CE, CS, Incentive

CE (Contribution Evaluation)
- 사례: LOO, SV, 신성국/성균관대
- 이슈: Combination of Aggregation에 대한 계산량 부담
- 이슈: Evaluation에 사용하는 data set의 적절성

CS (Client Selection)
- CE (Contribution Evaluation) 을 사용해서 다음 라운드에 참여할 Client 선별 (cross-device)
- CE (Contribution Evaluation) 을 사용해서 weighted average 로 aggregation (cross-silo)
- 사례: Oort, FedBalancer : Loss 기반 Evaluation 이 정확/유효 할까?
- 사례: McDuff : https://github.com/Kwangkee/rPPG/blob/main/rPPG%40DanielMcduff.md#fl-rppg, CE (Contribution Evaluation) 보완 필요

Incentive (CE를 사용해서 Reward 제공)
- https://github.com/Kwangkee/FL/blob/main/FL%40Incentive.md

CE, CS, Incentive

Contract Theory Based Incentive Mechanism for Federated Learning in Health CrowdSensing, https://ieeexplore.ieee.org/abstract/document/9932890

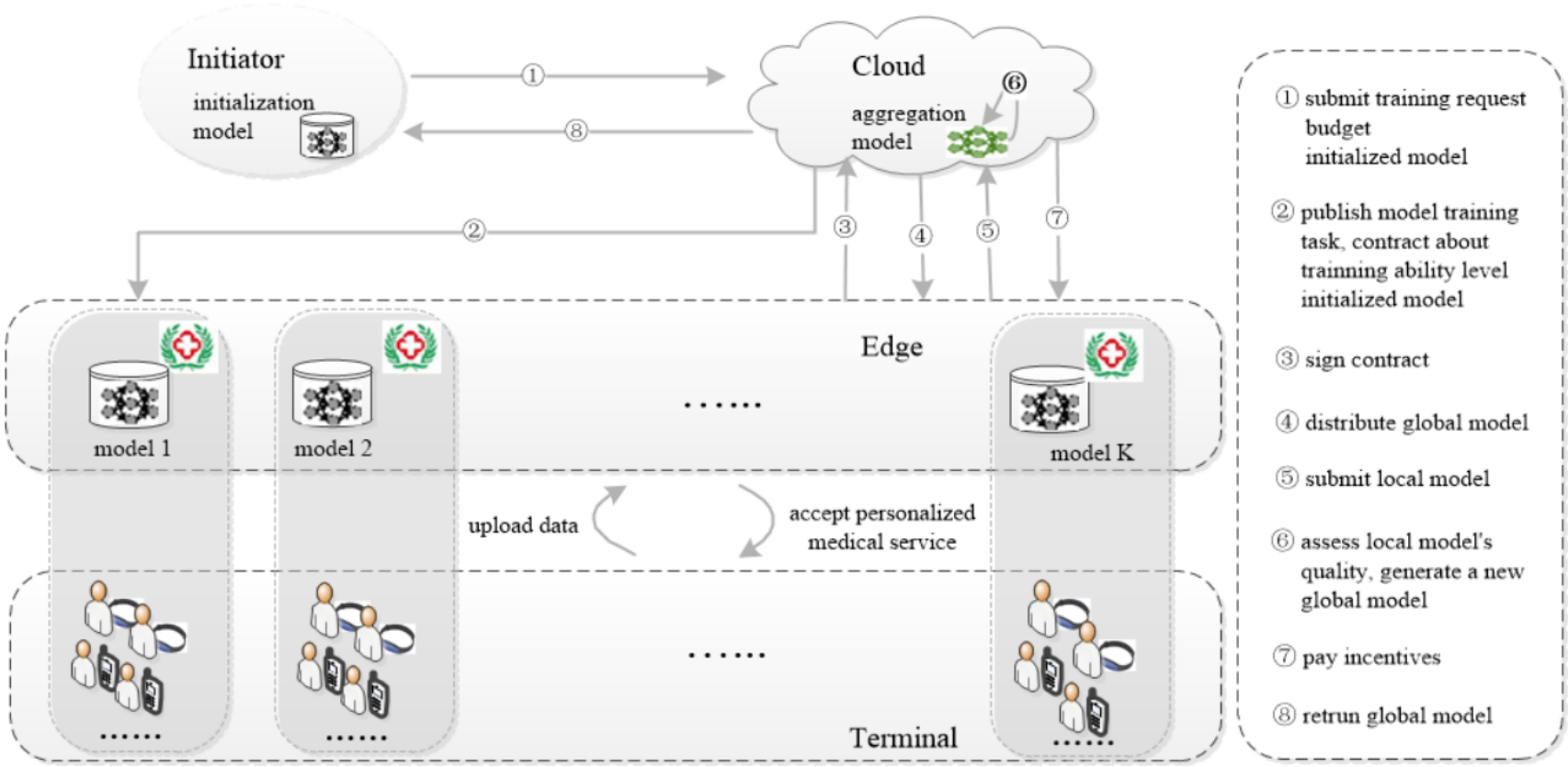
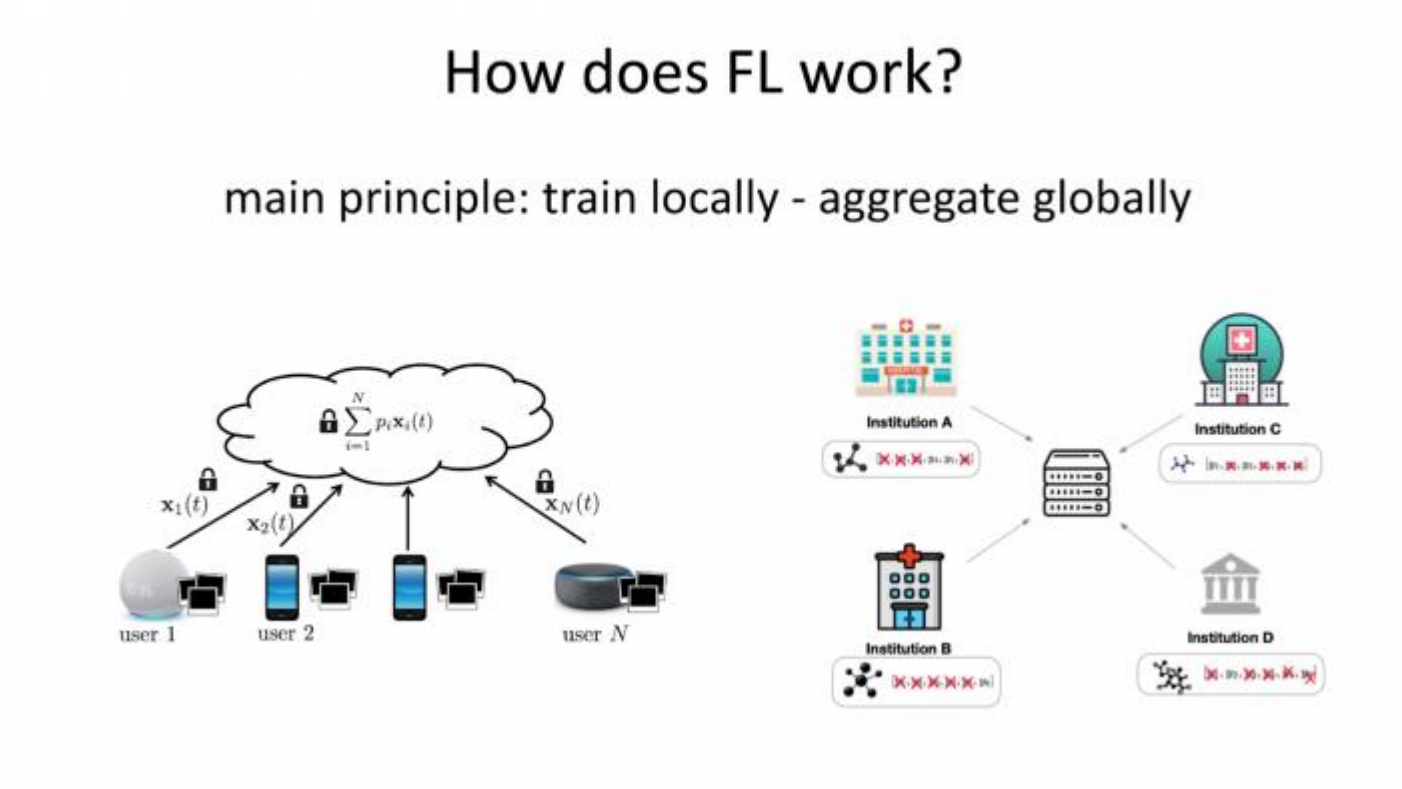


