



HAL
open science

Decolonizing Federated Learning: Designing Fair and Responsible Resource Allocation Position Paper

Genoveva Vargas-Solar, Javier-Alfonso Espinosa-Oviedo, Nadia Bennani, Andrea Mauri, José-Luis Zechinelli-Martini, Barbara Catania, Claudio A. Ardagna, Nicola Bena

► **To cite this version:**

Genoveva Vargas-Solar, Javier-Alfonso Espinosa-Oviedo, Nadia Bennani, Andrea Mauri, José-Luis Zechinelli-Martini, et al.. Decolonizing Federated Learning: Designing Fair and Responsible Resource Allocation Position Paper. ACS/IEEE 21st International Conference on Computer Systems and Applications (AICCSA 2024), Oct 2024, Sousse, Tunisia. hal-04743266

HAL Id: hal-04743266

<https://hal.science/hal-04743266v1>

Submitted on 18 Oct 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Decolonizing Federated Learning: Designing Fair and Responsible Resource Allocation

Position Paper

Geneveva Vargas-Solar*, Nadia Bennani*, Javier A. Espinosa-Oviedo*[†] Andrea Mauri*, José-Luis Zechinelli-Martini[‡] Barbara Catania[§] Claudio Ardagna[¶] and Nicola Bena[¶]

*CNRS, Univ. Lyon, INSA Lyon, CPE, UCBL, LIRIS, UMR5205, Lyon, France
geneveva.vargas-solar@cnrs.fr, Nadia.Bennani@insa-lyon.fr, andrea.mauri@univ-lyon1.fr

[†] Univ. Lyon, Univ. Lyon 1, UR ERIC, EA 3083, Villeurbanne, France
javier.espinosa@univ-lyon1.fr

[‡]Fundación Universidad de las Américas Puebla, San Andrés Cholula, Mexico
joseluis.zechinelli@udlap.mx

[§] University of Genoa, Italy
barbara.catania@unige.it

[¶] University of Milano, Italy

claudio.ardagna@unimi.it, nicola.bena@unimi.it

Abstract—This position paper explores the challenges, existing solutions, and open issues related to resource allocation in federated learning environments. The focus is on how to allocate resources effectively while adhering to service level objectives (SLOs) and fairness requirements, which include factors such as server location, data provenance, energy consumption, sovereignty, carbon footprint, and economic cost. The goal is to optimise resource distribution across different stages of the federated learning process within a given architecture, ensuring that these fairness criteria are integrated into the allocation strategy. This approach aligns with decolonial methodologies that seek to offer more sustainable and equitable alternatives to the resource-intensive artificial intelligence processes prevalent today.

Keywords— Federated learning, fairness, data sovereignty, resources allocation, responsible AI.

I. INTRODUCTION

Digital colonialism refers to the dominance and exploitation of digital technologies, data, and infrastructure by powerful entities, often multinational corporations or Global North countries, over less powerful nations, particularly in the Global South. This phenomenon highlights significant power imbalances in the global digital economy and raises concerns about its impact on the sovereignty, culture, and economic development of affected nations. Large tech companies frequently collect and exploit vast amounts of data from users in the Global South, benefiting corporations in the Global North without fair compensation or control for the data’s original owners. Many developing countries depend on technology infrastructure provided by foreign companies, leading to a lack of digital sovereignty and limited control over essential digital services. Additionally, digital platforms often promote the dominant cultures and languages, marginalizing local voices and leading to cultural homogenization. The economic benefits of digital services are concentrated among a few powerful corporations, exacerbating global inequalities and hindering the development of local digital industries [1].

Reversing digital colonialism involves empowering nations and communities in the Global South to gain greater control over their

digital infrastructure, data, and technological development. Strengthening data governance through robust data protection laws ensures that data collected within a country remains under local control, promoting fair data practices and equitable benefit sharing. Investing in local tech ecosystems is crucial, to build homegrown technologies that can compete globally. Equally important is expanding digital literacy and education, particularly in underserved communities, to empower individuals to actively participate in the digital economy and ensure ethical use of technology.

The rapid development of artificial intelligence-based data analytics methods has generated significant momentum but also raised concerns about the immense computing, economic, and ecological resources required for training models and deploying final solutions. The resource demands are colossal, leading to substantial ecological and economic impacts. Additionally, the bias and incompleteness inherent in models trained on partial content—often in English, digital, and predominantly from the Global North—exacerbate epistemic violence¹ by marginalizing diverse perspectives and knowledge systems from the Global South.

The use of the underlying infrastructure, such as computing servers, storage locations, and governance, is often obscured from users. Critical questions arise regarding the origins and handling of data: Where does the data originate? When the data involves specific territories, communities, or individuals, are they aware of its use for analytics purposes, and do they consent to it? Furthermore, do they agree with the interpretations and conclusions drawn from the analytics? Are they informed about where the hardware processing their data is located, and do they consent to that as well? Finally, are they aware of, and do they approve, the vast amount of computing resources consumed by AI models to analyze their data? Various

¹Epistemic violence refers to the harm done to individuals or groups through the marginalization, silencing, or devaluation of their knowledge, experiences, and ways of understanding the world. It occurs when certain forms of knowledge, often those from dominant or powerful groups, are privileged over others, leading to the suppression or erasure of alternative perspectives. This can happen in various contexts, such as academia, media, and technology, where the voices and knowledge of marginalized communities are ignored, misrepresented, or dismissed.

approaches in computer science are now addressing these challenges by designing algorithms, data collection, preparation, querying, and storage with a focus on fairness, explainability, and human-centric methodologies. The aim is to minimize the harmful and exclusionary effects of AI and technological solutions on vulnerable communities, mitigate negative economic impacts, and counter biased perceptions.

