

FL4NLP 2022

**The First Workshop on Federated Learning for Natural  
Language Processing**

**Proceedings of the Workshop**

May 27, 2022

The FL4NLP organizers gratefully acknowledge the support from the following sponsors.

**Gold**



©2022 Association for Computational Linguistics

Order copies of this and other ACL proceedings from:

Association for Computational Linguistics (ACL)  
209 N. Eighth Street  
Stroudsburg, PA 18360  
USA  
Tel: +1-570-476-8006  
Fax: +1-570-476-0860  
[acl@aclweb.org](mailto:acl@aclweb.org)

ISBN 978-1-955917-37-7

# Introduction

Due to increasing concerns and regulations about data privacy (e.g., General Data Protection Regulation), coupled with the growing computational power of edge devices, emerging data from realistic users have become much more fragmented, forming distributed private datasets across different clients (i.e., organizations or personal devices). Respecting users' privacy and restricted by these regulations, we have to assume that users' data in a client are not allowed to transfer to a centralized server or other clients. For example, a hospital does not want to share its private data (e.g., conversations, questions asked on its website/app) with other hospitals. This is despite the fact that models trained by a centralized dataset (i.e., combining data from all clients) usually enjoy better performance on downstream tasks (e.g., dialogue, question answering). Therefore, it is of vital importance to study NLP problems in such a scenario, where data are distributed across different isolated organizations or remote devices and cannot be shared for privacy concerns.

The field of federated learning (FL) aims to enable many individual clients to jointly train their models, while keeping their local data decentralized and completely private from other users or a centralized server. A common training schema of FL methods is that each client sends its model parameters to the server, which updates and sends back the global model to all clients in each round. Since the raw data of one client has never been exposed to others, FL is promising to be an effective way to address the above challenges, particularly in the NLP domain where many user-generated text data contain sensitive, personal information.

Our workshop website is <https://fl4nlp.github.io/>.

## **Organizing Committee**

### **Organizing Committee**

Bill Yuchen Lin, University of Southern California  
Chaoyang He, University of Southern California  
Chulin Xie, University of Illinois Urbana-Champaign  
Fatemehsadat Miresghallah, University of California San Diego  
Ninareh Mehrabi, University of Southern California  
Tian Li, Carnegie Mellon University  
Mahdi Soltanolkotabi, University of Southern California  
Xiang Ren, University of Southern California

## **Program Committee**

### **Program Chairs**

Bill Yuchen Lin, University of Southern California  
Chaoyang He, University of Southern California  
Tian Li, Carnegie Mellon University  
Ninareh Mehrabi, University of Southern California  
Fatemehsadat Miresghallah, University of California, San Diego  
Xiang Ren, University of Southern California  
Mahdi Soltanolkotabi, University of Southern California  
Chulin Xie, University of Illinois, Urbana Champaign

### **Program Committee**

Hongyuan Zhan  
Anit Kumar Sahu  
Bahareh Harandizadeh  
Basak Guler  
Dimitris Stripelis  
Eugene Bagdasaryan  
Farzin Haddadpour  
Gerald Penn  
Hongyi Wang  
Jinhyun So  
Jun Yan  
Kshitiz Malik  
Kevin Hsieh  
Ninareh Mehrabi  
Roozbeh Yousefzadeh  
Saurav Prakash  
Shen Li  
Shengyuan Hu  
Sijie Cheng  
Sunwoo Lee  
Tao Yu  
Umang Gupta  
Xin Dong  
Xuechen Li  
Yae Jee Cho  
Zheng Xu

### **Invited Speakers**

Salman Avestimehr, University of Southern California  
Virginia Smith, CMU  
Bo Li, UIUC  
Tong Zhang, HKUST  
Manzil Zaheer, Google DeepMind

Rahul Gupta, Amazon Alexa

## Table of Contents

### *ActPerFL: Active Personalized Federated Learning*

Huili Chen, Jie Ding, Eric William Tramel, wushuan@amazon.com wushuan@amazon.com, Anit Kumar Sahu, Salman Avestimehr and taozhng@amazon.com taozhng@amazon.com . . . . . 1

### *Scaling Language Model Size in Cross-Device Federated Learning*

Jae Hun Ro, Theresa Breiner, Lara McConnaughey, Mingqing Chen, Ananda Theertha Suresh, Shankar Kumar and Rajiv Mathews . . . . . 6

### *Adaptive Differential Privacy for Language Model Training*

Xinwei Wu, li Gong and Deyi Xiong . . . . . 21

### *Intrinsic Gradient Compression for Scalable and Efficient Federated Learning*

Luke Melas-Kyriazi and Franklyn Wang . . . . . 27



# Program

## Friday, May 27, 2022

09:00 - 09:10	<i>Opening Remarks</i>
09:10 - 10:05	<i>Invited Speaker 1</i>
10:05 - 11:00	<i>Invited Speaker 2</i>
11:05 - 12:00	<i>Invited Speaker 3</i>
12:00 - 12:30	<i>Lunch Break</i>
12:35 - 12:50	<i>Paper Presentation Slot 1</i>
12:50 - 13:05	<i>Paper Presentation Slot 2</i>
13:05 - 13:20	<i>Paper Presentation Slot 3</i>
13:20 - 13:35	<i>Paper Presentation Slot 4</i>
13:35 - 14:30	<i>Invited Speaker 4</i>
14:30 - 15:25	<i>Invited Speaker 5</i>
15:25 - 16:20	<i>Invited Speaker 6</i>
16:20 - 16:35	<i>Paper Presentation Slot 5</i>
16:35 - 16:50	<i>Paper Presentation Slot 6</i>
16:50 - 17:05	<i>Paper Presentation Slot 7</i>
17:05 - 17:20	<i>Paper Presentation Slot 8</i>
17:20 - 17:35	<i>Paper Presentation Slot 9</i>
17:35 - 18:30	<i>Panel Discussion (TBA)</i>

**Friday, May 27, 2022 (continued)**

18:40 - 18:30      *Closing Remarks*