APHERIS

Privacy and security for federated ML

How to assess privacy & security in readiness for federated machine learning & analytics

Introduction

Data is the lifeblood of an organization, yet the value of that data is often left untapped. Due to the inherent value and sensitivity of data, it must be protected.

As model architectures become increasingly commoditized, data becomes an organization's key differentiator. However, businesses need to safeguard their data assets and IP while leveraging it for ML.

Apheris enables governed, private, and secure access to data, helping connect multiple federated data sets, without moving or copying data to a central location, for the purpose of ML or advanced analytics.

In this document, we will explore the importance of proper data governance and share steps you can take to assess the privacy and security of your organization's data.

Why Data Governance and Regulation Matters

“AI development is happening at breakneck speed, underpinned by massive investments from Big Tech companies such as Microsoft, Google and Amazon. The consequence of this escalating AI arms race is that commercial applications are reaching consumers without sufficient information on the data upon which these AI models have been trained.

For regulators around the world, this has set alarm bells ringing.

AI regulation need not be an obstruction to innovation. In order to build and sustain public trust – which will ultimately be beneficial to the industry as it seeks to implement greater AI adoption into technology – comprehensive but flexible regulation is the only way to build a safe and transparent system that aligns with our human values.”

Robin Röhm
Co-Founder, Apheris

Original publication in
The AI Journal

Keeping up with Regulations

For organizations operating globally, data privacy laws such as GDPR, CCPA, HIPAA are increasingly complex and expected to change. Under GDPR's principle of integrity and confidentiality, for example, personal data should never move outside of its sovereignty.

Keeping data with the data custodian is the best starting point and because our Compute Gateway is deployed in the same environment as the data, it doesn't need to move and remains isolated from other party's data.

Adopting a federated data infrastructure is a good place to start in complying with any data privacy regulations now and in the future.

Power your data infrastructure with federated machine learning and analytics

Any data, any size, anywhere

5 Safes: a framework for considering privacy & security in federated ML

The "5 Safes" is a framework for managing access to sensitive data. It's flexible and can be customized for specific needs of the organization.

The framework is security-focused and covers all key aspects of data access and use, including projects, people, settings, data, and outputs. When implemented well, it can ensure sensitive data is handled systematically, and because of its flexibility, it can be adapted to different types of data and use cases.

The 5 Safes underpin best practice when it comes to keeping data secure and private.

Safe projects

What is the purpose of using the data and why is it needed?
Ensure projects are legal, ethical, and aligned between parties.

Safe people

Who needs access to data and what are their roles?
Govern activities on the platform with fine-grained identity and access management using roles and asset policies.

Safe settings

Where will the data be accessed and how will it be stored?
Is the location and environment in which data will be used secure and data storage and transmission done in a way that reduces risk of unauthorized access?

Safe data

What type of data is being used and how sensitive is it?
Ensure data is properly classified, the sensitivity understood, and that it's protected by appropriate security measures.

Safe outputs

What type of results will be produced and how will they be used or distributed?
Leverage controls in asset policies so that computation results are non-disclosive, are versioned, and can be reproduced at a later point.