Fig. 1. The typical research scenarios of Federated Learning in HCS

FedML

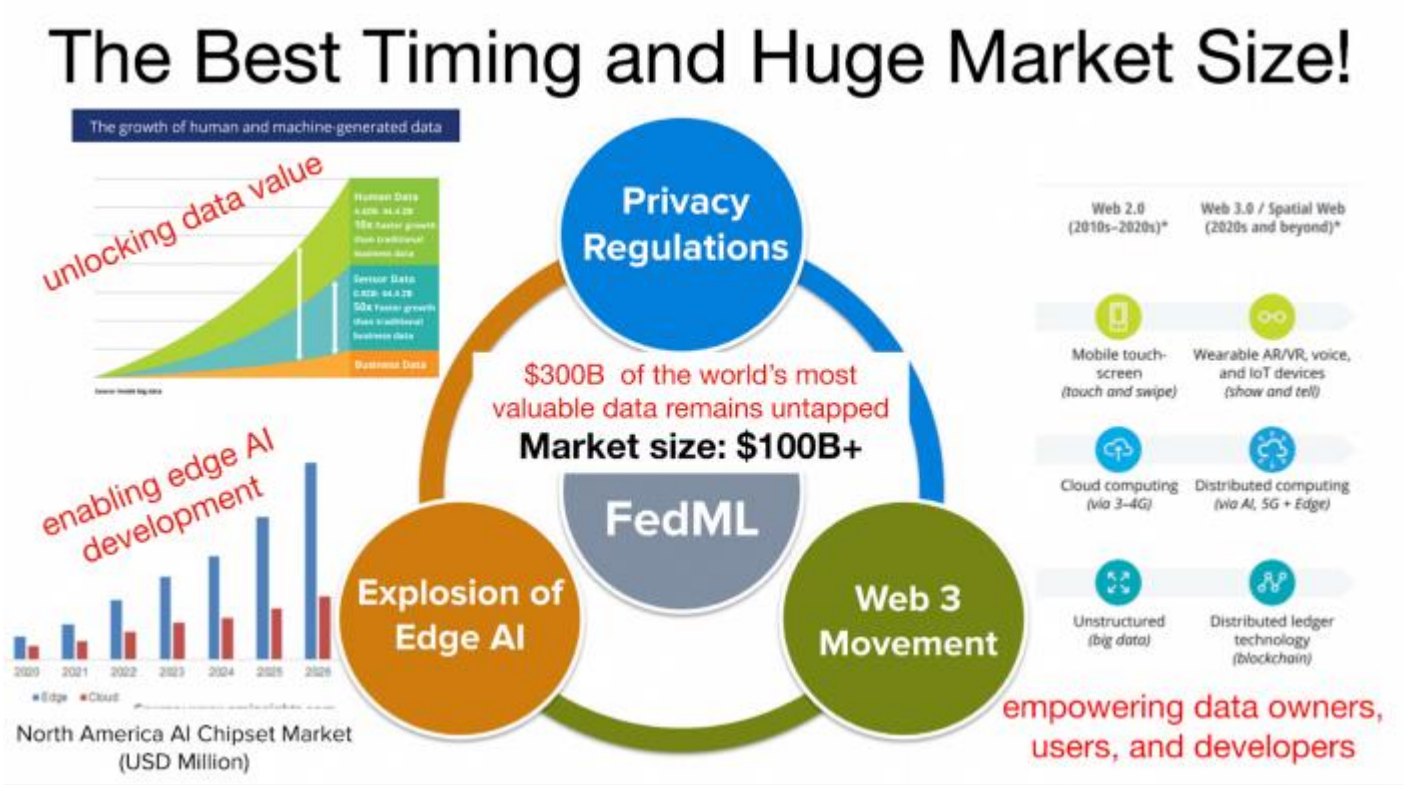
FedML Product Overview, https://medium.com/@FedML/fedml-ai-platform-releases-the-worlds-federated-learning-open-platform-on-public-cloud-with-an-8024e68a70b6