In relation to the above questions and emerging approaches, this paper explores the challenges, existing solutions, and open issues related to resource allocation in federated learning environments. The focus is on how to allocate resources effectively while adhering to service level objectives (SLOs) and fairness requirements, which include factors such as server location, data provenance, energy consumption, sovereignty, carbon footprint, and economic cost. The goal is to optimise resource distribution across different stages of the federated learning process within a given architecture, ensuring that these fairness criteria are integrated into the allocation strategy. This approach aligns with decolonial methodologies that seek to offer more sustainable, equitable and responsible alternatives to the resource-intensive artificial intelligence (AI) processes prevalent today.

Accordingly, the remainder of the paper is organised as follows. Section II states the problem of designing fair resource allocation strategies for federated learning processes. Section III enumerates and discusses the main scientific and technical challenges to address the problem and discusses the main associated implications. Section IV gives an overview of related work concerning approaches and results associated with the ambition of providing architectures that consider sobriety in the allocation of resources. Finally, Section V concludes the paper and discusses future work.

II. FAIRNESS IN RESOURCE ALLOCATION FOR FEDERATED LEARNING PROCESSES: PROBLEM STATEMENT

The analysis of data through AI models must respect the exchange, sharing and processing of raw data. AI risk undermining privacy and respect for data. Federated learning techniques provide a solution to these requirements. The principle is to be able to train an algorithm on several independent nodes, by forming local models on local data samples. This strategy gives control over the data that feeds the algorithms to independent nodes, which exchange with a coordinator node the minimum amount of data necessary to integrate the results and build a global model.

a) Executing federated learning processes: involve training machine learning models across multiple decentralized devices or servers that hold local data samples, without the need to transfer this data to a central server. The process begins with the initialization of a global model, which is then distributed to all participating nodes. Each node uses its own local data to train the model, updating the model parameters based on the specific data it holds. After local training, the nodes send their updated models back to a central server, where these models are aggregated—typically by averaging the parameters—to update the global model.

This iterative process of distributing the global model, performing local training, and aggregating the results continues over several rounds until the model reaches an optimal state. Once the model is sufficiently trained and has converged, it is finalised and deployed for use across the network or in specific applications. This method enables effective model training while preserving data privacy, as the raw data remains decentralised and never needs to be shared with a central server. By decentralising the learning process, federated learning reduces the need for centralised data storage, thereby enhancing data privacy.

When allocating resources for federated learning processes, key aspects to consider include:

- **Computational Power:** Ensuring that each participating node has sufficient computing resources to effectively train the model locally.

- **Network Bandwidth:** Managing the data transfer between nodes and the central server, particularly during model aggregation, to avoid bottlenecks.
- **Scalability:** Allocating resources in a way that supports the addition of more nodes without degrading performance.
- **Latency:** Minimizing delays in communication between nodes and the central server to maintain efficient model updates.
- **Storage:** Providing adequate storage at each node for the model and data, and at the central server for aggregating and updating the global model.

Additional aspects (constraints) can be integrated into the conditions under which a federated learning process is executed to ensure it is conducted responsibly. A responsible execution aims to minimize resource and energy consumption, reduce the CO₂ footprint, and adhere to region-specific process localisation. It ensures data and execution sovereignty by giving users full control over the sampled and selected data, its provenance, storage location, and the types of analytics applied. Additionally, it optimizes costs related to model training and deployment while promoting transparency in resource allocation and data management, including storage, transfer, and processing. These attributes can effectively guide and influence the execution of federated learning processes, ensuring that they align with ethical and sustainable practices.

b) Fairness and resources allocation: The next steps for advancing fairness in resource allocation and analytics processes involve several key actions:

- **Develop Comprehensive Fairness Metrics for Data:** Create robust and multi-dimensional fairness metrics that account for intersectionality, ensuring that multiple overlapping attributes (e.g., age, gender, ethnicity) are considered in fairness assessments.
- **Promote Intersectional² Approaches:** Encourage the adoption of intersectional strategies in fairness research and practice, ensuring that the unique experiences and challenges of diverse groups are adequately addressed.
- **Integrate Fairness into Systems:** Design and implement mechanisms for real-time measurement and monitoring of fairness during the execution of tasks and resource allocation, allowing for dynamic adjustments to maintain fairness.
- **Enhance Compliance Verification:** Develop automated tools and frameworks to continuously verify and enforce compliance with fairness metrics across all nodes and the federation as a whole.
- **Incorporate Fairness in Decision-Making:** Embed fairness considerations into the decision-making algorithms used for process deployment, job scheduling, and resource allocation, ensuring that these processes are guided by fairness constraints.
- **Continuous Evaluation and Adaptation:** Establish ongoing processes for evaluating the effectiveness of fairness measures and adapting them as new challenges and contexts emerge [2].
- **Collaborate Across Disciplines:** Foster interdisciplinary collaboration to integrate insights from social sciences, ethics, and technology, creating holistic approaches to fairness in data analytics and resource allocation.

c) Discussion: The aim of our work is to ensure fair resource allocation for federated learning phases across nodes and models. We focus on meeting fairness metrics established by the communities that produce and consume data and analytics results. Broadly, the research questions we suggest that should be explored concern the following key areas:

- **Defining Fairness Metrics:** How can we establish measures that accurately define fairness in the context of resource allocation?

