

IEEE BigData 2024 Program Schedule

Washington, DC, USA

Conference time zone: US Eastern Standard Time

The link to download the BigData 2024 conference proceedings for BigData24 and access the pre recorded videos is:

Meeting Room

Floor	Room
Lobby	CONGRESSIONAL E
Lobby	CONGRESSIONAL A
Lobby	CONGRESSIONAL B
Lobby	CONGRESSIONAL C/D
Lobby	CAPITOL ROOM A
Lobby	CAPITOL ROOM B
Ballroom	BUNKER HILL
Ballroom	LEXINGTON
Ballroom	CONCORD
Ballroom	COLUMBIA A
Ballroom	COLUMBIA B
Ballroom	COLUMBIA C
Ballroom	REGENCY BALLROOM - A
Ballroom	REGENCY BALLROOM - B
Ballroom	REGENCY BALLROOM - C
Ballroom	REGENCY BALLROOM - D
2nd floor	YELLOWSTONE
2nd floor	EVERGLADES
2nd floor	GRAND TETON
2nd floor	BRYCE
2nd floor	YOSEMITE
2nd floor	GLACIER
2nd floor	SEQUOIA
2nd floor	OLYMPIC

Sponsored by

Platinum Sponsor:



U.S. National
Science
Foundation

Gold Sponsor:



VIRGINIA
TECH.

Silver Sponsor:

IEEE *DataPort*TM

ankura 

IEEE Big Data 2024 Program Schedule

Washington DC, USA

December 15 - December 18, 2024

Keynote Lecture: **60 minutes** (50-min talk + 10-min Q/A)

Main conf. regular: **25 mins** (20-minute talk + 5-min Q/A)

Main conf. short: **15 minutes** (11-min talk + 4-min Q/A)

Registration: Saturday, Dec 14 16:00-18:00 PM

(Location: Regency Foyer)

Day 1: Sunday, December 15, 2024			
Time	Location	Workshop, Tutorial, Special Symposium, Special Session	Chairs
7:30-17:00	Regency Foyer	Registration	
10:00-10:30	Columbia Foyer	Coffee Break	
12:30-14:00	Lunch (on your own)		
16:00-16:30	Columbia Foyer	Coffee Break	
8:30-12:30	Congressional - A	Workshop #15: The 4th International Workshop on Multi-Modal Medical Data Analysis	Bo Zhang
14:00-18:00	Congressional - A	Workshop #48: DCAI: The 3rd International Workshop on Data-Centric AI	Yanjie Fu
8:30-12:30	Congressional - B	Workshop #51: Big Data and Artificial Intelligence for Personalized Health Care	Ryan Benton
9:00-17:00	Congressional - C/D	Workshop #13: The 7th Workshop on Benchmarking, Performance Tuning and Optimization for Big Data Applications (BPOD)	Zhiyuan Chen
8:30-12:30	Capitol Room - A	Workshop #17: The 3rd Workshop on Big Data Analytic in Healthcare	Alex Kuo
14:00-18:00	Capitol Room - A	Workshop #20: Next-Generation Big Graph Learning and Analytics (Next-Gen-BiGLA)	Xiaofei Zhang