FedML

FedML Product Overview, https://medium.com/@FedML/fedml-ai-platform-releases-the-worlds-federated-learning-open-platform-on-public-cloud-with-an-8024e68a70b6

In this era, federated learning technology is critical because it is at the historical intersection of the three major technological hotspots: privacy computing, edge AI, and Web3/Blockchain.



FedML

FedML Product Overview, https://medium.com/@FedML/fedml-ai-platform-releases-the-worlds-federated-learning-open-platform-on-public-cloud-with-an-8024e68a70b6

What never before have we considered is distributed AI algorithms on a large number of nodes with scattered small data while simultaneously taking into account security, system efficiency, and model accuracy.
In addition, many concepts in blockchain, including data decentralization computing, ownership verification, traceability, incentive mechanism, and trusted security, all coincide with the idea of federated learning to make the data value flow safely.

FedML

FedML Product Overview, https://medium.com/@FedML/fedml-ai-platform-releases-the-worlds-federated-learning-open-platform-on-public-cloud-with-an-8024e68a70b6

In academia, the most popular open-source framework is undoubtedly FedML (https://github.com/FedML-AI), which is widely used around the world (see https://fedml.ai/use-cases/). (there are also some other frameworks, but FedML is the one cited most: FedML-196; Flower-96; PySyft-32; FedScale-20; FATE-14).

Today, the FedML team has further upgraded these academic achievements into an industrialized platform. Its mission is to build open and collaborative AI anywhere at any scale. In other words, FedML supports both federated learning for data silos and distributed training for acceleration with MLOps and Open Source support, covering cutting-edge academia research and industrial grade use cases.

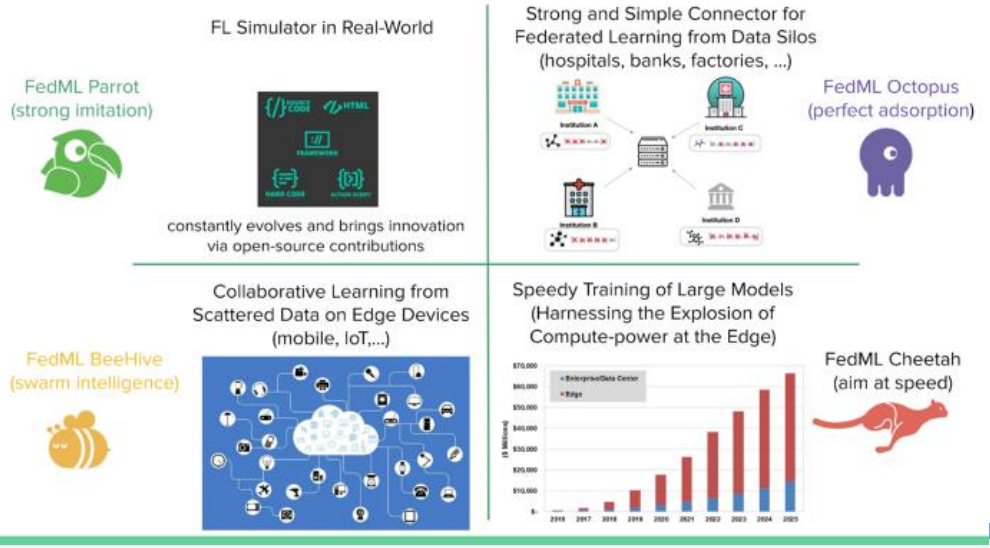
Different from the closed-source commercial platforms of most companies or the restricted mode of "applying for a trial," FedML focuses on building a public MLOps (ML Operations) platform, which is open to global users for free.

FedML

FedML Product Overview, https://medium.com/@FedML/fedml-ai-platform-releases-the-worlds-federated-learning-open-platform-on-public-cloud-with-an-8024e68a70b6

At its current stages, FedML provides the following services:

- FedML Parrot — Simulating federated learning in the real world.
- FedML Octopus — Cross-silo Federated Learning for cross-organization/account training, including Python-based edge SDK.
- FedML Beehive — Cross-device Federated Learning for Smartphones and IoTs, including edge SDK for Android/iOS and embedded Linux.
- FedML MLOps: FedML's machine learning operation pipeline for AI running anywhere at any scale.
- Model Serving: providing a better user experience for edge AI.