²Intersectionality is a conceptual framework that examines how different forms of social identities—such as race, gender, class, sexual orientation, ability, and other characteristics—intersect and interact to create unique and overlapping systems of oppression or privilege.

- Measuring Fairness at runtime: What methods can be used to assess fairness during execution, and how can this metric be integrated into process deployment and resource allocation strategies once the federation is established?
- Ensuring Compliance: How can we verify that both individual nodes and the entire federation adhere to the established fairness index?
- Respecting Fairness Constraints: What strategies can be employed to ensure that deployment decisions, job assignments, and resource allocations align with the fairness constraints the federation seeks to uphold?

III. TOWARDS FAIR RESOURCE ALLOCATION WITH A DECOLONIAL AND RESPONSIBLE PERSPECTIVE

Our work focuses on developing a data management approach within federated learning environments to ensure that these federations are equitable, providing a fair analysis of data from a decolonial and responsible perspective (see Figure 1). Three key challenges must be addressed to integrate this approach into the execution of federated learning processes: (1) the declaration of target combination of metrics that reflect the fairness expectations of the application or user community from the federated learning process; (2) the specification of metrics that can be dynamically monitored to assess the extent to which the execution aligns with given fairness thresholds; and (3) the enforcement of the “contracts” (e.g., SLO agreements) established between the federated learning execution environment and the users or applications to ensure adherence to these agreed-upon standards.

A. Specifying a fairness index for resources allocation

Our vision for defining a Fairness Index (FI) with a decolonial perspective involves combining several key metrics into a weighted formula, allowing for the prioritization of user or application specific preferences. The proposed metrics encompass considerations related to the social, economic, and ecological costs of processing data, executing AI algorithms, and deploying technology. They are inspired by discussions on decolonizing technology and represent an initial effort to address the decolonization of resource allocation³. The proposed Fairness Index for federated learning environments integrates metrics using the following pondered formula:

$$FairnessIndex(FI) = \alpha_1 L_s + \alpha_2 P_s + \alpha_3 DS_p + \alpha_4 MP + \alpha_5 T_t + \alpha_6 C_{GPU} + \alpha_7 C_{cal} + \beta_1 TC_{CO_2} + \beta_2 EC \quad (1)$$

- L_s represents servers location,
- P_s represents provenance,
- DS_p denotes data sovereignty and provenance,
- MP indicates model performance,
- T_t is training time,
- C_{GPU} number of CPU/GPU cores used,
- C_{cal} refers to calibrating cycles,
- TC_{CO_2} represents the training cost in terms of CO₂ footprint,
- EC denotes the economic cost.
- The coefficients $\alpha_1 \dots \alpha_7, \beta_1, \beta_2$ are weights assigned to each metric. They reflect their relative importance within the fairness index. These weights can be adjusted based on the specific priorities or fairness objectives of the federated learning process.

This formula allows for a balanced consideration of various critical factors, ensuring that the federated learning process is fair, efficient, and responsible in terms of both environmental, social and economic impacts. The weights assigned to metrics are determined by the application context. The sum of all preference weights must equal 1.

³Yásnaya Elena Aguilar, Giltequiologías charla sobre tecnologías colaborativas en resistencia, <https://centroculturaldigital.mx/revista/tequiologias>

More precisely, to feed the FI we propose observing “simple” metrics such as *servers location, data sovereignty* and *provenance, model performance, training time, number of CPU/GPU cores* and *number of local model calibrating cycles* to achieve an optimal global decolonial performance within the federation. Each FL phase in the envisioned process involves three key tasks: job distribution across participating servers, result integration, and overall performance evaluation.

On the basis of the above attributes, two additional complex attributes are considered in our approach: (i) *training cost* of AI models in terms of CO₂ footprint; and (ii) *economic cost*, determined by factors such as *number of CPU/GPU cores* and *storage space*.

The CO₂ footprint is computed using metrics *training time* and *number of CPU/GPU cores* and *number of local model calibrating cycles*. The CO₂ footprint also involves calculating the Carbon Intensity of Electricity (in grams of CO₂ per kWh), which varies by region and is influenced by the local energy mix (e.g., coal, gas, renewables).

```
EnergyConsumption (kWh) =
PowerConsumption (kW) × TrainingTime (hours)

CO2Emissions (g CO2) =
EnergyConsumption (kWh) × CarbonIntensity

// CarbonIntensity (g CO2/kWh)
// Convert to kilograms (kg)
// by dividing by 1000.
```

When estimating the CO₂ footprint of training an AI model, it is essential to consider additional factors beyond the energy consumed by computational hardware. Data centers often require cooling systems and other infrastructure to maintain optimal operating conditions, which increases overall energy consumption. This added energy use is quantified by the Power Usage Effectiveness (PUE) factor, which typically ranges from 1.1 to 1.5, reflecting the efficiency of the data center⁴. Additionally, the energy costs associated with data transfers—especially when moving large volumes of data during the training process—should be included in the total energy consumption. These factors are critical for accurately estimating the environmental impact of AI model training. We assume that each federated learning job is associated with the set of simple and complex metrics. Each job can have an associated target FI threshold specified by the user/application requirements.

