DataSafe

Senior Design Team Contract

University of Cincinnati

College of Education, Criminal Justice and Human Services

School of Information Technology

Andrew Drabek
Jonathan Heasley
Akshat Rojora
Drew Miluk

Table of Contents

Intent

The following contract was written and agreed upon by Andrew, Jonathan ,Akshat and Drew. The contract provides expectations, objectives, and results for developing the application DataSafe.

The contract is effective for all team members participating in the Senior Design Capstone class series in the 2025-2026 academic year.

Senior Design Contract

Project Summary

DataSafe is a comprehensive data loss protection tool with the goal of helping companies keep their data secure across all of their platforms. It aims to manage sensitive information across many environments. It attempts to while maintaining an affordable cost model aiming at small to medium customers.

1. **Confirm and Validate Access**: DataSafe provides a way for companies to regularly review and confirm who has access to their data, ensuring that only authorized personnel can interact with sensitive resources. This helps mitigate risks associated with outdated or excessive permissions.
2. **Dashboarding and Reporting**: DataSafe will include powerful visualizations that allow companies to see their security posture in relation to many frameworks at a quick glance. There also will be ways for companies to download and export reports for them to share internally.
3. **Identify and Address Poor Access Controls**: DataSafe will scan for vulnerabilities and misconfigurations, such as permissions or unprotected endpoints.
4. **Sensitive Data Identification**: DataSafe will provide companies with a way to identify sensitive information in their storage locations.

Problem Statement

Many small businesses don’t think data loss prevention (DLP) tools are worth it because they’re so expensive. According to Berecki (2019), “Digital security issues threaten businesses of all sizes; while larger companies have more data to steal, smaller businesses have less secure networks, thus usually they become quick and easy targets in the eyes of cybercriminals.” This highlights the fact that all businesses are at risk, especially when it comes to data. While tools like DLP can help mitigate that risk, they are extremely difficult for smaller businesses to afford, often making them easier targets, as mentioned earlier. This difficulty arises because of the resource drain that DLP tools pose.

Dilmegani (2024) outlines cost comparisons for DLP tools, noting that the cheapest options start at around $50 per month per user, while the most expensive can reach up to $135 per month per user. While these tools can be beneficial, their cost often places them out of reach for small businesses.

It’s not just money that prevents organizations from adopting these tools—it’s also the time and employee hours required. Staff at Proofpoint (2018) explain that “many organizations who invest in DLPs quickly discover that these tools are difficult and time-consuming to deploy. It can take quite a bit of manpower to set them up before it’s possible to use them to their full extent.” This adds to the problem: DLP tools are not only expensive but also challenging to implement and maintain.

We need to make DLP tools more accessible by getting rid of things like minimum users, lowering the costs to get started, making them easier to maintain and being more flexible overall. If we can do that, smaller businesses will be more likely to adopt these tools, which will help them improve their security now and set themselves up for success later.

Solution

Develop an affordable and user-friendly DLP tool that encourages companies to perform scans frequently and or automatically. This encourages companies to keep data protection in their mind as well as making a tool for them that is more accessible and cost friendly. By prioritizing ease of maintenance and usability, the tool will empower organizations of all sizes to adopt automated data security practices without the common barriers of high costs and complexity.

In addition, the solution will provide achievable goals, guiding users to identify the types of data they possess and outlining clear steps to remediate potential issues. This functionality ensures smaller companies can better understand and address their data security needs while not bogging them down with constant maintenance and other non-priority tasks. By streamlining the process, the tool allows smaller businesses to confidently use it while focusing on scaling and growing their operations.

Contact Information

Team Member	Degree + Track	Email	Phone Number OR Other Contact Info
Andrew Drabek	BSIT – Software Dev	drabekam@mail.uc.edu	513-968-8457
Jonathan Heasley	BSIT - CyberSecurity & Software Dev	heaslejn@mail.uc.edu	614-464-7873
Drew Miluk	BSIT – Software Dev	milukaj@mail.uc.edu	216-777-9433
Akshat Rajora	BSIT- Cybersecurity	Rajoraat@mail.uc.edu	202-361-2948

Project Source

Andrew had the idea of making a data loss prevention (DLP) tool. He worked as a systems engineer at a startup and noticed that DLP tools never came up in tooling conversations. Curious about why, he asked some of his coworkers, many of whom are either CISOs or at a CISO level (the CEO was an ex CISO and they have a CISO and other CISO contractors). The answer was always something close to - "They're not worth it for small companies."
The reasoning made sense at first—DLP tools are really expensive, and while they're nice to have, the return just doesn't seem worth it for smaller businesses. But the more Andrew thought about it, the more he realized that this was the wrong way to look at it. It's not that DLP tools aren't worth it; it's that they should be, and they need to be affordable.
Companies store tons of data—whether it's internal or external—and there's just so much of it. We have to think about who has access and what kind of data is being stored. Sure, you can use manual controls, but as companies grow, having a DLP tool in place can make a huge difference in improving security posture. With poorly managed credentials, whether it's on assignment or employees mishandling them, becoming a bigger issue, it's time to focus on making DLP tools something that small companies can actually use.

