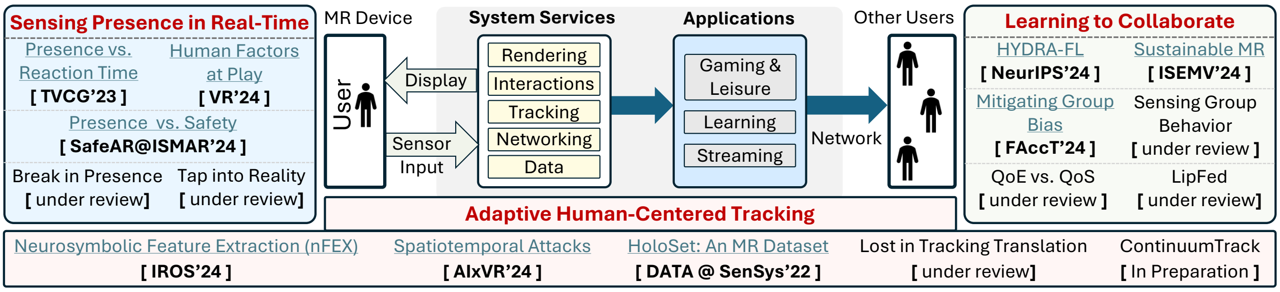
Research Statement
Yasra Chandio
yasrachandio.github.io

I am a Human-Centered Computing researcher focused on advancing Mixed Reality (MR) by measuring, interpreting, and adapting key system parameters and services to improve user immersion, safety, and overall experience. My research focuses on designing the next generation of adaptive Human-Centered MR by aligning technologies with user experiences and integrating human factors into every aspect of MR design and development, using tools and techniques from Systems, Human-Computer Interaction (HCI), and Machine Learning (ML).



MR combines physical and virtual worlds by sensing the user's environment and interactions, using tracking to align virtual elements with real-world objects accurately. The system processes this tracking data so virtual objects respond naturally to the user's actions in real-time, creating an immersive experience. This immersion creates the experiential aspect of MR: presence, the feeling of "being there" in the virtual environment (VE). MR applications place unique demands on both user experience and the underlying technology (such as tracking and real-time responsiveness), requiring the system to adapt. For instance, gaming prioritizes fast response times and high visual quality, while surgical applications require accurate tracking during critical procedures. In multi-user collaborative settings, the system must carefully manage what to share and when, balancing resources, security, privacy, and fairness to prevent a single user from disrupting the experience for others. Given the unique demands of MR applications, which must handle the unpredictability of human interactions and adapt to changing conditions, whether related to the environment, user behavior, collaboration, or resource constraints, my work focused on three main thrusts:

Experiential Thrust laid the scientific basis for understanding immersive experiences in MR by developing objective, quantifiable measures of presence, an important phenomenon that dictates human behavior and task performance in VE. These foundations guide how we build and evaluate future MR systems, shaping the design of interactions to align with cognitive processes and user perceptions of the VE. Traditional methods, often reliant on subjective questionnaires, fail to capture users' real-time, nuanced responses. In my work, I proposed new system metrics, such as reaction time, as an adaptable and objective measure of presence. Sensitive to range of factors, such as system fidelity [1], user cognitive states [2], interaction design [3], and real-world disruptions [4]. My approach enabled systems to dynamically adapt to users' cognitive states and sensory feedback [5], maintain immersion, and improve the overall experience [6].

Systemic Thrust contributed to core technological infrastructure, improving tracking accuracy, adaptability, and security in MR. Accurate, real-time tracking is essential for MR, yet current systems struggle to adapt to rapidly changing environments and unpredictable human interactions, leading to cognitive overload and reduced immersion. To address this, my research introduced Neurosymbolic Feature Extraction (nFEX) framework [7] that integrates a taxonomy of human interactions and environmental factors [8] into the tracking process. To support nFEX, I developed the HoloSet dataset [9], providing essential data for underexplored MR scenarios. Additionally, I am developing ContinuumTrack [10] to anticipate and resolve tracking failures, ensuring an uninterrupted user experience. By addressing vulnerabilities in MR systems [11], my work further boosts adaptability and security to improve the overall user experience.

