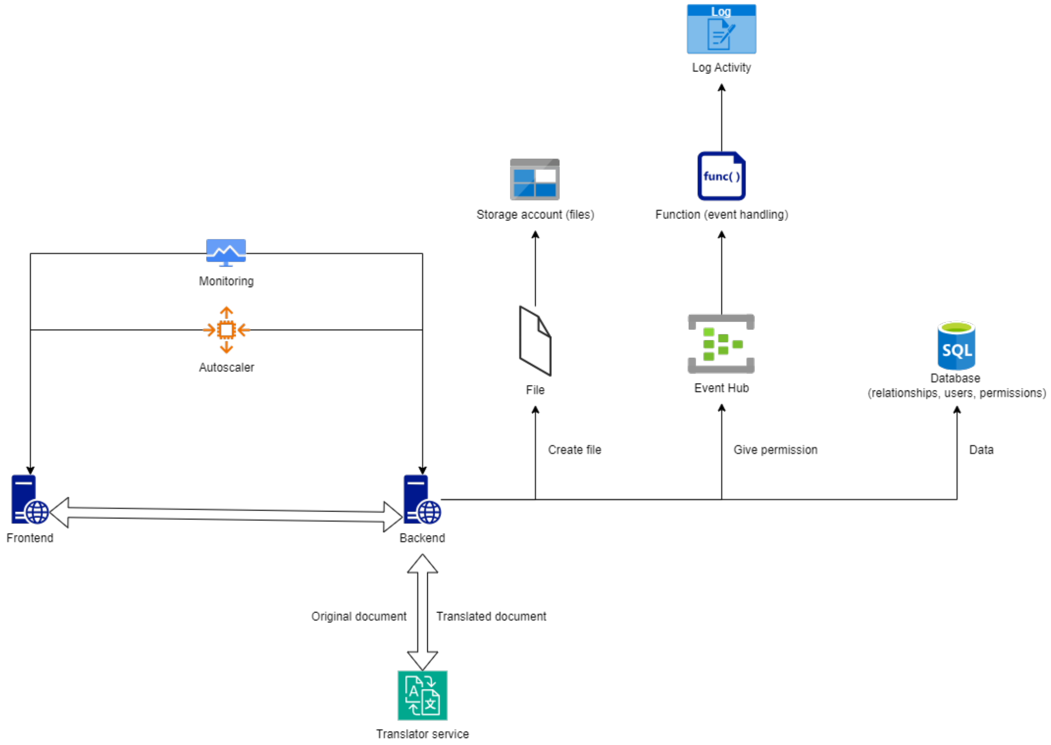
10 Business Model Canvas (see next page)

