

IEEE Guide for Framework for Trustworthy Federated Machine Learning

IEEE Computer Society

Developed by the
Artificial Intelligence Standards Committee

IEEE Std 3187™-2024

STANDARDS

IEEE Guide for Framework for Trustworthy Federated Machine Learning

Developed by the

Artificial Intelligence Standards Committee
of the
IEEE Computer Society

Approved 26 September 2024

IEEE SA Standards Board

Abstract: The development and application of federated machine learning are facing the critical challenges of balancing the tradeoff among privacy, security, performance, and efficiency, how to realize supervision covering the whole life cycle, and how to get explainable results. Thus, trustworthy federated machine learning is proposed to solve the above problem. In this standard, a general view of framework for trustworthy federated machine learning is provided in four parts: a principle in trustworthy federated machine learning, requirements from the perspective of different principles and different federated machine learning participants, and methods to realize trustworthy federated machine learning. Also provided is guidance on how trustworthy federated machine learning is used in various scenarios.

Keywords: federated machine learning, framework, IEEE 3187™, machine learning, principle, trustworthy federated machine learning

The Institute of Electrical and Electronics Engineers, Inc.
3 Park Avenue, New York, NY 10016-5997, USA

Copyright © 2024 by The Institute of Electrical and Electronics Engineers, Inc.
All rights reserved. Published 19 December 2024. Printed in the United States of America.

IEEE is a registered trademark in the U.S. Patent & Trademark Office, owned by The Institute of Electrical and Electronics Engineers, Incorporated.

PDF: ISBN 979-8-8557-1471-5 STD27493
Print: ISBN 979-8-8557-1472-2 STDPD27493

*IEEE prohibits discrimination, harassment, and bullying.
For more information, visit <https://www.ieee.org/about/corporate/governance/p9-26.html>.
No part of this publication may be reproduced in any form, in an electronic retrieval system or otherwise, without the prior written permission of the publisher.*

Important Notices and Disclaimers Concerning IEEE Standards Documents

IEEE Standards documents are made available for use subject to important notices and legal disclaimers. These notices and disclaimers, or a reference to this page (<https://standards.ieee.org/ipr/disclaimers.html>), appear in all IEEE standards and may be found under the heading “Important Notices and Disclaimers Concerning IEEE Standards Documents.”

Notice and Disclaimer of Liability Concerning the Use of IEEE Standards Documents

IEEE Standards documents are developed within IEEE Societies and subcommittees of IEEE Standards Association (IEEE SA) Board of Governors. IEEE develops its standards through an accredited consensus development process, which brings together volunteers representing varied viewpoints and interests to achieve the final product. IEEE standards are documents developed by volunteers with scientific, academic, and industry-based expertise in technical working groups. Volunteers involved in technical working groups are not necessarily members of IEEE or IEEE SA and participate without compensation from IEEE. While IEEE administers the process and establishes rules to promote fairness in the consensus development process, IEEE does not independently evaluate, test, or verify the accuracy of any of the information or the soundness of any judgments contained in its standards.

IEEE makes no warranties or representations concerning its standards, and expressly disclaims all warranties, express or implied, concerning all standards, including but not limited to the warranties of merchantability, fitness for a particular purpose and non-infringement. IEEE Standards documents do not guarantee safety, security, health, or environmental protection, or compliance with law, or guarantee against interference with or from other devices or networks. In addition, IEEE does not warrant or represent that the use of the material contained in its standards is free from patent infringement. IEEE Standards documents are supplied “AS IS” and “WITH ALL FAULTS.”

Use of an IEEE standard is wholly voluntary. The existence of an IEEE standard does not imply that there are no other ways to produce, test, measure, purchase, market, or provide other goods and services related to the scope of the IEEE standard. Furthermore, the viewpoint expressed at the time a standard is approved and issued is subject to change brought about through developments in the state of the art and comments received from users of the standard.

In publishing and making its standards available, IEEE is not suggesting or rendering professional or other services for, or on behalf of, any person or entity, nor is IEEE undertaking to perform any duty owed by any other person or entity to another. Any person utilizing any IEEE Standards document should rely upon their own independent judgment in the exercise of reasonable care in any given circumstances or, as appropriate, seek the advice of a competent professional in determining the appropriateness of a given IEEE standard.