FedML

FedML Product Overview, https://medium.com/@FedML/fedml-ai-platform-releases-the-worlds-federated-learning-open-platform-on-public-cloud-with-an-8024e68a70b6

FEDML MLOPS USER GUIDE, https://open.fedml.ai/octopus/userGuides/index

Write Once, Run Anywhere: Seamlessly Migrate Your Local Development to the Real-world Edge-cloud Deployment

Video Tutorial: https://www.youtube.com/embed/Xgm0XEaMlVQ

FEDERATED LEARNING ON ANDROID SMARTPHONES,
https://open.fedml.ai/beehive/userGuides/index

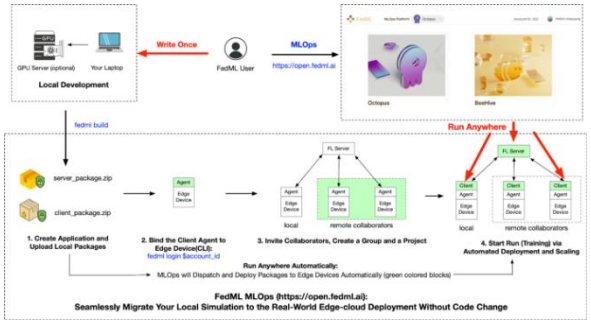
FedML

FedML Product Overview, https://medium.com/@FedML/fedml-ai-platform-releases-the-worlds-federated-learning-open-platform-on-public-cloud-with-an-8024e68a70b6

3.1 Seamless migration between simulated experiment and real deployment, zero code modification

FedML can help users seamlessly migrate the code of experimental simulation (POC) to the actual system (Production) to carry out experiments under real private data and distributed training systems of edge devices.

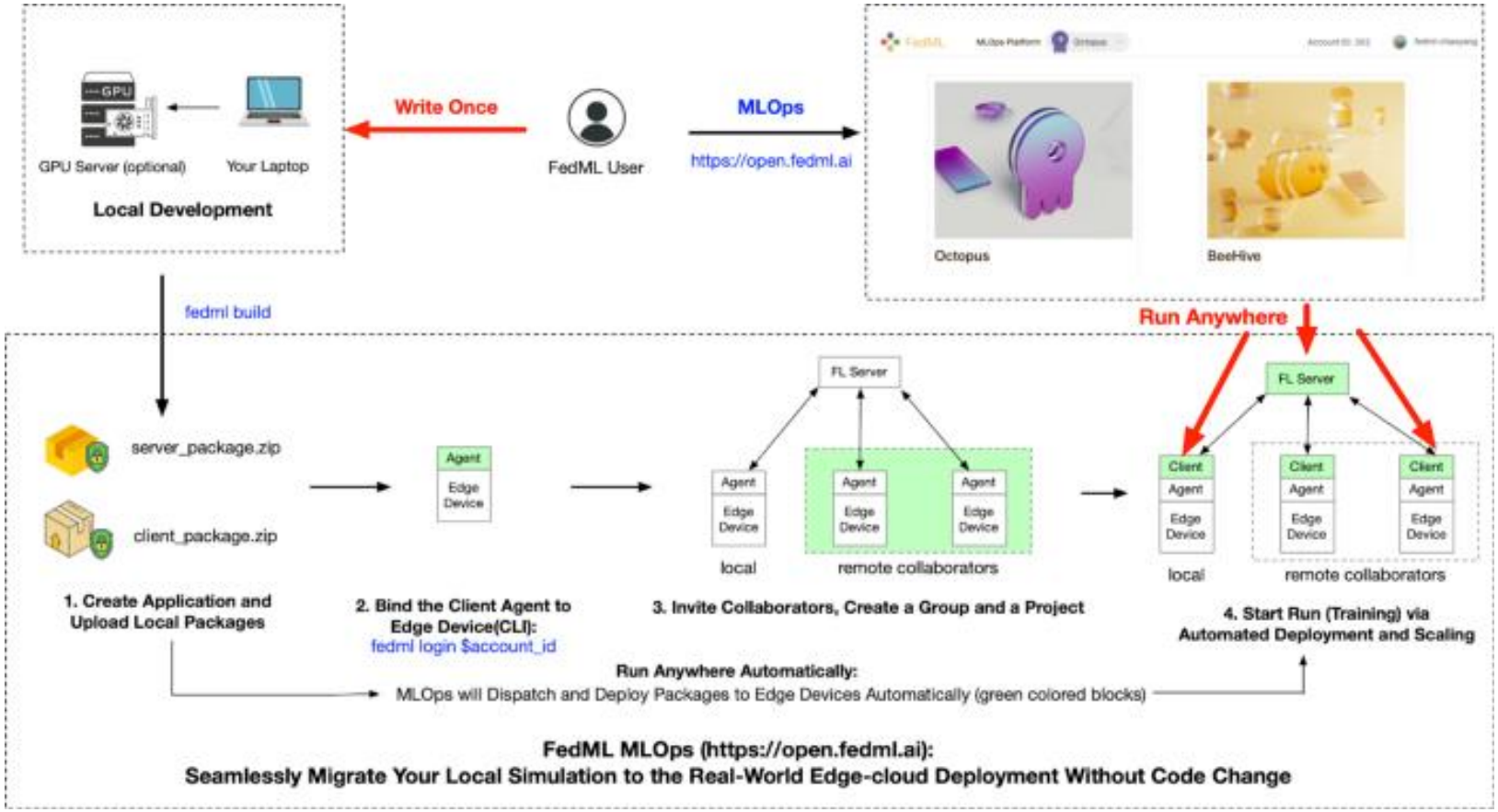
- After it is verified that federated learning can produce modeling benefits on specific applications, users can use FedML MLOps to upgrade the simulation into production without modifying the code.
- The simulated source code can be deployed directly to edge devices with real data. As shown in the upper left of Figure 1, the user first completes local development and debugging in "Local Development" and then generates installation packages.
- Finally, through simple commands and UI interactions, these installation packages and scripts can be distributed to any private device (shown in red on the right of Figure 1).