Project Objectives/Goals

- Create a comprehensive DLP tool that allows users to scan file storage and know what data is being stored and if there are any access issues.
    - Making it easier for startups and smaller companies to focus on file security.
- Integrate ways for users to have complete visibility into their current security posture.
- Create methods for users to be alerted.
    - This allows startups to set automated scans and then alert them so that they do not need to dedicate users to this tool.

Team Members and Responsibilities

There are three main areas that the team is going to focus on as far as assigning work. Those areas are the frontend (external facing items), backend(codebase), and the database. While the team might have adopted a full stack ideology, certain users will be focused or more attuned to those items.

Frontend - Drew
Backend - Andrew
Database – Jonathan
Infrastructure/security - Akshat

Project Scope

While we realize that we might not be able to get all these items accomplished, here are the overall goals for DataSafe.

1. Create a basic file storage scanning service

    a. Must be able to get credential access
    a. Must be able to see what type of data is stored
    a. Must be able to scan multiple file stores - S3, Storage Account etc.
        . Get statistics from all the items above but will not store file contents

2. Create dashboarding solution
    a. Using the information gathered from above create a dashboard
    a. Show all information types stored and percentages of each
    a. Show large risk areas
    a. Show any immediate action items

1. Create alerting and reporting
    a. Set up way to alert users no major items
    a. Allow users to create a single exportable report

Quick Project Timeline

Task #	*Task Name*	*objectives*	*Duration – 2 Semesters ~ 32 weeks*	*Start Date*	*End Date*
1	Research and Requirements map out	Finalize frameworks and technologies that might need to be changed. Confirm open source licensing for technologies used.	2 Weeks	September 15	September 29
2	Wireframes design and backend planning	Create wireframes as a plan for the frontend. Create a plan for the backend and what services are going to be broken up. Plan out any backend	4 Weeks	September 29	October 20

		environments and what technologies will be needed to host the app. Plan repo structure.			
3	Basic File Scanning and data classification with start dashboarding	Create the planned feature. This one will allow for core functionality and for users to input sources to be scanned and then have a ui that shows what was found. We will need to create a way to run automated and manual scans	7 weeks	October 20	December 8
4	Alerting implementation	Create the planned feature. This one will allow for alerting when there is a serious violation found.	3 weeks	December 8	December 29
5	Frontend dashboarding finalization	Create the planned feature. This one will be	8 weeks	December 29	February 23

		allowing users to see what issues they might have and how many alerts we have given them.			
6	Final internal review and tests	Test and find any issues. Fix the bugs that are found.	3 Weeks	February 23	March 16
7	Documentation and presentation	Document projects and create the presentation.	4 Weeks	March 16	April 13

Technologies Used

While designing the application, we will focus on a **microservice approach**. This involves utilizing technologies like Docker containers and Docker clusters to enable modular and scalable architecture. Each company will have its own dedicated database, hosted by a single database host. This ensures data isolation and enhances security. Using robust permission controls, we can effectively manage who has access to which database, maintaining strict data segregation.

- **Frontend Codebase**: React with JavaScript, Auth0 for authentication(There is a free tier)
- **Container Platform**: Docker
- **Backend**: Python with Fast API and Pandas and pyshiny(if needed)
- **Database**: MSSQL

Ethical Considerations

The primary ethical concern for this type of application revolves around the security of accessing the platform and the nature of the information being stored. According to IEEE (2016), “Individuals should uphold the following principles in order to follow ethical guidelines: Do no harm while seeking to improve the quality of life for all people. Establish accountability practices. Respect confidentiality.”

This principle will guide our approach. We will create an application that respects users' privacy and does not store information that could potentially harm them in the event of a breach. Additionally, we will ensure that user information is not used for purposes beyond what the customer has explicitly agreed to. By adhering to best practices and employing appropriate tools, we aim to uphold these ethical standards.

To address these concerns, we will implement the following measures:

- **Role-Based Access Control (RBAC):** We will enforce strong RBAC policies to ensure users have appropriate permissions and establish robust user management systems.
- **Avoid Storing Sensitive Content:** Sensitive user file contents will not be stored. Instead, we will scan documents to extract and store only relevant metadata, minimizing the need to retain original files or sensitive content.
- **Secure Tokenization:** Customer information will not be embedded in tokens. Instead, we will use identifiers such as GUIDs and other mechanisms that conceal user identities.

While these are just a start, these measures will help ensure our application adheres to ethical and privacy-centric design principles, fostering trust and minimizing potential risks.