Collaborative Thrust focused on fairness, privacy, security, and resource efficiency in shared MR experiences to promote equitable interactions. My contributions included developing techniques to sense and interpret individual behaviors, allowing MR systems to infer group dynamics to optimize interactions [12]. I introduced methods to mitigate biases [13–15] and enhance privacy and security [16, 17] through Federated Learning (FL), to support secure and fair data handling across multiple users. Additionally, by optimizing resource efficiency to balance quality-of-service (QoS) with quality-of-experience (QoE) [18], my work supported the sustainability of MR systems [19]. This aligns technological innovation with broader societal goals, such as decarbonization and resource conservation.

Broader Impact. While my research advances MR systems, it has far-reaching societal implications. The behavioral and safety tools I develop in my experiential and systemic thrusts have broader applications in fields like accessibility and healthcare, particularly in areas such as obesity, diabetes, and aging, aligning with National Institutes of Health (NIH) goals. My tools, designed to monitor and interpret real-time user interactions and tracking, can be applied to monitoring behaviors in neurodivergent individuals or those with autism, tracking physical activity and cognitive changes in aging populations, and supporting diabetes management by tracking health-related behavior patterns[1]. Finally, with my interdisciplinary background, I am well-qualified for CAREER programs from both NSF and NIH.

[1] NSF IIS:HCC, NSF CCF:FET, NIH BRAIN Initiative, NIH AI and Technology Collaboratory, NIH VR Technologies for Obesity and Diabetes

Present Work

I will now elaborate on my three research thrusts, clarifying their focus and how they relate to my broader research.

① **Rethinking How We Sense Presence in Real-Time** *(VR, TVCG, SafeAR@ISMAR).* This research advanced the theory of immersive environments in MR by developing new system-level experiential metrics that link human perception, cognitive state, and real-time interaction, as improving experience requires the ability to measure it accurately. First, I identified reaction time (the time it takes for a user to respond to a cue) as a key measure of presence (a subjective feeling that reflects the user's immersive experience). Though often overlooked, reaction time is readily available and offers a real-time measure of how sensory inputs like visuals and audio impact user interaction. This metric captured the impact of sensory elements on the user's experience by *linking human perception (subjective experiential element) to human actions (systemic performance indicator)* [1]. However, measuring reaction time as an indicator of presence was challenging. It required isolating reaction times in MR, where multiple sensory inputs and interactions happen simultaneously. Noise, such as background movements or unintended triggers, had to be filtered out, and measurements needed to be timed to align with the task's cognitive load and the user's actions. To address these challenges, I carefully controlled the appearance and behavior of virtual objects to create consistent conditions for accurate and repeatable reaction time measurements across different levels of presence. My human-subject studies showed that increased reaction time correlates with reduced presence.

Second, while reaction time was sensitive to changes in perception due to variations in the visual stimuli, my analysis showed that it varied in ways that could not be fully explained by sensory changes alone. This led me to investigate the *link between the human's cognitive state and their performance* [2]. After accounting for the baseline change, I found that factors like the user's familiarity with the environment or task (cognitive conditioning) further influenced the change in reaction time magnitude. These human factors shape how users process sensory information and respond in MR environments. It showed that reaction time reflects both scene fidelity and the user's cognitive state and offered a more comprehensive understanding of how presence is experienced.

After confirming the perceptual and cognitive correlation between reaction time and presence, I evaluated its validity as a presence metric by assessing how interaction design, as the sensory tool, impacts this relationship. My research showed that a designer's choice directly affects user perception, and the cognitive load required for users to learn and adapt to these interactions further shapes the user presence. With this insight, I updated my measurement framework to use reaction time as a way to understand how interactions shaped presence in MR, allowing for *assessing presence (experiential) and fine-tuning interactions (systemic) to improve the overall experience and design* [3]. I observed that familiar, intuitive interactions that mimic real-world experiences enhance presence by reducing cognitive load and reaction time, aligning with users' expectations and sensory processing. In contrast, abstract interactions increase cognitive load, leading to increased reaction times and diminished presence. Our findings emphasized the importance of *designing MR interactions that are both perceptually intuitive and cognitively manageable for users* [5].