FedML

FedML Product Overview, https://medium.com/@FedML/fedml-ai-platform-releases-the-worlds-federated-learning-open-platform-on-public-cloud-with-an-8024e68a70b6

3.1 Seamless migration between simulated experiment and real deployment, zero code modification



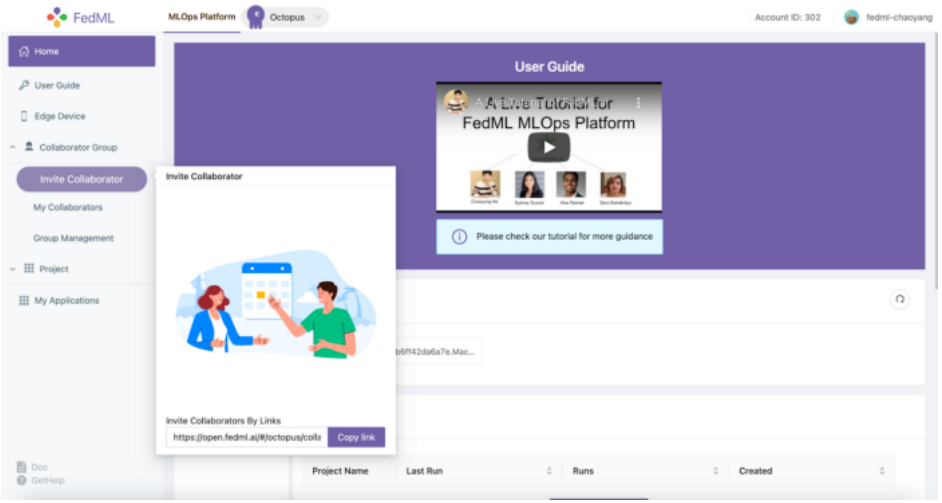
Federated Learning

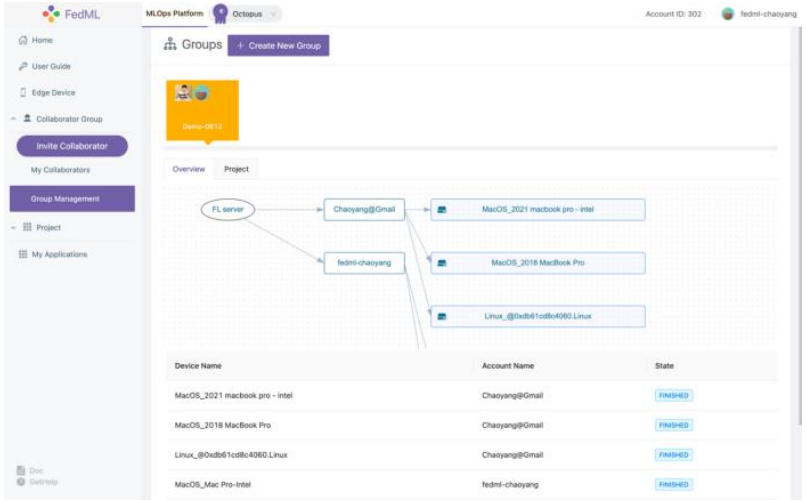
FedML Product Overview, https://medium.com/@FedML/fedml-ai-platform-releases-the-worlds-federated-learning-open-platform-on-public-cloud-with-an-8024e68a70b6

3.2 One-line command to complete the edge deployment

3.3 Support simplified collaboration anywhere: multinational, cross-city, multi-tenant

FedML collaboration has become extremely simple. Just as you can collaborate on documents with friends, you can easily create a federated learning group by sending an invitation link. There are no geographical, national, or city restrictions anywhere in the world.



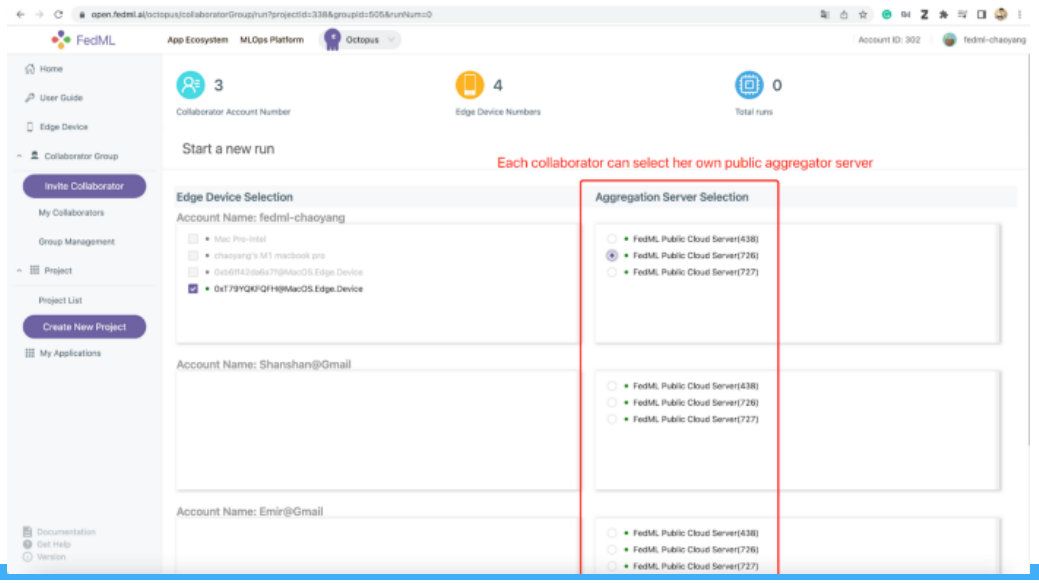


FedML

FedML Product Overview, https://medium.com/@FedML/fedml-ai-platform-releases-the-worlds-federated-learning-open-platform-on-public-cloud-with-an-8024e68a70b6

3.4 Provide free public cloud aggregation server and private cloud deployment with Docker

To reduce the difficulty of federated training, FedML's open platform provides a public cloud aggregation server for everyone. Users can arbitrarily select free service nodes offered on the public cloud when initiating training.
FedML also considers more secure and strict deployment requirements for platform users. For this reason, FedML platform has also developed a private aggregation server library that can be deployed freely. It still only needs a one-line command, that is, the "fedml login -s" command to run a secure docker environment at any self-hosted server.