IN NO EVENT SHALL IEEE BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO: THE NEED TO PROCURE SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE PUBLICATION, USE OF, OR RELIANCE UPON ANY STANDARD, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE AND REGARDLESS OF WHETHER SUCH DAMAGE WAS FORESEEABLE.

Translations

The IEEE consensus balloting process involves the review of documents in English only. In the event that an IEEE standard is translated, only the English language version published by IEEE is the approved IEEE standard.

Use by artificial intelligence systems

In no event shall material in any IEEE Standards documents be used for the purpose of creating, training, enhancing, developing, maintaining, or contributing to any artificial intelligence systems without the express, written consent of IEEE SA in advance. “Artificial intelligence” refers to any software, application, or other system that uses artificial intelligence, machine learning, or similar technologies, to analyze, train, process, or generate content. Requests for consent can be submitted using the Contact Us form.

Official statements

A statement, written or oral, that is not processed in accordance with the IEEE SA Standards Board Operations Manual is not, and shall not be considered or inferred to be, the official position of IEEE or any of its committees and shall not be considered to be, or be relied upon as, a formal position of IEEE or IEEE SA. At lectures, symposia, seminars, or educational courses, an individual presenting information on IEEE standards shall make it clear that the presenter’s views should be considered the personal views of that individual rather than the formal position of IEEE, IEEE SA, the Standards Committee, or the Working Group. Statements made by volunteers may not represent the formal position of their employer(s) or affiliation(s). News releases about IEEE standards issued by entities other than IEEE SA should be considered the view of the entity issuing the release rather than the formal position of IEEE or IEEE SA.

Comments on standards

Comments for revision of IEEE Standards documents are welcome from any interested party, regardless of membership affiliation with IEEE or IEEE SA. However, **IEEE does not provide interpretations, consulting information, or advice pertaining to IEEE Standards documents.**

Suggestions for changes in documents should be in the form of a proposed change of text, together with appropriate supporting comments. Since IEEE standards represent a consensus of concerned interests, it is important that any responses to comments and questions also receive the concurrence of a balance of interests. For this reason, IEEE and the members of its Societies and subcommittees of the IEEE SA Board of Governors are not able to provide an instant response to comments or questions, except in those cases where the matter has previously been addressed. For the same reason, IEEE does not respond to interpretation requests. Any person who would like to participate in evaluating comments or revisions to an IEEE standard is welcome to join the relevant IEEE SA working group. You can indicate interest in a working group using the Interests tab in the Manage Profile and Interests area of the [IEEE SA myProject system](#).¹ An IEEE Account is needed to access the application.

Comments on standards should be submitted using the [Contact Us](#) form.²

Laws and regulations

Users of IEEE Standards documents should consult all applicable laws and regulations. Compliance with the provisions of any IEEE Standards document does not constitute compliance to any applicable regulatory

¹Available at: <https://development.standards.ieee.org/myproject-web/public/view.html#landing>.

²Available at: <https://standards.ieee.org/about/contact/>.

requirements. Implementers of the standard are responsible for observing or referring to the applicable regulatory requirements. IEEE does not, by the publication of its standards, intend to urge action that is not in compliance with applicable laws, and these documents may not be construed as doing so.

Data privacy

Users of IEEE Standards documents should evaluate the standards for considerations of data privacy and data ownership in the context of assessing and using the standards in compliance with applicable laws and regulations.

Copyrights

IEEE draft and approved standards are copyrighted by IEEE under US and international copyright laws. They are made available by IEEE and are adopted for a wide variety of both public and private uses. These include both use by reference, in laws and regulations, and use in private self-regulation, standardization, and the promotion of engineering practices and methods. By making these documents available for use and adoption by public authorities and private users, neither IEEE nor its licensors waive any rights in copyright to the documents.