Assessing privacy & security in readiness for federated learning & analytics

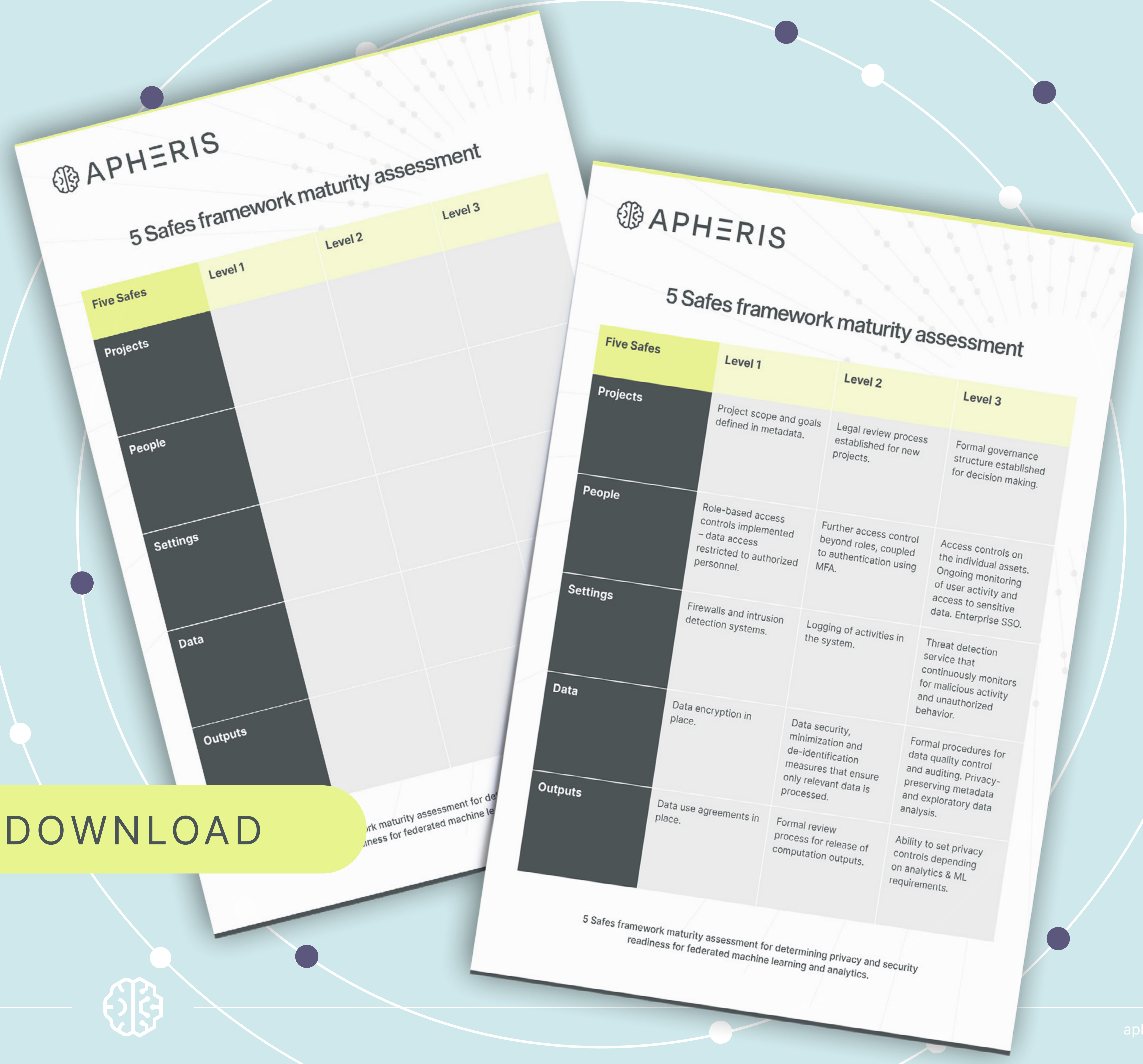
Our 5 Safes Maturity Assessment is a useful framework for assessing capability and readiness for implementing federated learning.

It provides a set of criteria that can be used to assess your current state, identify gaps, and develop a roadmap for improving capabilities ready for successfully implementing federated learning.

Each step is additive – level 1 provides the foundation for 2; level 2 builds on level 1 and together are the basis for 3.

Take a look at following example, then download a copy for you to fill out yourself.

DOWNLOAD



APHERIS

5 Safes framework maturity assessment

Five Safes	Level 1	Level 2	Level 3
Projects			
People			
Settings			
Data			
Outputs			

rk maturity assessment for de
ness for federated machine le

APHERIS

5 Safes framework maturity assessment

Five Safes	Level 1	Level 2	Level 3
Projects	Project scope and goals defined in metadata.	Legal review process established for new projects.	Formal governance structure established for decision making.
People	Role-based access controls implemented – data access restricted to authorized personnel.	Further access control beyond roles, coupled to authentication using MFA.	Access controls on the individual assets. Ongoing monitoring of user activity and access to sensitive data. Enterprise SSO.
Settings	Firewalls and intrusion detection systems.	Logging of activities in the system.	Threat detection service that continuously monitors for malicious activity and unauthorized behavior.
Data	Data encryption in place.	Data security, minimization and de-identification measures that ensure only relevant data is processed.	Formal procedures for data quality control and auditing. Privacy-preserving metadata and exploratory data analysis.
Outputs	Data use agreements in place.	Formal review process for release of computation outputs.	Ability to set privacy controls depending on analytics & ML requirements.

5 Safes framework maturity assessment for determining privacy and security readiness for federated machine learning and analytics.

In Conclusion

A comprehensive data governance framework such as the “Five Safes” should be implemented to strike the right balance between privacy, security, and collaboration.

By implementing this framework, companies can ensure that their data is handled in a responsible and ethical manner, with appropriate measures in place to protect the privacy and security of their users’ data.

By taking a proactive approach to data governance and privacy, technical leaders can foster trust and confidence among their users, while unlocking the full potential of federated learning and analytics.

Learn more about privacy & security with Apheris.
Get in touch at apheris.com

APHERIS

Contact us

Let’s start a conversation. Get in touch with our specialists to learn more.

General enquiries:
info@apheris.com

Media enquiries:
press@apheris.com

apheris.com