a) *Adjusting weights for defining the level of fairness for a given FL setting.*: For non-critical problems, reducing factors such as processing capacity, training time, and calibration cycles may take priority. In contrast, for social issues, data sovereignty and provenance become crucial. These aspects are vital to avoid exclusion and bias against human groups. Therefore, users can express preference items as an n-tuple, which pairs the attribute with a preference weight expressed as a percentage (used in the FI formula). For instance, the following n-tuple might represent the preference item corresponding to the server location attribute, using the coordinates of the geographical region in which we want the servers to be located as the threshold value, with a “preference weight” of 35%:

```
< location, < northernmost-point,
southernmost-point,
easternmost-point,
westernmost-point
>, 0,35 >
```

⁴A PUE of 1,1 indicates a highly efficient data center, while a PUE of 1,5 suggests that for every 1 kWh consumed by IT equipment, an additional 0,5 kWh is used for cooling and other overhead.

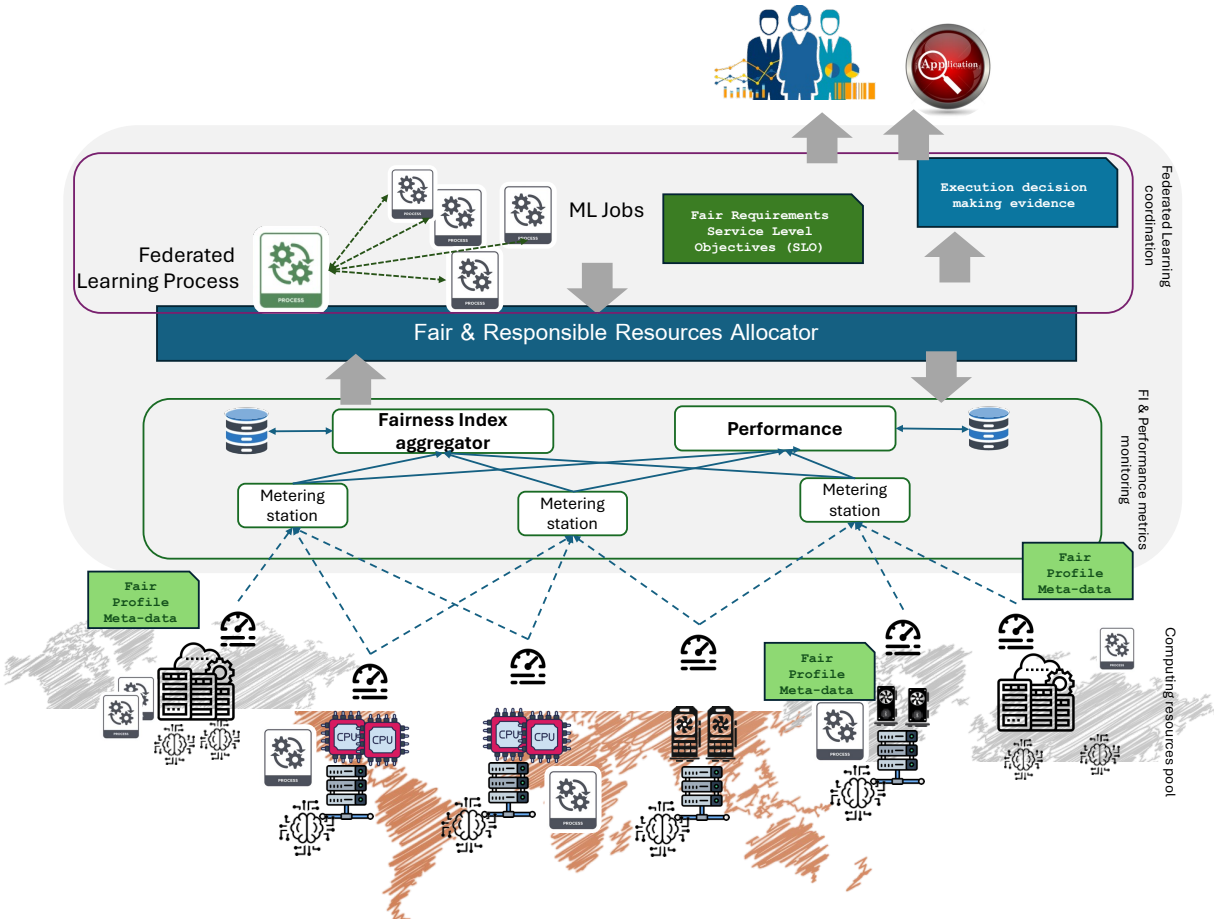


Fig. 1. Fair Resources Allocation for Federated Learning Processes

Using the previous item, one might specify that servers for a given federated learning job should be located in the region of Chiapas, Mexico. Similarly, a preference item for provenance might specify a set of acceptable owners for the servers used by the job. Additionally, data sovereignty should be reinforced through a certification process for the servers involved in the federated learning execution. This certification is conducted offline by the data and server owners, who must attest that data sovereignty is maintained and provide detailed logs documenting how the data are processed and utilized.