9:00-17:00	Capitol Room - B	Workshop #23: The 5th International Workshop on Big Data & AI Tools, Methods, and Use Cases for Innovative Scientific Discovery (BTSD) 2024	Sangkeun Lee
8:30-12:30	Concord	Workshop #36: Second International Workshop for securing IoT systems using BigData	Cristian Bucur
14:00-18:00	Concord	Workshop #9: 4th Workshop on Knowledge Graphs and Big Data In conjunction with the IEEE Big Data 2024	Yuan An
9:00-17:00	Columbia - A	Workshop #46: 7th Workshop on Big Data for CyberSecurity (BigCyber-2024)	Lavanya Elluri
9:00-17:00	Columbia - B	Workshop #38: Workshop on Big Data and AI for Healthcare	Ping Chen
8:30-12:30	Columbia - C	Workshop #10: International Workshop on Data Engineering and Modeling for AI (DEMAI)	Carlos Ordonez
14:00-18:00	Columbia - C	Workshop #39: The Seventh Workshop on Big Data for Economic and Business Forecasting	Wei Shang
9:00-17:00	Regency - B	High School and Undergrad Symposium	Xuan Wang & Yanjie Fu
9:00-17:00	Regency - C	Special Session: Federated Learning on Big Data	Francesco Piccialli
9:00-17:00	Regency - D	REU Symposium	Xuechen Zhang, Xinhui Zhao, and Xiaokun Yang
9:00-17:00	Yellowstone	Workshop #2: 7th Annual Workshop on Cyber Threat Intelligence and Hunting (CyberHunt 2024)	Vasileios Mavroeidis
9:00-17:00	Everglades	Workshop #7: 2nd Workshop on AI Music Generation with AI Music Competition (AIMG 2024)	Callie Liao
9:00-17:00	Grand Teton	Workshop #3: 2nd workshop on Big Data Applications for Fight against Crime and Terrorism (BDA4FCT)	Georgios Stavropoulos
9:00-17:00	Bryce	Workshop #5: The 9th International Workshop on Application of Big Data for Computational Social Science (ABCSS2024)	Isamu Okada
8:30-12:30	Yosemite	Workshop #4: 8th International Workshop on Applications of Big Data Methods and	John Easton

		Technology in the Transport Industry	
14:00-18:00	Yosemite	Workshop #47: 9th Workshop on Advances in High Dimensional (AdHD) Big Data	Aristidis Vrahatis
9:00-17:00	Sequoia	Special Session: Intelligent Data Mining	Uraz Yavanoglu

Day 2: Monday, December 16, 2024			
8:00-17:30	Registration (Regency Foyer)		
8:00-8:30	Poster Setup		
8:20-9:00	Opening Ceremony Conference Chairs: Chang-Tien Lu (Virginia Tech) Fusheng Wang (Stony Brook U.) PC Chairs: Wei Ding (U. of Massachusetts, Boston) Liping Di (George Mason U.) Ricardo Baeza-Yates (Northeastern U.) Steering Committee Chair: Xiaohua (Tony) Hu (Drexel U.) Location: Regency Ballroom		
9:00-10:00	Keynote Speech I Dr. Haixun Wang, Instacart, USA Generative Information Retrieval and E-commerce Location: Regency Ballroom		
10:00-10:30	Columbia Foyer	Coffee Break	
10:00-10:30	Regency Foyer Columbia Foyer	Poster Display (Session Chairs: Kaiqun Fu, Yufei Tang)	
Time	Location	Session/Workshop	Session Chair
10:30-12:30	Capital A	D1S1: Foundation Models for Big Data	
10:30-12:30	Capital B	D1S2: Big Data Learning and Analytics I	
10:30-12:30	Columbia A	D1S3: Big Data Infrastructure I	

10:30-12:30	Columbia B	D1S4: Big Data Science and Foundations I	
10:30-12:30	Congressional A	D1S5: Big Data Management I	
10:30-12:30	Congressional B	D1S6: Big Data Search and Mining I	
9:00-17:00	Congressional C/D	Workshop #42: From Theory to Practice: Workshop on Large Language and Foundation Models	Tania Cerquitelli
10:30-12:30	Yellowstone	Challenge Cup: Predicting Chess Puzzle Difficulty	Dominik Ślęzak
14:00-18:00	Yellowstone	Workshop #8: 4th IEEE Big Data Workshop on Multimodal AI (MMAI 2024)	Lindi Liao
14:00-18:00	Everglades	Workshop #34: The 8th Annual Workshop on Applications of Artificial Intelligence in the Legal Industry	Han Qin
14:00-18:00	Grand Teton	Workshop #6: 10th Workshop on Methods to Improve Big Data Science Projects (MIDP-2024)	Jeffrey Saltz
8:30-12:30	Yosemite	Workshop #14: 2nd WS on Large-scale Data Utilization in Economics of Information and Management Sciences: Theory, Computation, and Experiment	Naoki Watanabe
10:30-12:30	Columbia C	I&G-1: Generative AI	
10:30-12:30	Bryce	SS: Understanding New Markets by Data Science, Social Science, and Economics	Teruaki Hayashi
10:30-12:30	Glacier	SS: Federated Learning on Big Data	Francesco Piccialli
12:30-14:00	Lunch		
14:30-16:00	Regency Ballroom	Panel: Big Data and AI: How they benefit each other	Wei Ding
14:00-16:00	Capital A	D1S7: Data Ecosystem I	
14:00-16:00	Capital B	D1S8: Big Data Learning and Analytics II	
14:00-16:00	Columbia A	D1S9: Big Data Learning and Analytics III	