Photocopies

Subject to payment of the appropriate licensing fees, IEEE will grant users a limited, non-exclusive license to photocopy portions of any individual standard for company or organizational internal use or individual, non-commercial use only. To arrange for payment of licensing fees, please contact Copyright Clearance Center, Customer Service, 222 Rosewood Drive, Danvers, MA 01923 USA; +1 978 750 8400; <https://www.copyright.com/>. Permission to photocopy portions of any individual standard for educational classroom use can also be obtained through the Copyright Clearance Center.

Updating of IEEE Standards documents

Users of IEEE Standards documents should be aware that these documents may be superseded at any time by the issuance of new editions or may be amended from time to time through the issuance of amendments, corrigenda, or errata. An official IEEE document at any point in time consists of the current edition of the document together with any amendments, corrigenda, or errata then in effect.

Every IEEE standard is subjected to review at least every 10 years. When a document is more than 10 years old and has not undergone a revision process, it is reasonable to conclude that its contents, although still of some value, do not wholly reflect the present state of the art. Users are cautioned to check to determine that they have the latest edition of any IEEE standard.

In order to determine whether a given document is the current edition and whether it has been amended through the issuance of amendments, corrigenda, or errata, visit [IEEE Xplore](#) or [contact IEEE](#).³ For more information about the IEEE SA or IEEE's standards development process, visit the IEEE SA Website.

Errata

Errata, if any, for all IEEE standards can be accessed on the [IEEE SA Website](#).⁴ Search for standard number and year of approval to access the web page of the published standard. Errata links are located under the Additional

³Available at: <https://ieeexplore.ieee.org/browse/standards/collection/ieee>.

⁴Available at: <https://standards.ieee.org/standard/index.html>.

Resources Details section. Errata are also available in [IEEE Xplore](#). Users are encouraged to periodically check for errata.

Patents

IEEE standards are developed in compliance with the [IEEE SA Patent Policy](#).⁵

Attention is called to the possibility that implementation of this standard may require use of subject matter covered by patent rights. By publication of this standard, no position is taken by the IEEE with respect to the existence or validity of any patent rights in connection therewith. If a patent holder or patent applicant has filed a statement of assurance via an Accepted Letter of Assurance, then the statement is listed on the IEEE SA Website at <https://standards.ieee.org/about/sasb/patcom/patents.html>. Letters of Assurance may indicate whether the Submitter is willing or unwilling to grant licenses under patent rights without compensation or under reasonable rates, with reasonable terms and conditions that are demonstrably free of any unfair discrimination to applicants desiring to obtain such licenses.

Essential Patent Claims may exist for which a Letter of Assurance has not been received. The IEEE is not responsible for identifying Essential Patent Claims for which a license may be required, for conducting inquiries into the legal validity or scope of Patents Claims, or determining whether any licensing terms or conditions provided in connection with submission of a Letter of Assurance, if any, or in any licensing agreements are reasonable or non-discriminatory. Users of this standard are expressly advised that determination of the validity of any patent rights, and the risk of infringement of such rights, is entirely their own responsibility. Further information may be obtained from the IEEE Standards Association.

IMPORTANT NOTICE

Technologies, application of technologies, and recommended procedures in various industries evolve over time. The IEEE standards development process allows participants to review developments in industries, technologies, and practices, and to determine what, if any, updates should be made to the IEEE standard. During this evolution, the technologies and recommendations in IEEE standards may be implemented in ways not foreseen during the standard's development. IEEE standards development activities consider research and information presented to the standards development group in developing any safety recommendations. Other information about safety practices, changes in technology or technology implementation, or impact by peripheral systems also may be pertinent to safety considerations during implementation of the standard. Implementers and users of IEEE Standards documents are responsible for determining and complying with all appropriate safety, security, environmental, health, data privacy, and interference protection practices and all applicable laws and regulations.

⁵Available at: <https://standards.ieee.org/about/sasb/patcom/materials.html>.