B. Measuring and monitoring fairness metrics

The principle of a federated learning architecture is that participating servers operate independently, functioning as almost black boxes concerning the specifics of the data used for training and testing models. This approach ensures that the results do not inadvertently reveal any information about the underlying data. To uphold this level of privacy and security, it is crucial to develop strategies that can collect and estimate fairness metrics associated with these servers without compromising data confidentiality. Additionally, other relevant metrics can be provided through metadata associated with the servers that opt to participate in the federated learning architecture, further enhancing the transparency and accountability of the process.

We build upon our previous work [3], [4], focused on estimating trust metrics—such as data freshness and update frequency—from black box data services. In these environments, where direct access to the underlying data is restricted, we developed methods to infer these metrics based on observable outputs and interaction patterns. This approach allows us to assess the reliability and timeliness of the

data provided by these services, even when the internal processes and data management practices are not transparent. By extending these techniques to the context of federated learning, we aim to estimate similar trust metrics for participating servers, ensuring that they contribute valuable and up-to-date information to the learning process without compromising the privacy and autonomy of the individual datasets.

To ensure analytics sovereignty in the use of AI models for addressing research questions specific to communities or issues in the Global South, it is essential to collect detailed data about the conditions under which the federated analytics process is conducted. This data should encompass the number of FL cycles required for convergence to a global result, the parameters used for calibrating the global model, the selection and number of participating servers, and their adherence to fairness metrics. Additionally, it is important to gather information on the global model itself, including its performance metrics, associated findings, and the environmental and economic costs incurred. This comprehensive approach helps maintain control over the analytics process, ensuring that it remains aligned with the interests and needs of the communities it serves.

C. Responsible resources allocation in target computing architectures

Existing Data Science (DS) environments, are typically "one-size-fits-all" cloud systems. They address the analytics and data management divide by offering integrated backends that efficiently execute analytics pipelines while allocating the necessary infrastructure resources, such as CPUs, FPGAs, GPUs, and TPUs, alongside

platform services like Spark and TensorFlow. These environments are designed to support the execution of DS tasks that demand significant storage and computational resources, with pipelines that often transition from in-house executions to cloud deployments. This shift necessitates elastic architectures that can dynamically scale resources to meet varying demands. Disaggregated data center solutions offer a promising approach to meet these needs. Disaggregated data centres are emerging as promising alternatives to fulfil the requirements of these applications. They promote composable infrastructures that can be customized to dynamically provide resources according to the workloads submitted by applications and fulfil their Service Level Objectives (SLOs). Our previous work has focussed on addressing the challenges involved in coupling disaggregated data center architectures with the execution of DS workflows, aiming to optimize resource allocation and efficiency [5].

Our goal is to refine our resource dispatching algorithm to dynamically meet Service Level Objectives (SLOs) such as response times while adhering to fairness constraints. By incorporating a fairness specification based on the previously mentioned metrics and optimizing for the Fairness Index (FI in equation 1), the dispatching algorithm will allocate computing resources and delegate the execution of analytics tasks in alignment with both the SLOs and FI. As the algorithm monitors the execution status of jobs assigned to various servers and the overall progress of the federated learning process, it can dynamically adjust resource allocation. For example, the algorithm might allocate GPU resources to train a model calibrated to prioritize the FI. The FI, which in turn, might involve lowering performance expectations to reduce energy consumption. Subsequently, it could switch to allocating CPU resources for the testing phases or model integration tasks at the global level within the federated learning architecture. This dynamic approach ensures that resource allocation is continuously optimized to balance fairness, performance, and energy efficiency. Since it may be unfeasible to jointly optimize both SLO and FI, a negotiation strategy can be applied and develop a best effort optimisation approach for implementing the resources allocation strategy. In this sense we will elaborate on our current work regarding trust in distributed systems [6].

a) *General lines of a use case:* Let us consider a dispatching algorithm that dynamically adjusts the allocated resources while the FL is executing using an *adaptive, cost- and energy-aware strategy*. It balances the preferences in terms of fairness, environmental impact, and economic cost, defined by the FI (equation 1), as well as other metrics in terms of SLOs. Several regions with different characteristics contribute to the FL task and specify different FI thresholds: (i) *high* prioritizes efficient hardware and low carbon emissions while limiting disproportionate advantages; (ii) *medium* prioritizes moderate computing resources and mid-level energy costs; (iii) *low* prioritizes minimal resources, high energy costs, or stringent data sovereignty rules, limiting regions' load while still encouraging participation.

Adaptive dispatching and monitoring. The dispatch algorithm dynamically adjusts task allocation based on real-time monitoring of each region's resources and FI value. Compute-intensive tasks (e.g., model training) are directed to regions with high-performance GPUs and low emissions (*high* FI). Data processing tasks (e.g., data cleaning) go to regions with moderate capabilities (*medium* FI), and lightweight tasks (e.g., validation) are allocated to less capable regions (*low* FI), respecting data privacy rules and minimizing environmental and economic burdens. When a region exceeds its FI threshold during the FL execution, the dispatching algorithm shifts the workload to regions with available capacity.

Cost and energy awareness. The algorithm further optimizes costs and energy by (i) dispatching more tasks to regions with surplus energy-efficient capacity or lower economic costs and (ii) scheduling tasks during periods of lower energy demand in regions with fluctuating electricity rates.

We note that a simpler, *static* dispatching algorithm can also be

implemented. It fixes allocated resources prior to execution thus reducing its overhead sacrificing real-time, optimal allocation.