14:00-16:00	Columbia B	D1S10: Deep Learning I	
14:00-16:00	Congressional A	D1S11: Machine Learning Algorithms I	
14:00-16:00	Congressional B	D1S12: Deep Learning II	
14:00-16:00	Columbia C	I&G-2: Big data Infra & Multi-modality	
14:00-16:00	Concord	I&G-3: Modeling	
14:00-16:00	Bryce	SS: Understanding New Markets by Data Science, Social Science, and Economics	Teruaki Hayashi
14:00-16:00	Glacier	SS: Federated Learning on Big Data	Francesco Piccialli
16:00-16:30	Columbia Foyer	Coffee Break	
16:00-16:30	Regency Foyer Columbia Foyer	Poster Display (Session Chairs: Kaiqun Fu, Yufei Tang)	
16:30-18:00	Capital A	D1S13: Big Data Applications I	
16:30-18:00	Capital B	D1S14: Big Data Applications II	
16:30-18:00	Columbia A	D1S15: Big Data Applications III	
16:30-18:00	Columbia B	D1S16: Big Data Analytics I	
16:30-18:00	Congressional A	D1S17: Big Data Analytics II	
16:30-18:00	Congressional B	D1S18: Big Data Analytics III	
16:30-18:00	Bryce	SS: Understanding New Markets by Data Science, Social Science, and Economics	Teruaki Hayashi
16:30-18:00	Glacier	SS: Federated Learning on Big Data	Francesco Piccialli
19:00-21:00		OC Business Meeting	

Day 3: Tuesday, December 17, 2024

8:30-17:00	Registration (Regency Foyer)		
8:30-9:00	PhD Forum Poster Setup		
9:00-10:00	Keynote Speech II Prof. Cornelia Caragea, USA and U. of Illinois Chicago, USA Improving Semi-Supervised Learning with Pseudo-Margins Location: Regency Ballroom		
10:00-10:30	Columbia Foyer	Coffee Break	
10:00-10:30	Regency Foyer Columbia Foyer	PhD Forum Poster Display (Session Chairs: Feng Chen, Shuhan Yuan)	
Time	Location	Session/Workshop	Session Chair
10:30-12:00	Regency Ballroom	Panel B: Big Data and AI – Shaping the Next Frontier of Data-Driven Intelligence	Duoduo (Lindi) Liao
10:30-12:30	Capital A	D2S1: Foundation Models and Analytics	
10:30-12:30	Capital B	D2S2: Data Ecosystem II	
10:30-12:30	Columbia A	D2S3: Big Data Infrastructure II	
10:30-12:30	Columbia B	D2S4: Big Data Science and Foundations II	
10:30-12:30	Congressional A	D2S5: Big Data Search and Mining II	
10:30-12:30	Congressional B	D2S6: Big Data Applications IV	
9:00-17:00	Congressional C/D	Workshop #1: 9th Computational Archival Science (CAS) Workshop	Mark Hedges
10:30-12:30	Columbia C	I&G-4: Business and Energy	
10:30-12:30	Yosemite	Special Session SS-5: Synergizing Mobility Data for Creating and Discovering Valuable Places	Yukio Ohsawa
10:30-12:30	Grand Teton	Special Session: Social Cognitive Computing in Digital Education and Learning	Jerry Chun-Wei Lin
10:30-12:30	Concord	PhD Forum	Feng Chen, Shuhan Yuan