Participants

At the time this guide was completed, the FTML Working Group had the following membership:

Zuping Wu, *Chair*
Shuangping Huang, *Vice Chair*

Yang Bian	Xiaoyun Jia	Huajie Shen
Tai Chen	Jiewu Leng	Lin Sun
Xiaofeng Chen	Huifan Li	Qi Sun
Yazheng Chen	Kepeng Li	Qiuyue Wang
Liangkun Cui	Lianzhu Li	Wanwan Wang
Di Dong	Xiangxue Li	Zuping Wu
Lan Du	Xiaohong Li	Yunlong Xie
Lian Fan	Ximeng Liu	Hao Xu
Lixin Fan	Yang Liu	Jing Xu
Xiaoliang Fan	Yonggang Luo	Qiang Yang
Xuehui Hu	Limeng Ma	Tianya Yang
Cong Huang	Yining Meng	Bo Yu
Cuiting Huang	Chao Pei	Nianyin Zeng
Shuangping Huang	Yanhong Pu	Jian Zhang
Yimin Huang	Jinxian Qu	Peiyu Zheng

The following members of the individual Standards Association balloting group voted on this guide. Balloters may have voted for approval, disapproval, or abstention.

Abhishek Alladi	T. Leopold	Walter Struppler
Boon Chong Ang	Yongfu Li	Qi Sun
Pieter Botman	Haiying Lu	Lisa Ward
Zhiman Chen	Yonggang Luo	Yan Yan
Diego Chiozzi	Palak Mathur	Zhen Yang
Madhukar Gaganam	Joanna Olszewska	Xiaobo Yu
Louis Gullo	Bansi Patel	Yu Yuan
Marco Hernandez	George Percivall	Janusz Zalewski
Werner Hoelzl	Edwin Puerta	Nianyin Zeng
Piotr Karocki	R. K. Rannow	Zhen Zhang
	Jhony Sembiring	

When the IEEE SA Standards Board approved this guide on 26 September 2024, it had the following membership:

David J. Law, *Chair*
Jon Walter Rosdahl, *Vice Chair*
Gary Hoffman, *Past Chair*
Alpesh Shah, *Secretary*

Sara R. Biyabani
Ted Burse
Stephen Dukes
Doug Edwards
J. Travis Griffith
Guido R. Hiertz
Ronald W. Hotchkiss

Hao Hu
Yousef Kimiagar
Joseph L. Koepfinger*
Howard Li
Xiaohui Liu
John Haiying Lu
Kevin W. Lu
Hiroshi Mano

Paul Nikolich
Robby Robson
Lei Wang
F. Keith Waters
Sha Wei
Philip B. Winston
Don Wright

*Member Emeritus

Introduction

This introduction is not part of IEEE Std 3187-2024, IEEE Guide for Framework for Trustworthy Federated Machine Learning.

Federated machine learning is one of the key technologies in privacy computing, which has attracted more and more attention from the industry. The implementation of relevant policies in privacy computing are accelerating the commercialization and technology iteration of federated machine learning. Generally, federated machine learning is used as an effective privacy-preserving technology by continuously optimizing the model to achieve better performance. However, the application and commercialization of federated machine learning are facing the critical challenge about how to balance the trade-off among privacy/security, performance, and efficiency. Besides, due to the black-box property of federated machine learning, people are increasingly concerned about whether the supervision covers the whole life cycle and whether they can get explainable results.

Thus, trustworthy federated machine learning is proposed to optimize multiple objectives and meet the needs of users, regulators, and other aspects. This standard proceeds on the assumption that all legal rights and obligations, ethical principles, and values remain mandatory and must be duly observed. That is, this standard mainly focuses on the technical level of trustworthy federated machine learning, including the principles of trustworthy federated machine learning, the requirements corresponding to the different principles and different user roles, and how federated machine learning and other privacy computing technologies work together. The guideline of how trustworthy federated machine learning is used in various scenarios to build a unified solution required by the industry is also provided.