Finally, even with finely tuned sensory inputs and interactions, interruptions in MR pose a significant challenge to maintaining a consistent presence when the user's attention shifts from the virtual to the real world. My research categorized these distractions into congruent (aligned with the task) and incongruent (disruptive). Users recover quickly from congruent distractions by rationalizing them as minor glitches or familiar events. In contrast, incongruent distractions leave users without a mental model, leading to longer reaction times and significantly reduced presence due to cognitive overload, known as break-in-presence (BIP). To ensure consistent overall MR experience, my research quantified BIP by modeling this *mediation between distraction, cognitive load, reaction time, and presence* [4], showing how incongruent distractions increase reaction times and reduce presence.

② **Scene-Driven Adaptive Human-Centered Tracking** *(IROS, AIxVR, DATA@SenSys).* One of the fundamental problems in MR is the ability of tracking systems to adapt to rapidly changing environments and unpredictable human interactions, accumulating tracking drift and misalignment. These errors lead to cognitive overload, loss of user immersion, and physical safety. In MR, users seamlessly move between different settings, indoors, outdoors, varied lighting conditions, all while interacting with both physical and virtual elements. To address this, my research focused on developing a novel approach to tracking that dynamically adapts to the complexities of human behavior and diverse settings. At the core of my approach is *Neurosymbolic Feature Extraction (nFEX), a novel adaptive framework for feature extraction in tracking* [7]. The neurosymbolic architecture avoids the rigidity of symbolic methods while retaining their interpretability and reasoning. It reduces deep learning's data dependency while integrating its learning capabilities, which are needed for adapting to new and unseen environments. I developed a domain-specific language (DSL) and a knowledge graph to structure and reason about environmental factors such as light, textures, and user interactions such as variations in motion speed and sudden starts or stops. `nFEX` integrated this domain knowledge with adaptive capabilities through a two-step neural network process, enabling dynamic optimization of feature extraction. This balance of clarity and consistency of symbolic reasoning with the flexibility and adaptability of neural networks makes neurosymbolic architecture particularly suited for MR, enabling real-time adaptation to environmental factors and unpredictable user interactions while minimizing tracking errors.

To generate this symbolic domain knowledge, I developed two foundational resources: (1) symbolic rules that systematically define MR interactions to inform *the design of tracking solutions suited to unpredictable human interactions* [8], and (2) *HoloSet dataset* [9], *to validate the unexplored human interactions in MR*. First, to apply symbolic rules to the DSL and knowledge graph, I systematically defined MR interactions by identifying and evaluating the computational, environmental, and locomotion challenges faced by tracking algorithms due to unpredictable human interactions. Mapping these challenges into the taxonomy provides a symbolic backbone for `nFEX`'s ability to adapt in real time.

Second, to validate `nFEX`'s ability to adapt to complex MR scenarios, I released HoloSet, a foundational resource with approximately 100k samples of both macro and micro-movements essential for critical applications like surgery. Each sample includes high-resolution RGB images, grayscale images, inertial data, depth images, and precise ground truth pose trajectories, providing the essential data for evaluating and improving tracking systems. Evaluations showed that `nFEX` reduced pose errors by up to 90% compared to traditional tracking methods like ORB, significantly improving system accuracy and adaptiveness in complex scenarios, including autonomous vehicles, drones, and MR environments.

Despite this added adaptability, MR systems remain vulnerable to stealthy manipulations that could compromise performance and erode user trust. MR systems are vulnerable to malicious threats, particularly as they become more integrated into everyday life. To address these risks, I explored *a novel attack surface that exploits the multimodal and spatiotemporal nature of MR tracking systems* [11, 20]. I developed techniques to manipulate visual and inertial data streams, launching undetectable attacks during stable system periods to disrupt accuracy and safety. These attacks manifest gradually, using small manipulations over extended periods to trigger tracking failures that remain undetectable by current tracking methods, highlighting how coordinated, multi-modal attacks can undermine the accuracy and safety of MR systems. Even nFEX, which primarily addresses immediate challenges like error correction and real-time adaptability to user actions and environmental changes, could miss these subtle attacks. To address long-term tracking failures caused by internal system errors, unpredictable environments, or external malicious attacks, I developed *ContinuumTrack* [10], *a system designed to proactively anticipate and resolve tracking failures before they occur*. By profiling system behavior, predicting potential tracking failures, and implementing corrective mechanisms, it maintains long-term reliability to improve safety for critical applications, complementing `nFEX`'s real-time adaptability.