<table>
<tr><td colspan="6"><b><i>DocCraft Business Model Canvas</i></b></td></tr>
<tr><td><b>Key Partners</b></td><td><b>Key Activities</b></td><td colspan="2"><b>Value Propositions</b></td><td><b>Customer Relationships</b></td><td><b>Customer segments</b></td></tr>
<tr><td rowspan="3">- <i>Cloud Service Providers</i>: Companies like AWS, Google Cloud, or Azure for reliable cloud infrastructure.</td><td>- <i>Product Development</i>: Continuous improvement and updating of the platform with new features.
- <i>Data Management</i>: Ensuring data security, privacy and efficient cloud storagemanagement.</td><td colspan="2" rowspan="3">- <i>Ease of Use</i>: User-friendly interface for seamless document creation and editing.
- <i>Collaboration</i>: Real-time editing and sharing features that facilitate teamwork.
- <i>Cloud Storage</i>: Secure, cloud-based storage for easy access to documents from anywhere.
- <i>Security</i>: Advanced security features to ensure document privacy and protection.</td><td>- <i>Customer Support</i>: 24/7 customer support via chat, email and phone.</td><td rowspan="3">- <i>Individual Users</i>: Students, professionals and freelancers who need a reliable document creation and sharing platform. Small and Medium Businesses (SMBs): Teams and businesses requiring collaborative tools for document management.
- <i>Educational Institutions</i>: Schools and universities looking for platforms to facilitate document collaboration among students and faculties.
- <i>Non-Profit Organizations</i>: Organizations needing affordable document management and collaboration tools.</td></tr>
<tr><td><b>Key Resources</b></td><td><b>Channels</b></td></tr>
<tr><td>- <i>Technology Infrastructure</i>: Servers, cloud storage and data centers to support the platform.
- <i>Development Team</i>: Software engineers, UX/UI designers and quality assurance professionals.
- <i>Customer Support Team</i>: Staff for providing 24/7 support to users.</td><td>- <i>Online Platform</i>: Website</td></tr>
<tr><td colspan="3"><b>Cost Structure</b></td><td colspan="3"><b>Revenue Streams</b></td></tr>
<tr><td colspan="3">- <i>Technology and Infrastructure</i>: Costs for servers, cloud storage and data centers.
- <i>Customer Support</i>: Expenses related to maintaining a customer support team and infrastructure.</td><td colspan="3">- <i>Freemium Model</i>: Basic features available for free with premium plans offering advanced features.
- <i>Advertisements</i>: Optional, non-intrusive ads for free users.</td></tr>
</table>

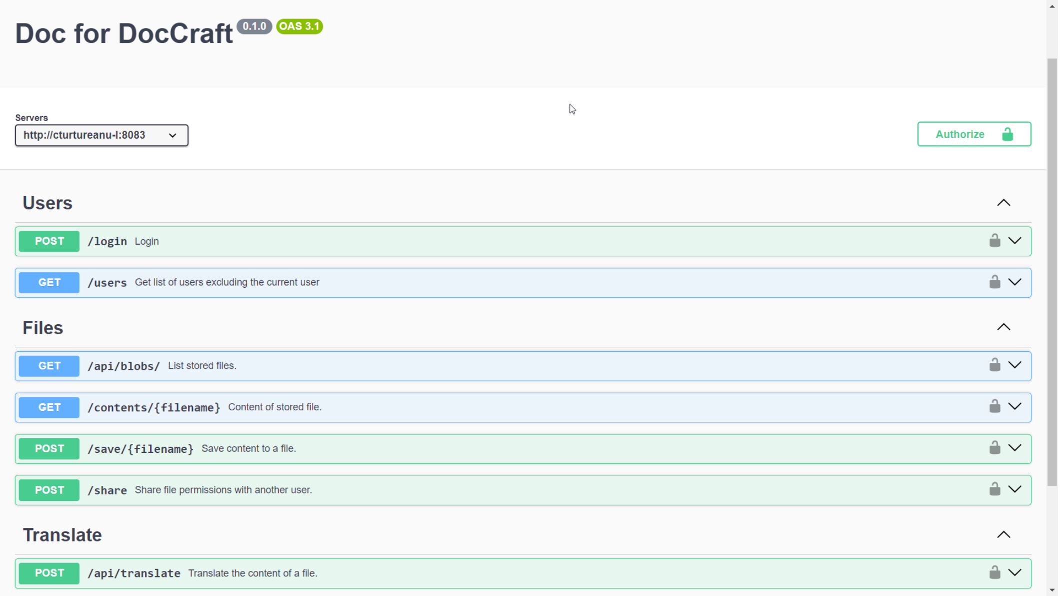
11 Use case diagram



12 Architectural Diagram



13 OpenAPI specification



Doc for DocCraft 0.1.0 OAS 3.1

Servers
http://cturtureanu-l:8083

Authorize

Users

POST /login Login

GET /users Get list of users excluding the current user

Files

GET /api/blobs/ List stored files.

GET /contents/{filename} Content of stored file.

POST /save/{filename} Save content to a file.

POST /share Share file permissions with another user.

Translate

POST /api/translate Translate the content of a file.