Contents

1. Overview	11
1.1 Scope	11
1.2 Purpose	11
1.3 Word usage	11
2. Normative references	12
3. Definitions, acronyms, and abbreviations	12
3.1 Definitions	12
3.2 Acronyms and abbreviations	13
4. Principles in trustworthy federated machine learning	13
4.1 Overview	13
4.2 Privacy and security	13
4.3 Complexity	14
4.4 Utility	14
4.5 Regulatory	14
4.6 Confidence	14
5. Framework of trustworthy federated machine learning	14
6. Aims of trustworthy federated machine learning	16
6.1 From the perspective of different principles	16
6.2 From the perspective of different user roles	18
7. Methods to realized T-FML	20
7.1 Overview	20
7.2 Blockchain based T-FML	21
7.3 TEE-based T-FML	23
7.4 MPC-based T-FML	25
7.5 Personalized federated machine learning	27
7.6 Edge federated machine learning	29
7.7 Other methods to realize trustworthy federated machine learning	32
Annex A (informative) Requirements at the system level	37
Annex B (informative) Mathematical description of confidence	38
Annex C (informative) Scenarios: Mobile edge network	39
Annex D (informative) Scenarios: Bank anti-fraud model	41
Annex E (informative) Scenarios: Joint risk control	43
Annex F (informative) Scenarios: Automotive application	45
Annex G (informative) Bibliography	47

IEEE Guide for Framework for Trustworthy Federated Machine Learning

1. Overview

1.1 Scope

This document provides a reference framework for trustworthy federated machine learning, including the principles of trustworthy federated machine learning, requirements for different roles and principles of trustworthy federated machine learning, and several technologies to realize trustworthy federated machine learning. It also lists some scenarios where trustworthy federated machine learning can be applied.

1.2 Purpose

The purpose of this guide is to provide a credible, practical, and controllable solution guidance for trustworthy federated machine learning and other privacy computing applications.

1.3 Word usage

The word *shall* indicates mandatory requirements strictly to be followed in order to conform to the standard and from which no deviation is permitted (*shall* equals *is required to*).^{6,7}

The word *should* indicates that among several possibilities one is recommended as particularly suitable, without mentioning or excluding others; or that a certain course of action is preferred but not necessarily required (*should* equals *is recommended that*).

The word *may* is used to indicate a course of action permissible within the limits of the standard (*may* equals *is permitted to*).

The word *can* is used for statements of possibility and capability, whether material, physical, or causal (*can* equals *is able to*).

⁶The use of the word *must* is deprecated and cannot be used when stating mandatory requirements; *must* is used only to describe unavoidable situations.

⁷The use of *will* is deprecated and cannot be used when stating mandatory requirements; *will* is only used in statements of fact.

2. Normative references

The following referenced documents are indispensable for the application of this document (i.e., they must be understood and used, so each referenced document is cited in text and its relationship to this document is explained). For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments or corrigenda) applies.

IEEE Std 1633™-2016, IEEE Recommended Practice on Software Reliability.

IEEE Std 2894™-2023, IEEE Guide for an Architectural Framework Explainable Artificial Intelligence.

IEEE Std 2986™-2023, IEEE Recommended Practice for Privacy and Security for Federated Machine Learning.

IEEE Std 3652.1™-2020, IEEE Guide for Architectural Framework and Application of Federated Machine Learning.

3. Definitions, acronyms, and abbreviations

3.1 Definitions

For the purposes of this document, the following terms and definitions apply. The *IEEE Standards Dictionary Online* should be consulted for terms not defined in this clause.⁸

blockchain: A distributed ledger of data records, which is jointly maintained by multiple parties with certain cryptography to help ensure the transmission and access security, data storage consistency and non-manipulation.

complexity: The model training cost without database, which is measured by communication and computing overheads, storage, and training time.

confidence: The probability of the federated machine learning system achieving a goal (such as utility), or the level of user's certainty in the federated machine learning system achieving this goal.

edge federated machine learning: A type of machine learning that fully utilizes the computing resources of terminal and can perform more extensive and efficient machine learning while protecting privacy.

federated machine learning model: The result of the model-training process of a federated machine learning system. The learned model can be used to make certain machine-learning inference tasks on new data, e.g., classification, recognition, prediction, and recommendation⁴.

personalized federated machine learning: A technology that combines federated machine learning and personalized models based on heterogeneous data from each participant. The impact of data heterogeneity is reduced and high-quality personalized models for each participant can be obtained.

secure multi-party computation: A computing protocol used for collaborative computing among multiple participants while protecting their privacy and data security. The encrypted calculations are performed through the execution protocol, ultimately obtaining the expected computational results.

trusted execution environment: A security zone constructed by software and hardware methods on a data computing platform, can help ensure that the data loaded inside the security zone are protected.