③ **Learning to Collaborate** *(NeurIPS, FAccT, ISEMV, AIChallengeIoT@SenSys)*. The first part of my work on collaborative MR focused on sensing audio, video, proximity, and depth to *sense and interpret individual behaviors in a group to infer group dynamics* [12], such as identifying whether users are collaborating, competing, or working independently. I developed a system prototype that passively sensed the individual user behavior in collaborative MR using sensor data and used social network analysis tools, such as sociometry, to autonomously infer the group behavior in real time. This real-time learning helps MR systems to dynamically optimize collaboration by aligning with specific interaction patterns. However, continuous sensing and readily available data raise security and privacy concerns by potentially capturing sensitive information like private conversations or personal habits. Disproportionate data collection from users with specific speech patterns or gestures, may introduce biases, leading to misinterpretation and unfair treatment of some users, skewing collaboration dynamics. Additionally, malicious users manipulating virtual objects could inject biased data, expose private interactions, or cause security breaches, compromising the MR environment. These challenges underscore the need for strong privacy, security, and fairness measures in collaborative MR systems, which my work addresses by protecting sensitive data and supporting fair model training [13–17, 21].

Federated Learning (FL) is a widely used paradigm for developing tools for shared experiences. I made fundamental contributions in this field to lay the groundwork for future collaborative MR environments where data from multiple users must be handled securely and fairly in a privacy-preserving manner. I *developed techniques using regularization and weighting to mitigate biases* [13, 14, 21] while preserving privacy, and also *developed techniques to combat model poisoning attacks* [16, 17] by adjusting model training and redistributing knowledge distillation across model layers. Finally, beyond addressing experiential challenges like user behavior, security, fairness, and privacy, optimizing the systemic performance of collaborative MR requires focusing on resource efficiency. Network factors like bandwidth and latency are key for maintaining a seamless sense of presence; my work focused on the *critical trade-offs between QoS and QoE* [18], essential for enabling smooth interactions in resource-constrained environments. This balance is crucial to *minimize the carbon footprint and enhance resource efficiency without sacrificing performance* [19]. By gracefully navigating these tradeoffs, I developed MR systems that are not only high-performing but also environmentally sustainable, supporting meaningful virtual experiences with broader societal impact.

Research Vision

Grounded in an extensive theoretical background and informed by human subjects research, my research will advance MR and adjacent human-centered systems by developing prototypes that integrate user behavioral state estimation and physical tracking with collaborative model training. My research will create new tools, introduce novel control parameters and services, and use existing data in previously unexplored ways. Emphasizing safety, well-being, and fairness, with implications for societal goals like climate change and health, I aim to create adaptive, immersive experiences scalable across diverse environments and hardware platforms.

Ⓧ **Tracking with Real-Time Human Input** This research will focus on integrating human expertise (knowledge) and human experience (user interaction, sensory perception, cognitive load) into tracking systems, not just for real-time feedback but as a catalyst for scaling expertise and fostering autonomy in complex environments. Humans can provide critical insights for systems, such as in 3D reconstruction, where algorithms often struggle with complex geometries or occlusions. An expert can manually flag problematic areas and offer valuable context by identifying more significant occlusions. This perception will allow the system to refine its parameters and improve accuracy. While human input can be slow and costly, I aim to use it as a foundation for scaling expert knowledge and enabling real-time adaptation.
In the **short-term**, I will develop interfaces and algorithms for real-time collaboration between humans and tracking systems in areas such as healthcare, autonomous vehicles, and MR. These interfaces will also capture human experience metrics (perception, cognition) to refine real-time system adjustments and serve as training grounds for the system.
The **long-term goal** is to create systems that learn from human input and progressively reduce dependence on human intervention. I will scale this process, enabling the system to self-sustain and self-correct in real time. With ML and transfer learning, I will develop systems that can autonomously adapt to new and unseen scenarios.

This interdisciplinary research expands my adaptive tracking work by introducing fundamentally new dimensions and a broader scope that incorporates real-time human collaboration. I will collaborate with experts in 3D visualization and computer vision to develop advanced interfaces and algorithms for MR systems and demonstrate the value of human-in-the-loop approaches in improving visual system accuracy, adaptability, and safety. Additionally, I will build multimodal transfer learning frameworks to scale across diverse environments, preserving human reasoning insights and interpretability while processing multimodal data from sensors and human input.