FedML

FedML Product Overview, https://medium.com/@FedML/fedml-ai-platform-releases-the-worlds-federated-learning-open-platform-on-public-cloud-with-an-8024e68a70b6

3.5 Experiment monitoring and analysis capabilities tailored for distributed training

In addition to the features mentioned above of lowering the user access threshold, the FedML platform also helps AI application modeling by providing experimental tracking, management, visualization, and analysis capabilities. Key capabilities currently supported include:

1. Edge device training status tracking.

2. Custom metrics reporting, such as the accuracy in common classification tasks in deep learning, the error rate of regression tasks, and even the running time of the system state, memory usage, and GPU utilization.

3. Profiling flow and edge device system performance. It can help users to view the execution performance of different subtasks on each edge device, which is convenient for analyzing the bottleneck of the system.

4. Distributed logging. This is a capability that the current general-purpose machine learning platform does not have, and it is convenient to track and analyze the anomalies that occur on each device in a real-time manner.

5. The experimental report allows users to compare multiple experimental results.

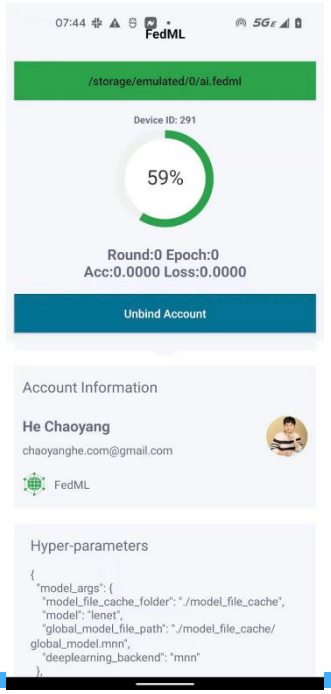
FedML

FedML Product Overview, https://medium.com/@FedML/fedml-ai-platform-releases-the-worlds-federated-learning-open-platform-on-public-cloud-with-an-8024e68a70b6

3.6 Unified cross-platform design, supporting smartphones and IoT devices

FedML also recently released the Android platform for mobile devices; details can be found at the following links:

FedML Android Platform: https://github.com/FedML-AI/FedML/tree/master/android
FedML IoT Platform: https://github.com/FedML-AI/FedML/tree/master/iot



FedML

FedML Product Overview, https://medium.com/@FedML/fedml-ai-platform-releases-the-worlds-federated-learning-open-platform-on-public-cloud-with-an-8024e68a70b6

4. Three-in-one Strategy for Open Collaboration: Open Source, Open Platform, Collaborative Application Ecosystem

In addition to operating open source communities, FedML promotes open collaboration and open source research and development from multiple product perspectives.

- Besides the open source library (https://github.com/FedML-AI),
- and open platform (https://open.fedml.ai),

FedML has also developed the collaborative App Ecosystem (users can visit https://open.fedml.ai and find "App Ecosystem" on the left top).

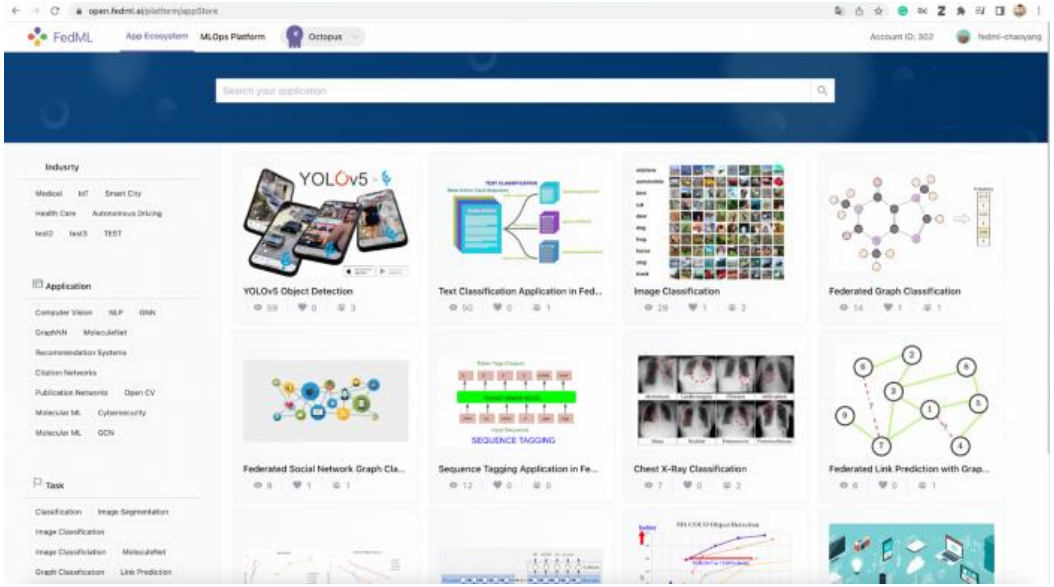
- The App Ecosystem and the platform cooperate with each other to continuously enrich the application ecosystem. The first version has completed the open collaboration of over 20 applications.
- Users can contribute and share the application. Each application includes all the FedML-based source code of an AI application, including model definitions, training scripts, and configuration files.
- At present, the App Ecosystem covers mainstream AI application scenarios such as computer vision, natural language processing, graph data mining, and the Internet of Things.