Legal Considerations

The creation and management of a Data Loss Prevention (DLP) tool come with significant legal challenges, primarily due to varying regulations across jurisdictions. According to ISACA (2024), "Regulations are not internationally harmonized, causing severe complications (especially between the United States and the European Union) on a cross-border basis, which is the rule rather than the exception in modern business." This difference creates obstacles in designing a DLP solution that can consistently comply with global requirements.

A DLP tool faces heightened scrutiny due to its responsibility for managing sensitive information. As a company dealing directly with data protection, any breach not only raises ethical concerns but also triggers legal requirements. These requirements include disclosing the breach, identifying affected parties, and reporting the incident according to specific regulatory requirements. Failing to meet these requirements can lead to significant penalties, legal liabilities, and long-term reputational damage.

As Borner (2019) highlights, "A data breach can also damage stock prices for large multinationals. For example, the ICO fine imposed on British Airways saw shares slide down by 2 percent. But the reputational damage incurred can have a long-term effect on consumer loyalty and reduced trust in the marketplace." This demonstrates how legal consequences of a breach extend beyond fines and into areas that affect business sustainability.

To mitigate these risks, our DLP tool will:

1. **Ensure Regulatory Compliance:** Make sure that we are in compliance with whatever country we are working in.
2. **Be Proactive:** Monitor and track application usage and calls. Allowing us to stay on top of incidents and prevent incidents before they happen.
3. **Encrypted Data Management:** Implement encryption standards and secure handling to reduce exposure.
4. **Promote Accountability:** Include logging and audit trails to provide clear logging of what is happening in the background.

Team Rules

1. It is expected that all users will come to an agreed meeting. The team will need to give at least 72-hour notice before scheduling the meeting.
    a. If all members agree to the meeting it is expected that they will be there
2. Each team member will be respectful of others and their opinions.
3. If there are problems the team will do its best to solve them internally, but members are encouraged to reach out to the professor if they need to.
4. All members are expected to keep up with their tasks and do research on their own as needed. If the team uses a management platform it is expected that they will use it.

Team Signatures:

Signature: _____Andrew Drabek__________________

Date: __________09/14/2025_________________

Signature: ______Akshat Rajora_________________
Date: ________9/14/2025___________________

Signature: __________Jonathan Heasley_______________

Date: ____________9/14/2025_________________

Signature: ___________Drew Miluk______________
Date: _____________9/14/2025________________

Signature: _______Samuel Bricking_____________
Date: _______9/22/2025_________________

References

24, F., & Staff, P. (2024, October 22). *4 reasons data loss prevention tools aren't cutting it: Proofpoint us*. Proofpoint. https://www.proofpoint.com/us/blog/insider-threat-management/4-reasons-data-loss-prevention-tools-arent-cutting-it

Achieving Ethical Protection of Data Privacy. ISACA. (n.d.-a). https://www.isaca.org/resources/news-and-trends/industry-news/2024/achieving-ethical-protection-of-data-privacy

Berecki, P. byBeata. (2020, December 15). *3 reasons why SMBS should use DLP*. Endpoint Protector Blog. https://www.endpointprotector.com/blog/top-3-reasons-to-use-dlp-for-smbs/

Borner, P. (2024, August 26). *The impact of a data breach: Consequences of data and security breaches*. The Data Privacy Group. https://thedataprivacygroup.com/blog/2019-9-17-data-breach-the-legal-implications/?utm_source=chatgpt.com#

Coos, P. byAndrada. (2020, December 10). *Debunking the top 3 myths about DLP*. Endpoint Protector Blog. https://www.endpointprotector.com/blog/debunking-the-top-3-myths-about-dlp/

DLP pricing: Save by comparing top 3 vendors. AIMultiple. (2024, October 16). https://research.aimultiple.com/dlp-pricing/

An ethical approach to data privacy protection. ISACA. (n.d.-b). https://www.isaca.org/resources/isaca-journal/issues/2016/volume-6/an-ethical-approach-to-data-privacy-protection?utm_source=chatgpt.com

Ethical issues related to data privacy and security: Why we must balance ethical and legal requirements in the connected world - IEEE digital privacy. (n.d.). https://digitalprivacy.ieee.org/publications/topics/ethical-issues-related-to-data-privacy-and-security-why-we-must-balance-ethical-and-legal-requirements-in-the-connected-world

FASTAPI. FastAPI. (n.d.). https://fastapi.tiangolo.com/

Ibm. (2024, November 1). *What is Data Loss Prevention (DLP)?*. IBM. https://www.ibm.com/topics/data-loss-prevention

Pandas. (2018). Python Data Analysis Library. Pydata.org. https://pandas.pydata.org/

React. React Blog RSS. (n.d.). https://react.dev/

What is Data Loss Prevention (DLP)?. Microsoft Security. (n.d.). https://www.microsoft.com/en-us/security/business/security-101/what-is-data-loss-prevention-dlp