Ⓧ **Robust User State Estimation For Safety and Immersiveness** In this work, I will focus on managing user states to balance immersion and safety in immersive environments. To achieve this, the system should non-intrusively and continuously monitor the user's cognitive and physical state in real-time by combining explicit feedback, such as reaction time, with implicit signals like eye gaze shifts and voice interactions. It can leverage implicit cues, such as subtle changes in the environment detected by the system, and implicit interactions, like spontaneous gaze shifts or verbal reactions, that provide non-intrusive insights into user presence and experience. These cues and interactions can inform how the system assesses user state, allowing it to make real-time adjustments that maintain an immersive experience while protecting against risks like cognitive overload or disorientation. My research goal is to balance user immersion and safety, providing an optimized, responsive experience without compromising well-being.
The **short-term goal** is to develop systems that monitor explicit feedback and implicit cues to assess user states in real time. These systems will use this data to dynamically adjust immersion levels to ensure that users remain engaged without being overwhelmed, particularly in healthcare and emergency response training simulations where cognitive overload or physical disorientation poses a critical risk. The question I aim to answer is how to account for individual variability in users' responses and tailor safety interventions to meet each user's needs.
The **long-term goal** is to develop systems that autonomously adapt to user behavior by learning from both explicit feedback and implicit cues. A key challenge is mitigating variability in user responses, such as differences in gaze patterns or speech recognition, while ensuring consistency across diverse MR platforms. Addressing this challenge required foundational work in areas like eye tracking and voice recognition systems, which are essential for accurately interpreting implicit user behaviors and adapting the system to individual needs.

Building on my work in quantifying presence, I aim to refine how MR systems interpret implicit user behaviors, creating adaptive systems across applications like training simulations and entertainment. This includes fundamental theoretical contributions, system prototypes, and human subject studies in HCI, Systems, and ML. I plan to continue my collaboration with cognitive science experts and begin working with neuroscience experts to integrate complex human behaviors, infer the causality of these behaviors, and align technical innovation with human-centered design.

Ⓧ **Collaborative Model Training** My research will address challenges in collaborative model training in MR, focusing on fairness, bias detection, protection against malicious collaborators, and resource efficiency. The main goal is to design systems that maintain balanced shared MR experiences, preventing any one user from dominating or negatively affecting others. The secondary goal is to scale these systems across diverse hardware environments while maintaining fairness and minimizing energy consumption and environmental impact.
The **short-term goal** is to design algorithms resilient to fairness concerns that can detect bias and mitigate the risk of malicious actors or harmful contributions in MR. This will involve developing strategies and metrics for safety and fairness so that incorrect or biased inputs do not propagate and that all users have transparency in their contributions.
My **long-term goal** is to develop scalable, resource-efficient systems that can support large collaborative environments across diverse hardware platforms, such as different MR headsets or integrating other wearables in this ecosystem. I aim to optimize resource efficiency by integrating energy-efficient algorithms that address broader societal objectives, like climate change mitigation. These systems will autonomously manage interactions, minimize environmental impact, and support fairness, transparency, and trust across a wide range of MR and other human-in-the-loop systems.

These contributions will advance both technical innovation and environmental responsibility, addressing societal challenges like climate change while pushing the boundaries of distributed machine learning and collaborative systems.

References

[1] **Yasra Chandio**, Noman Bashir, Victoria Interrante, and Fatima Anwar. "Investigating the Correlation Between Presence and Reaction Time in Mixed Reality". In: *IEEE Transactions on Visualization and Computer Graphics (TVCG)*. 2023 [PDF].

[2] **Yasra Chandio**, Victoria Interrante, and Fatima Anwar. "Human Factors At Play: Understanding The Impact Of Conditioning On Presence And Reaction Time In Mixed Reality". In: *IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. selected to appear in IEEE TVCG, (top 12% papers). 2024 [PDF].

[3] **Yasra Chandio**, Victoria Interrante, and Fatima Anwar. "Tap into Reality: Understanding the Impact of Interactions on Presence and Reaction Time in Mixed Reality". In: submission.

[4] **Yasra Chandio**, Victoria Interrante, and Fatima Anwar. "Reaction Time as a Proxy for Presence in Mixed Reality with Distraction". In: submission.