12:30-14:00	Lunch		
14:30-16:00	Regency Ballroom	Panel	Liping Di
14:00-16:00	Capital A	D2S7: Big Data Search and Mining III	
14:00-16:00	Capital B	D2S8: Algorithms and Systems for Big Data Search	
14:00-16:00	Columbia A	D2S9: Link and Graph Mining I	
14:00-16:00	Columbia B	D2S10: Link and Graph Mining II	
14:00-16:00	Congressional A	D2S11: Machine Learning Algorithms II	
14:00-16:00	Congressional B	D2S12: Machine Learning Algorithms III	
14:00-16:00	Columbia C	I&G-5: Recommender Systems	
14:00-16:00	Everglades	I&G-6:AI4Science	
14:00-16:00	Yosemite	Special Session SS-5: Synergizing Mobility Data for Creating and Discovering Valuable Places	Yukio Ohsawa
14:00-16:00	Grand Teton	Special Session: Data-Driven Designation and Implementation of Automated Guided Vehicles	Jerry Chun-Wei Lin
14:00-17:30	Concord	PhD Forum	
16:00-16:30	Columbia Foyer	Coffee Break	
16:00-16:30	Regency Foyer Columbia Foyer	PhD Forum Poster Display (Session Chairs: Feng Chen, Shuhan Yuan)	
17:30-18:00		Walk to the National Museum of American History 1300 Constitution Ave. NW, Washington, DC 20560	
18:00-21:00	National Museum of American History	Banquet Award Ceremony Social Program	

Day 4: Wednesday, December 18, 2024			
8:30-17:00	Registration (Regency Foyer)		
9:00-10:00	Keynote Speech III Prof. Alexander "Sasha" Rush, Cornell Tech, USA Location: Regency Ballroom		
10:00-10:30	Columbia Foyer	Coffee Break	
Time	Location	Session/Workshop	Session Chair
10:30-12:30	Capital A	D3S1: Big Data Science and Foundations III	
10:30-12:30	Capital B	D3S2: Big Data Infrastructure III	
10:30-12:30	Columbia A	D3S3: Machine Learning Algorithms IV	
10:30-12:30	Columbia B	D3S4: Predictive Analytics	
10:30-12:30	Congressional A	D3S5: Machine Learning Algorithms V	
10:30-12:30	Congressional B	D3S6: Learning from Complex Data	
10:30-12:30	Yosemite	SS-6: Special Session on Dataspaces and DFFT	Noboru Koshizuka
10:30-12:30	Grand Teton	SS: HealthCare Data	Ozgun Pinarer
12:30-14:00	Lunch		
14:00-16:00	Capital A	D3S7: Multimodal and Scalable Big Data	
14:00-16:00	Capital B		
14:00-16:00	Columbia A		
14:00-16:00	Columbia B		
8:30-12:30	Columbia C	Workshop #24: 2nd International Workshop on Big Data Analytics with Artificial Intelligence for Climate Change	Rajiv Misra
14:00-16:00	Congressional A		
14:00-16:00	Congressional B		

14:00-16:00	Yosemite	SS-6: Special Session on Dataspaces and DFFT	Noboru Koshizuka
14:00-16:00	Grand Teton	Special Session: HealthCare Data	Ozgun Pinarer
16:00-16:30	Columbia Foyer	Coffee Break	
17:00-17:30	Regency Ballroom	Closing Ceremony	

Industry and Government Program Sessions

I&G-1: Generative AI Dec. 16, Monday, 10:30-12:30, Columbia C	
Paper 2 (N221) LLM Chain Ensembles for Scalable and Accurate Data Annotation (regular)	
Paper 6 (N257) Experimental Study of In-Context Learning for Text Classification and Its Application to Legal Document Review in Construction Delay Disputes (regular)	
Paper 9 (IDNet) A Novel Identity Document Dataset via Few-Shot and Quality-Driven Synthetic Data Generation (regular)	
Paper 23 (N233) AUTORED: Automated Attack Scenario Generation Framework for Red Teaming of LLMs (short)	
Paper 26 (N253) Generating Phishing Attacks and Novel Detection Algorithms in the Era of Large Language Models (short)	
Paper 28 (N237) Enriching Skill Taxonomies through Vector Space Models (short)	
I&G-2: Big data Infra & Multi-modality Dec. 16, Monday, 14:00-16:00, Columbia C	
Paper 1 (N239) DIVERSE: A Dataset of YouTube Video Comment Stances with a Data Programming Model (regular)	
Paper 15 (N209) INFA-FinOps for Cloud Data Integration (regular)	
Paper 20 (N218) Capturing Infrastructure Interdependencies for Power Outages Prediction During Extreme Events (regular)	
Paper 29 (N204) GPURank: A Cloud GPU Instance Recommendation System (short)	
Paper 32 (N271) Addressing Spend Leakage and Optimization of Cloud Costs (short)	
Paper 33 (N262) Forecasting Application Counts in Talent Acquisition Platforms: Harnessing Multimodal Signals using LMs (short)	