FedML

FedML Product Overview, https://medium.com/@FedML/fedml-ai-platform-releases-the-worlds-federated-learning-open-platform-on-public-cloud-with-an-8024e68a70b6

4. Three-in-one Strategy for Open Collaboration: Open Source, Open Platform, Collaborative Application Ecosystem

If the open platform reduces the difficulty of actual building deployment of the federated learning system to the lowest level, then the App Ecosystem is used to lower the AI application R&D threshold for practitioners: A company needs not to hire high-cost machine learning teams but rather needs only one engineer who can do "one-click import" on the basis of community results and use the application directly without intensive development circles.



FedML

FedML Product Overview, https://medium.com/@FedML/fedml-ai-platform-releases-the-worlds-federated-learning-open-platform-on-public-cloud-with-an-8024e68a70b6

5. Simple and flexible APIs, boosting innovation in algorithm and system optimization

- First, from the application point of view, FedML does its best to shield all code details and complex configurations of distributed training. Data scientists and engineers at the application level, such as computer vision, natural language processing, and data mining, only need to write the model, data, and trainer in the same way as a stand-alone program and then pass it to the FedMLRunner object to complete all the processes. This greatly reduces the bar for application developers to perform federated learning.
- Secondly, the FedML team believes that the design of the API should conform to the current technology development trend and should not assume that today's technology is the final solution; rather, it should be iterated as it progresses. We can see that the algorithm innovation of the open source community is still very active, and many more user-valued algorithms continue to be innovated every month. It is based on this background that FedML considers making custom APIs flexible enough to empower algorithm innovation. To this end, FedML abstracts the core trainer and aggregator and provides users with two abstract objects, FedML.core.ClientTrainer and FedML.core.ServerAggregator, which only need to inherit the interfaces of these two abstract objects and pass them to FedMLRunner. Such customization provides machine learning developers with maximum flexibility. Users can define arbitrary model structures, optimizers, loss functions, etc. These customizations can also be seamlessly connected with the open source community, open platform, and application ecology mentioned above with the help of FedMLRunner, which completely solves the long lag problem from innovative algorithms to commercialization.

FedML

FedML Product Overview, https://medium.com/@FedML/fedml-ai-platform-releases-the-worlds-federated-learning-open-platform-on-public-cloud-with-an-8024e68a70b6

5. Simple and flexible APIs, boosting innovation in algorithm and system optimization

- Finally, FedML believes that although FL is a comprehensive technology that combines security, system efficiency, and model accuracy, the first priority is still "ML-oriented Research and Development". For example, security and system optimization are definitely important, but it is not a good product design for ML users if they have a huge learning burden on security and system design — this would eventually cause the core users to abandon the product. Therefore, in terms of architecture, FedML considers that security, privacy, and system optimization should all serve ML. The details of these auxiliary modules are hidden throughout layered design, and it is ultimately through this that the best ML experience is achieved. This responsibility is accomplished through FedML Flow.

FedML

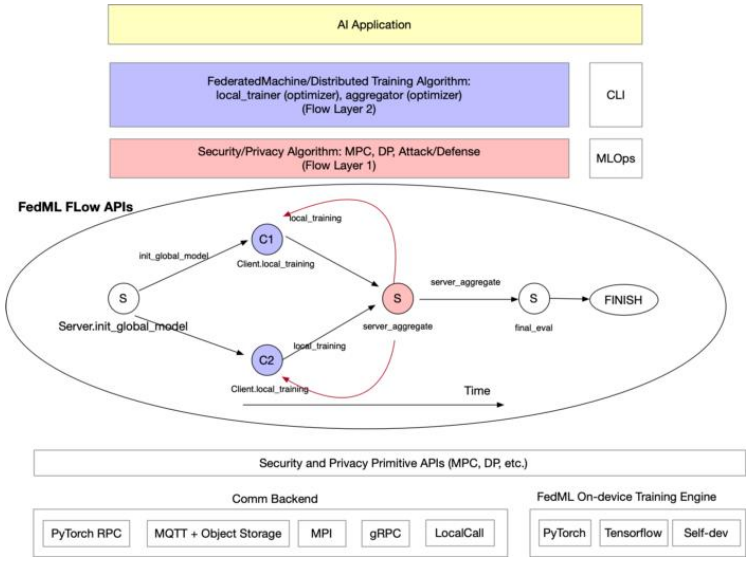
FedML Product Overview, https://medium.com/@FedML/fedml-ai-platform-releases-the-worlds-federated-learning-open-platform-on-public-cloud-with-an-8024e68a70b6

5. Simple and flexible APIs, boosting innovation in algorithm and system optimization

Specifically, as shown in the figure, FedML regards distributed computing processes such as complex security protocols and distributed training as a directed acyclic graph (DAG) flow computing process, making the writing of complex protocols similar to stand-alone programs.

- Based on this idea, the security protocol Flow Layer 1
- and the machine learning algorithm process Flow Layer 2

can be easily separated so that security engineers and machine learning engineers can perform their duties without having to master multiple technologies with the same mindset.



Federated Learning

FedML Product Overview, https://medium.com/@FedML/fedml-ai-platform-releases-the-worlds-federated-learning-open-platform-on-public-cloud-with-an-8024e68a70b6

5. Simple and flexible APIs, boosting innovation in algorithm and system optimization

For a more intuitive understanding of FedML Flow, the following example demonstrates the process of implementing the FedAvg algorithm and adding multiple distributed tasks through FedML Flow.