IV. RELATED WORK

Our related work encompasses several key areas that align with our research focus. These include fairness-aware federated learning, which addresses the equitable distribution of resources and fairness in machine learning models across decentralized networks. Fair resource allocation is another critical area, exploring how to allocate computational and data resources in a way that ensures justice and equity among all participants. Additionally, fair data preparation is essential, emphasizing the importance of unbiased and representative datasets in training AI models. Sobriety in AI model training and use highlights the need for responsible, measured approaches to AI development, ensuring ethical considerations are at the forefront. Lastly, feminism in technologies provides a crucial perspective on how gender and power dynamics influence technological development, advocating for more inclusive and equitable tech practices. Together, these areas form the foundation of our work, guiding our approach to creating fair, responsible, and inclusive AI systems.

A. Fairness-aware federated learning (FAFL)

It refers to the approach that addresses and mitigates biases and ensures equitable outcomes across all participating nodes or servers in a federated learning system. This can involve various strategies, such as adjusting the weights of contributions from different nodes, implementing fairness constraints during the aggregation of local models, or using optimization techniques that explicitly aim to achieve fairness across different groups. The goal is to ensure that the trained model performs consistently and equitably across all participants, reducing disparities and avoiding biased outcomes that could negatively impact underrepresented or disadvantaged groups.

The work in [7] seeks to address the issue of bias in collaborative machine learning, where models trained using standard federated algorithms may favour certain participants, leading to uneven performance across the network. To combat this, various approaches have been proposed. Li Ju et al. [8] introduced AdaFedAdam, an adaptive federated optimisation algorithm designed to accelerate fair federated learning by reducing bias. Similarly, Salazar et al. [9] developed FAIR-FATE, a fairness-aware federated learning algorithm that achieves group fairness while maintaining high utility through a specialised aggregation method. Ezzeldin et al. [10] proposed FairFed, another fairness-aware aggregation method that enhances group fairness in federated learning without compromising utility. Additionally, Papadaki et al. [11] introduced FedMinMax, an optimisation algorithm aimed at achieving minmax group fairness in federated learning. These approaches have demonstrated improved fairness and have shown superior performance in terms of fairness and convergence compared to existing algorithms in federated learning.

These methods primarily focus on issues such as unbalanced workload distribution and poisoning attacks, rather than addressing the potential biases inherent in the data managed by individual nodes and how these biases can contaminate both local and global models due to processes that are not aware of diversity, equity, and inclusion (DEI) considerations. In fact, recent works also demonstrated how the lack of quality in the data can exacerbate bias [12], [13]. The purely technical aspects of load distribution, node interactions, and the aggregation of local models can still produce biased results, potentially leading to unfair and unjust conclusions.

B. Fair resources allocation

Fair resource allocation in data centers involves distributing resources like CPU, memory, and bandwidth among users to ensure efficiency and fairness. Techniques such as Dominant Resource Fairness (DRF) and its extensions, like DRF-CE, address these challenges by considering multiple resource types and user demands

[14], [15]. These methods aim to satisfy fairness properties such as envy-freeness, Pareto efficiency, and strategy-proofness [16], [17]. Additionally, new mechanisms have been proposed to optimize resource allocation in cloud-edge systems, improving user efficiency by considering bandwidth demand compression and other factors [18].

C. Fair data preparation

Existing techniques strive to ensure fairness in data analysis by addressing bias and promoting the representation of diverse groups characterised by various attribute values within data collections. For example, when analysing the career development of a group of people, these techniques work to ensure that individuals of all ages and genders are fairly represented in the data. In implementing analytics processes, it is crucial to ensure the fair treatment of data throughout the entire data analysis pipeline, from data collection [19]^{Andrea} to processing and analysis. Fairness in data analytics [20], [21] involves addressing biases, ensuring equitable treatment of all participants, and developing methodologies that promote fairness throughout all stages of data collection, processing, and analysis. It is a continuous effort that engages researchers, practitioners, policymakers, and society at large. Existing work on fairness in data analytics includes approaches that advocate for the use of data in ways that avoid creating or perpetuating biases [22], [23]. Additionally, these efforts extend to incorporating fairness and justice into data science processes through the ethics of data analytics [24], with some works also integrating feminist perspectives [25]. When examining issues of equity, diversity, and inclusion, intersectional reasoning—considering multiple overlapping attributes—is essential for achieving a truly equitable perspective. This recognition motivates to propose intersectional strategies that more effectively address these challenges in data analysis.

D. Sobriety in artificial intelligence models training and use

Sobriety in artificial intelligence (AI) model training and deployment emphasizes a measured, responsible approach that avoids hype and reckless development. Key aspects of this approach include ethical considerations, such as thoroughly evaluating potential negative impacts, incorporating safeguards and human oversight, and prioritizing fairness, transparency, and accountability; and considering also the environmental impact of those models [26]. Rigorous testing ensures AI models are validated across diverse scenarios, with ongoing monitoring and redesign if needed. Thoughtful data practices involve high-quality, representative datasets, privacy respect, and bias mitigation. Responsible resource use considers environmental impacts, optimizing efficiency, and exploring less computationally intensive methods.