[5] **Yasra Chandio**, Victoria Interrante, and Fatima Anwar. "Towards Objective And Implicit Presence Measures In Mixed Reality". In: submission [short paper].

[6] **Yasra Chandio**, Victoria Interrante, and Fatima Anwar. "Balancing Presence And Safety Using Reaction Time In Mixed Reality". In: *Workshop at IEEE International Symposium on Mixed and Augmented Reality (SafeAR@ISMAR)*. 2024 [PDF].

[7] **Yasra Chandio**, Momin Ahmed Khan, Khotso Selialia, Luis Antonio Garcia, Joseph DeGol, and Fatima Anwar. "A Neurosymbolic Approach To Adaptive Feature Extraction In SLAM". In: *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. 2024 [PDF].

[8] **Yasra Chandio**, Khotso Selialia, Luis Antonio Garcia, Joseph DeGol, and Fatima Anwar. "Lost in Tracking Translation: A Comprehensive Analysis of Visual SLAM in Human-Centered XR and IoT Ecosystems". In: submission.

[9] **Yasra Chandio**, Noman Bashir, and Fatima Anwar. "Dataset: HoloSet - A dataset for visual inertial pose tracking in extended reality". In: *5th International SenSys/BuildSys Workshop on Data (DATA)*. 2022 [PDF].

[10] **Yasra Chandio**, Luis Antonio Garcia, Joseph DeGol, and Fatima Anwar. "ContinuumTrack: Adaptive Real-Time Profiling and Correction in Dynamic SLAM Systems". In: preparation.

[11] **Yasra Chandio**, Noman Bashir, and Fatima Anwar. "Stealthy and practical multi-modal attacks on mixed reality tracking". In: *IEEE International Conference on Artificial Intelligence & extended and Virtual Reality (AIxVR)*. Best Paper Finalist. 2024 [PDF].

[12] Diana Romero*, **Yasra Chandio***, Fatima Anwar, and Salma Elmalaki. "Sensing Group Behavior In Mixed Reality". In: submission [* co-primary author].

[13] Khotso Selialia, **Yasra Chandio**, and Fatima Anwar. "Mitigating feature heterogeneity biases in Federated Learning". In: *ACM Conference on Fairness, Accountability, and Transparency (FAccT)*. 2024 [PDF].

[14] Khotso Selialia, **Yasra Chandio**, Jimi Oke, and Fatima Anwar. "LipFed: Mitigating Subgroup Bias in Federated Learning with Lipschitz Constraints". In: submission.

[15] Khotso Selialia, **Yasra Chandio**, and Fatima Anwar. "Federated Learning Biases in Heterogeneous Edge-Devices - A Case-Study". In: *4th International SenSys/BuildSys Workshop on Challenges in Artificial Intelligence and Machine Learning for Internet of Things (AIChallengeIoT)*. 2022 [PDF].

[16] Momin Ahmed Khan, **Yasra Chandio**, and Fatima Anwar. "HYDRA-FL: Hybrid Knowledge Distillation for Robust and Accurate Federated Learning". In: *Annual Conference on Neural Information Processing Systems (NeurIPS)*. 2024 [PDF].

[17] Momin Ahmad Khan, Virat Shejwalkar, **Yasra Chandio**, Amir Houmansadr, and Fatima Anwar. "A Critical Analysis Of The Pitfalls In Federated Learning Defense Evaluations". In: submission.

[18] John Murray, **Yasra Chandio**, Fatima Anwar, and Michael Zink. "Exploring Relationship Between Quality Of Experience And Quality Of Service In Collaborative Mixed Reality". In: preparation.

[19] **Yasra Chandio**, Noman Bashir, Tian Guo, Elsa Olivetti, and Fatima Anwar. "Scoping The Decarbonization Of Collaborative Mixed Reality". In: *IEEE International Symposium on Emerging Metaverse (ISEMV)*. 2024 [PDF].

[20] **Yasra Chandio** and Fatima Anwar. "Poster: Spatiotemporal Security in Mixed Reality systems". In: *18th ACM Conference on Embedded Networked Sensor Systems (SenSys)*. 2020 [PDF].

[21] Khotso Selialia, **Yasra Chandio**, Jimi Oke, and Fatima Anwar. "Tackling Group Fairness Inconsistencies in Federated Learning". In: preparation.