- First, each distributed node can be regarded as an abstract FedMLExecutor, which is carried in an independent process and is responsible for executing a specific task. This task can be training or some protocol messages, thus maintaining a high degree of flexibility and abstraction.

- Flow is a framework that helps to transfer the behavior of these distributed Executors, and it can arrange the order of task execution and message passing between tasks. Specific to a FedAvg algorithm, we can define a Client Executor and a Server Executor object and use their custom functions as tasks in the flow.

- Through the Flow API, users can freely combine the execution processes of these Executors. The following code shows the entire process of model initialization, multiple rounds of training, and finally, distributed evaluation of the model.

- The most important thing is that is that this programming example only happens on the personal computer of FedML users and does not require any distributed system development skills

FedML

FedML Product Overview, https://medium.com/@FedML/fedml-ai-platform-releases-the-worlds-federated-learning-open-platform-on-public-cloud-with-an-8024e68a70b6

```
    if args.rank == 0:
        executor = Server(args)
        executor.init(device, dataset, model)
    else:
        executor = Client(args)
        executor.init(device, dataset, model)

    fedml_alg_flow = FedMLAlgorithmFlow(args, executor)
    fedml_alg_flow.add_flow("init_global_model", Server.init_global_model)
    fedml_alg_flow.add_flow("handle_init", Client.handle_init_global_model)
    for round_idx in range(args.comm_round):
        fedml_alg_flow.add_flow("local_training", Client.local_training)
        fedml_alg_flow.add_flow("server_aggregate", Server.server_aggregate)
    fedml_alg_flow.add_flow("final_eval", Server.final_eval)
    fedml_alg_flow.build()

    fedml_runner = FedMLRunner(args, device, dataset, model, algorithm_flow=fedml_alg_flow)
    fedml_runner.run()

class Client(FedMLExecutor):
    def local_training(self):
    def handle_init_global_model(self):

class Server(FedMLExecutor):
    def init_global_model(self):
    def server_aggregate(self):
    def final_eval(self):
```

FedML

FedML Product Overview, https://medium.com/@FedML/fedml-ai-platform-releases-the-worlds-federated-learning-open-platform-on-public-cloud-with-an-8024e68a70b6

6. Release Cross-silo FL dataset with Okwin, allowing research face real scenarios

Different from the synthetic or hypothetical data combed by papers such as LEAF and FedScale (note: only natural ID segmentation does not represent the authenticity of products and business scenarios), the datasets released by FedML and Okwin are taken from real federated learning scenarios.

As shown in the figure below, the current version mainly focuses on medical scenarios, containing 7 naturally partitioned medical datasets covering multiple tasks, models, and data modalities, each with baseline training codes for everyone to conduct research and development. All these results are published on the FedML open platform; see https://open.fedml.ai for details (click "App Ecosystem" after logging in).

Dataset	Fed-Camelyon16	Fed-LIDC-IDRI	Fed-IXI	Fed-TCGA-BRCA	Fed-KITS2019	Fed-ISIC2019	Fed-Heart-Disease
Input (x)	Slides	CT-scans	T1WI	Patient info.	CT-scans	Dermoscopy	Patient info.
Preprocessing	Matter extraction + tiling	Patch Sampling	Registration	None	Patch Sampling	Various image transform	Removing missing data
Task type	binary classificatior	3D segmentation	3D segmentation	survival	3D segmentation	multi-class classificatio	binary classification
Prediction (y)	Tumor on slide	Lung Nodule Mask	Brain mask	Risk of death	Kidney and tumor masks	Melanoma class	Heart disease
Center extraction	Hospital	Scanner Manufacturer	Hospital	Group of Hospitals	Group of Hospitals	Hospital	Hospital
Thumbnails							
Original paper	Litjens *et al.* 2018	Armato *et al.* 2011	Perez *et al.* 2021	Liu *et al.* 2018	Heller *et al.* 2019	Tschandl *et al.* 2018 / Codella *et al.* 2017 / Combalia *et al.* 2019	Janosi *et al.* 1988
# clients	2	5	3	5	6	5	4
# examples	399	1,018	566	1, 088	96	23, 247	740
# examples per center	239, 150	670, 205, 69, 74	311, 181, 74	311, 196, 206, 162 51	12, 14, 12, 12, 16, 30	12413, 3954, 3363, 225 819, 439	303, 261, 46, 130
Model	DeepMIL [63]	Vnet [97, 100]	3D U-net [23]	Cox Model [31]	nnU-Net [66]	efficientnet [117] + linear layer	Logistic Regression
Metric	AUC	DICE	DICE	C-index	DICE	Balanced Accuracy	Accuracy
Size	50G (850G total)	115G	444M	115K	54G	9G	40K
Image resolution	0.5 μm / pixel	~1.0 × 1.0 × 1.0 mm / voxel	~ 1.0 × 1.0 × 1.0 mm / voxel	NA	~1.0 × 1.0 × 1.0 mm / voxel	~0.02 mm / pixel	NA
Input dimension	10, 000 x 2048	128 x 128 x 128	48 x 60 x 48	39	64 x 192 x 192	200 x 200 x 3	13

FedML

FedML Product Overview, https://medium.com/@FedML/fedml-ai-platform-releases-the-worlds-federated-learning-open-platform-on-public-cloud-with-an-8024e68a70b6

7. Published 50+ top scientific papers, covering key challenges such as security, efficiency, weak supervision, and fairness

All papers are summarized at https://doc.fedml.ai/resources/papers.html

8. Academia Sponsorship

For more details, please visit: https://fedml.ai/academia-sponsorship/

9. FedML Team

AIFactory

https://aifactory.space/