Emerging approaches use several models with different sizes and performance to address queries with different performance requirements. The work in [27] proposes a novel hybrid approach to reduce the computational costs associated with deep neural networks while maintaining performance. The work proposes a cost-effective and quality-aware query routing for LLMs. The authors identify several approaches for reducing ML inference costs like model compression techniques (pruning, quantization, and knowledge distillation) and neural architecture search to optimize model architectures. Hybrid inference paradigms that use two models of different sizes to reduce costs [28], [29]. The concept involves directing simpler queries to a smaller model, reducing inference costs while maintaining response quality. A threshold based on query difficulty balances quality and cost dynamically.

E. Feminism in technologies

Techno-feminism⁵ seeks to transform the tech industry by embedding a feminist perspective at every stage of development. This involves increasing female representation in development teams, redesigning processes to meet diverse user needs, and addressing unconscious biases. It promotes inclusive practices and drives innovation, creating technologies that challenge and dismantle existing inequalities, fostering a more equitable digital future.

V. CONCLUSION AND PERSPECTIVES

Existing work on responsible resource allocation strategies for federated learning, particularly from a decolonial perspective, has several limitations. While progress has been made in fairness-aware federated learning and resource allocation, these approaches often fail to address the complex interplay of power dynamics, resource inequities, and biases in global systems. Current methods tend to focus on technical aspects like workload distribution and model aggregation, neglecting broader socio-economic issues such as reinforcing existing power structures and epistemic violence. Moreover, despite growing awareness of responsible AI practices, there is a lack of comprehensive frameworks that fully address fairness, equity, and inclusivity throughout the federated learning process. These gaps point to the need for interdisciplinary approaches and new strategies that prioritize marginalized communities and ensure equitable distribution of benefits.

Implementing these principles demands deep community engagement, interdisciplinary collaboration, and a commitment to challenging the dominant paradigms in AI/ML research and development. The ultimate goal is to create federated learning systems that are not only equitable but also empower communities rather than merely extracting value from them. Our initial studies and vision, grounded in a decolonial perspective, emphasize that fair resource allocation in federated learning requires a critical examination of power dynamics and a focus on the needs of marginalized communities. To develop effective solutions, it is essential to address several key considerations:

- *Data and Algorithms*
 - Representation and participation: Ensure diverse participation from communities that have historically been excluded from AI/ML development, especially those in the Global South. This means actively partnering with and empowering local organizations and researchers.
 - Data sovereignty: Respect indigenous data sovereignty principles. Allow communities to maintain ownership and control over their data, rather than extracting it for centralized models.
 - Pluriversal approach: Embrace multiple ways of knowing and allow for diverse, culturally-grounded approaches to machine learning rather than a one-size-fits-all model.
 - Local context: Design federated learning systems that can adapt to local contexts, languages, and knowledge systems rather than imposing Western/Global North frameworks.
- *Computing resources allocation in different target architectures*
 - Transparency: Make decision-making processes around resource allocation transparent and accountable to all participants.
 - Decolonial metrics: Develop new metrics for fairness and performance that go beyond Western notions of accuracy and efficiency to consider cultural relevance, sustainability, and social impact.

⁵Joy Buolamwini, Techno-féminisme : comment les femmes redéfinissent la technologie pour un avenir plus inclusif, Magazine 1981, <https://www.1981-magazine.fr/index.php/2024/08/21/techno-feminisme-comment-les-femmes-redefinissent-la-technologie-pour-un-avenir-plus-inclusif/> - accessed 22nd August 2024

To create a truly equitable federated learning system, it is crucial to ensure that benefits and resources are fairly distributed across all participating communities, rather than being concentrated among large tech companies or research institutions. This requires investment in building local AI and machine learning capacities, especially in underserved regions, rather than exploiting them for data and computation. Federated learning architectures must be power-aware, addressing how they can either reinforce or fairly redistribute power. Additionally, incorporating indigenous and non-Western ethical frameworks into the design and governance of these systems will ensure respect for diverse cultural perspectives, making federated learning more inclusive and responsive to all stakeholders.

ACKNOWLEDGEMENT

The work of this paper is funded by the project FRIENDLY⁶ of the program Transversal actions of the LIRIS laboratory (UMR 5205, France) bringing together the teams: BD (Genoveva Vargas Solar, CNRS), SOC (Chirine Ghedira-Guegan, iaelyon Université Lyon 3), and DRIM (Nadia Bennani, INSA Lyon).

REFERENCES

[1] G. Vargas-Solar, "Calling for a feminist revolt to decolonise data and algorithms in the age of datification," 2022. [Online]. Available: <https://arxiv.org/abs/2210.08965>

[2] M. Anisetti, C. A. Ardagna, and N. Bena, "Continuous certification of non-functional properties across system changes," in *Proc. of ICSOC 2023*, Rome, Italy, November–December 2023.

[3] S. Romdhani, G. Vargas-Solar, N. Bennani, and C. Ghedira-Guegan, "Qos-based trust evaluation for data services as a black box," in *2021 IEEE International Conference on Web Services (ICWS)*. IEEE, 2021, pp. 476–481.

[4] S. Romdhani, N. Bennani, C. Ghedira-Guegan, and G. Vargas-Solar, "Trusted data integration in service environments: a systematic mapping," in *Service-Oriented Computing: 17th International Conference, ICSOC 2019, Toulouse, France, October 28–31, 2019, Proceedings 17*. Springer, 2019, pp. 237–242.