⁸*IEEE Standards Dictionary Online* is available at: <http://dictionary.ieee.org>. An IEEE account is required for access to the dictionary, and one can be created at no charge on the dictionary sign-in page.

trustworthy federated machine learning: A specific multi-party distributed federated machine learning paradigm, with core characteristics as privacy, security, low complexity, high utility, regulatory, and confidence.

trustworthy federated machine learning model training: A trustworthy federated machine learning process under multiple given machine learning tasks can achieve a balance among privacy, security, complexity, and utility.

utility: The Trustworthy federated machine learning is the net proceeds of a federated machine learning system, which is measured by the difference between the algorithm performance and the cost. The algorithm performance is a comprehensive indicators of accuracy, consistency, and flexibility.

3.2 Acronyms and abbreviations

AI	artificial intelligence
D2D	device to device
E-FML	edge federated machine learning
FML	federated machine learning
IID	independent identically distribution
MEC	mobile edge computing
MPC	secure multi-party computation
Non-IID	non-independent identically distribution
P-FML	personalized federated machine learning
QoS	quality of service
RAN	radio access network
TEE	trusted execution environment
T-FML	trustworthy federated machine learning
UPF	user plane function

4. Principles in trustworthy federated machine learning

4.1 Overview

This clause sets out the foundation of trustworthy federated machine learning(T-FML), reflected by five principles that should be adhered to. The basic principles of federated machine learning (FML) stated in IEEE Std 3652.1™-2020 are followed, and only the enhanced features are mentioned.⁹

4.2 Privacy and security

The privacy disclosure and security risks can be quantified and assessed when using T-FML. Participants would know the extent to which their privacy would be disclosed and the level of their privacy disclosure. It means that the privacy and security are variable and that participants can trust the T-FML when using it. Detailed assessment methods can be referred to IEEE Std 2986™-2023, which is not included in this guide.

⁹Information on references can be found in [Clause 2](#).

4.3 Complexity

The overhead of privacy-preserving computing is controllable, which includes communication and computational efficiency. The adoption of cryptographic protection mechanisms and the distributed computing mode will bring about multiple rounds of information exchange among participants, which results in a certain amount of efficiency reduction. Complexity then includes simplifying the communication process, controlling the overall T-FML time and improving the calculation efficiency. It is a tradeoff between privacy and security as well as performance and efficiency.

4.4 Utility

The utility of T-FML is the comprehensive requirement of accuracy, consistency, and flexibility. The model is flexible to support various aggregation algorithms, manage data set heterogeneity, and personalize models. When privacy-preserving technology and highly heterogeneous data are used, the decreased accuracy of the global model is constrained to some degree. Adopting protection mechanisms and non-independent and identically distributed (Non-IID) data may lead to some degree of utility loss. The trained model or the decisions made by the model is consistent every time with the same data set and the same machine learning algorithm. Then the model utility of T-FML helps ensure that it can meet the requirements of different application scenarios and needs.

NOTE—For system-level definitions, please refer to [Annex A](#).¹⁰

4.5 Regulatory

It also means traceability and accountability. That is, the model and training data of T-FML can be traced back. The supervision of T-FML model should cover the whole life cycle.

4.6 Confidence

The model in T-FML, the results after using T-FML, the decision-making in T-FML and the whole T-FML process can be explained. It can also reflect the degree to which participants can understand the cause of a decision produced by a T-FML system. Due to the black-box property of FML system, assessment methods are needed to investigate qualitative and quantitative evaluations of explainability. Detailed properties of explainability can be referred to IEEE Std 2894™-2023, and reliability can be referred to IEEE Std 1633™-2016, which are not included in this guide.

NOTE—The mathematical description is detailed in [Annex B](#).

5. Framework of trustworthy federated machine learning

The framework of T-FML is described in [Figure 1](#).

¹⁰Notes in text, tables, and figures of a standard are given for information only and do not contain requirements needed to implement this standard.