[5] G. Vargas-Solar, M. S. Hassan, and A. Akoglu, "Jita4ds: disaggregated execution of data science pipelines between the edge and the data centre," *Journal of Web Engineering*, vol. 21, no. 1, pp. 1–26, 2022.

[6] C. A. Ardagna, N. Bena, N. Bennani, C. Ghedira-Guegan, N. Grecchi, and G. Vargas-Solar, "Revisiting trust management in the data economy: A roadmap," *IEEE Internet Computing*, 2024.

[7] Y. Shi, H. Yu, and C. Leung, "Towards fairness-aware federated learning," *IEEE Transactions on Neural Networks and Learning Systems*, 2023.

[8] L. Ju, T. Zhang, S. Toor, and A. Hellander, "Accelerating fair federated learning: Adaptive federated adam," *arXiv preprint arXiv:2301.09357*, 2023.

[9] T. Salazar, M. Fernandes, H. Araújo, and P. H. Abreu, "Fair-fate: Fair federated learning with momentum," in *International Conference on Computational Science*. Springer, 2023, pp. 524–538.

[10] Y. H. Ezzeldin, S. Yan, C. He, E. Ferrara, and A. S. Avestimehr, "Fairfed: Enabling group fairness in federated learning," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 37, no. 6, 2023, pp. 7494–7502.

[11] A. Papadaki, N. Martinez, M. Bertran, G. Sapiro, and M. Rodrigues, "Minimax demographic group fairness in federated learning," in *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency*, 2022, pp. 142–159.

[12] R. Feng, F. Calmon, and H. Wang, "Adapting fairness interventions to missing values," *Advances in Neural Information Processing Systems*, vol. 36, 2024.

[13] H. Jeong, H. Wang, and F. P. Calmon, "Fairness without imputation: A decision tree approach for fair prediction with missing values," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 36, no. 9, 2022, pp. 9558–9566.

[14] Z. Wang, M. Lu, X. Yuan, J. Zhang, and H. Van De Wetering, "Visual traffic jam analysis based on trajectory data," *IEEE transactions on visualization and computer graphics*, vol. 19, no. 12, pp. 2159–2168, 2013.

[15] T. Li, M. Sanjabi, A. Beirami, and V. Smith, "Fair resource allocation in federated learning," in *International Conference on Learning Representations*, 2020. [Online]. Available: <https://openreview.net/forum?id=ByexELSYDr>

[16] S. Jiang and J. Wu, "Multi-resource allocation in cloud data centers: A trade-off on fairness and efficiency," *Concurrency and Computation: Practice and Experience*, vol. 33, no. 6, p. e6061, 2021.

[17] H. Hamzeh, "Fairness for resource allocation in cloud computing." Ph.D. dissertation, Bournemouth University, 2021.

[18] T. Li, M. Sanjabi, and V. Smith, "Fair resource allocation in federated learning," *ArXiv*, vol. abs/1905.10497, 2019. [Online]. Available: <https://api.semanticscholar.org/CorpusID:166227978>

[19] W. Orr and K. Crawford, "The social construction of datasets: On the practices, processes, and challenges of dataset creation for machine learning," *New Media & Society*, vol. 26, no. 9, pp. 4955–4972, 2024. [Online]. Available: <https://doi.org/10.1177/14614448241251797>

[20] B. Catania, G. Guerrini, and C. Accinelli, "Fairness & friends in the data science era," *AI & SOCIETY*, vol. 38, no. 2, pp. 721–731, 2023.

[21] R. Madduri, K. Chard, M. D'Arcy, S. C. Jung, A. Rodriguez, D. Sulakhe, E. Deutsch, C. Funk, B. Heavner, M. Richards *et al.*, "Reproducible big data science: A case study in continuous fairness," *PLoS one*, vol. 14, no. 4, p. e0213013, 2019.

[22] M. Drosou, H. V. Jagadish, E. Pitoura, and J. Stoyanovich, "Diversity in big data: A review," *Big data*, vol. 5, no. 2, pp. 73–84, 2017.

[23] E. Pitoura, K. Stefanidis, and G. Koutrika, "Fairness in rankings and recommendations: an overview," *The VLDB Journal*, pp. 1–28, 2022.

[24] K. Martin, "Fairness and justice in data analytics," in *Ethics of Data and Analytics*. Auerbach Publications, 2022, pp. 249–253.

[25] T. C. Pargman, C. McGrath, O. Viberg, and S. Knight, "New vistas on responsible learning analytics: A data feminist perspective," *Journal of Learning Analytics*, vol. 10, no. 1, pp. 133–148, 2023.

[26] K. Crawford, "The atlas of ai: Power, politics, and the planetary costs of artificial intelligence," 2021.

[27] D. Ding, A. Mallick, C. Wang, R. Sim, S. Mukherjee, V. Ruhle, L. V. Lakshmanan, and A. H. Awadallah, "Hybrid llm: Cost-efficient and quality-aware query routing," *arXiv preprint arXiv:2404.14618*, 2024.

[28] D. Kang, T. Hashimoto, I. Stoica, and Y. Sun, "Scaling up trustless dnn inference with zero-knowledge proofs," *arXiv preprint arXiv:2210.08674*, 2022.

[29] D. Ding, S. Amer-Yahia, and L. V. Lakshmanan, "On efficient approximate queries over machine learning models," *arXiv preprint arXiv:2206.02845*, 2022.

⁶<http://vargas-solar.com/friendly>