I&G-3: Modeling Dec. 16, Monday, 14:00-16:00, Concord	
Paper 5 (N278) Data-Prep-Kit: getting your data ready for LLM application development (regular)	
Paper 10 (N272) WeightedMMoE: Weighted Multi-Gate Mixture-of-Experts for Multi-Source Unsupervised Domain Adaptation (regular)	
Paper 12 (N220) HLAT: High-quality Large Language Model Pre-trained on AWS Trainium (regular)	
Paper 14 (N206) Active Learning for Practical Misinformation Classification in Social Media: a Case Study on COVID-19 (regular)	
Paper 18 (N245) Machine Learning (ML) Classifier to Assist in Metadata Creation (regular)	
Paper 34 (N236) An explainable framework based on counterfactual explanations for multi-class financial distress prediction of small and medium enterprises) (short)	
I&G-4: Business and Energy Dec. 17, Tuesday, 10:30-12:30, Columbia C	
Paper 3 (N235) Targeted Detection of Anomalous Merchants on Integrated Payment Platforms via Multifaceted Transaction Representation Learning (regular)	
Paper 4 (N203) End-to-End Trip Energy Predictions Developed from Real-World Driving Data using Recurrent Neural Networks (regular)	
Paper 7 (N254) Counter Data Paucity through Adversarial Invariance Encoding: A Case Study on Modeling Battery Thermal Runaway (regular)	
Paper 17 (N244) ProbSAINT: Probabilistic Tabular Regression for Used Car Pricing (regular)	
Paper 30 (N266) Forecasting Migration Patterns and Land Border Encounters (short)	
I&G-5: Recommender systems Dec. 17, Tuesday, 14:00-16:00, Columbia C	

Paper 8 (N256) Multi-task Recommendation in Marketplace via Knowledge Attentive Graph Convolutional Network with Adaptive Contrastive Learning (regular)	
Paper 16 (N216) Optimal Transport for Efficient, Unsupervised Anomaly Detection on Industrial Data (regular)	
Paper 21 (N213) GRAINRec: Graph and Attention Integrated Approach for Real-Time Session-Based Item Recommendations (regular)	
Paper 24 (N224) Exploring Query Understanding for Amazon Product Search (short)	
Paper 27 (N259) CryptoPulse: Short-Term Cryptocurrency Forecasting with Dual-Prediction and Cross-Correlated Market Indicators (short)	
Paper 35 (N229) Harnessing the Power of Graph Neural Networks for Personalized Rail Recommendations (short)	
I&G-6: AI4Science Dec. 17, Tuesday, 14:00-16:00, Everglades	
Paper 11 (N252) Enabling Scalable Mineral Exploration: Self-Supervision and Explainability (regular)	
Paper 13 (N265) Accurate and Interpretable Radar Quantitative Precipitation Estimation with Symbolic Regression (regular)	
Paper 19 (N243) Monitoring Recombinant Protein Titer in Escherichia Coli Fermentations Using Filtering Techniques (regular)	
Paper 22 (N268) Detecting Fraud in a Large Anonymized Voter Registration Dataset (short)	
Paper 25 (N215) ARIM-mdx Data System: Towards a Nationwide Data Platform for Materials Science (short)	
Paper 31 (N202) Data Workbench For Earth and Atmospheric Science Research Community (short)	

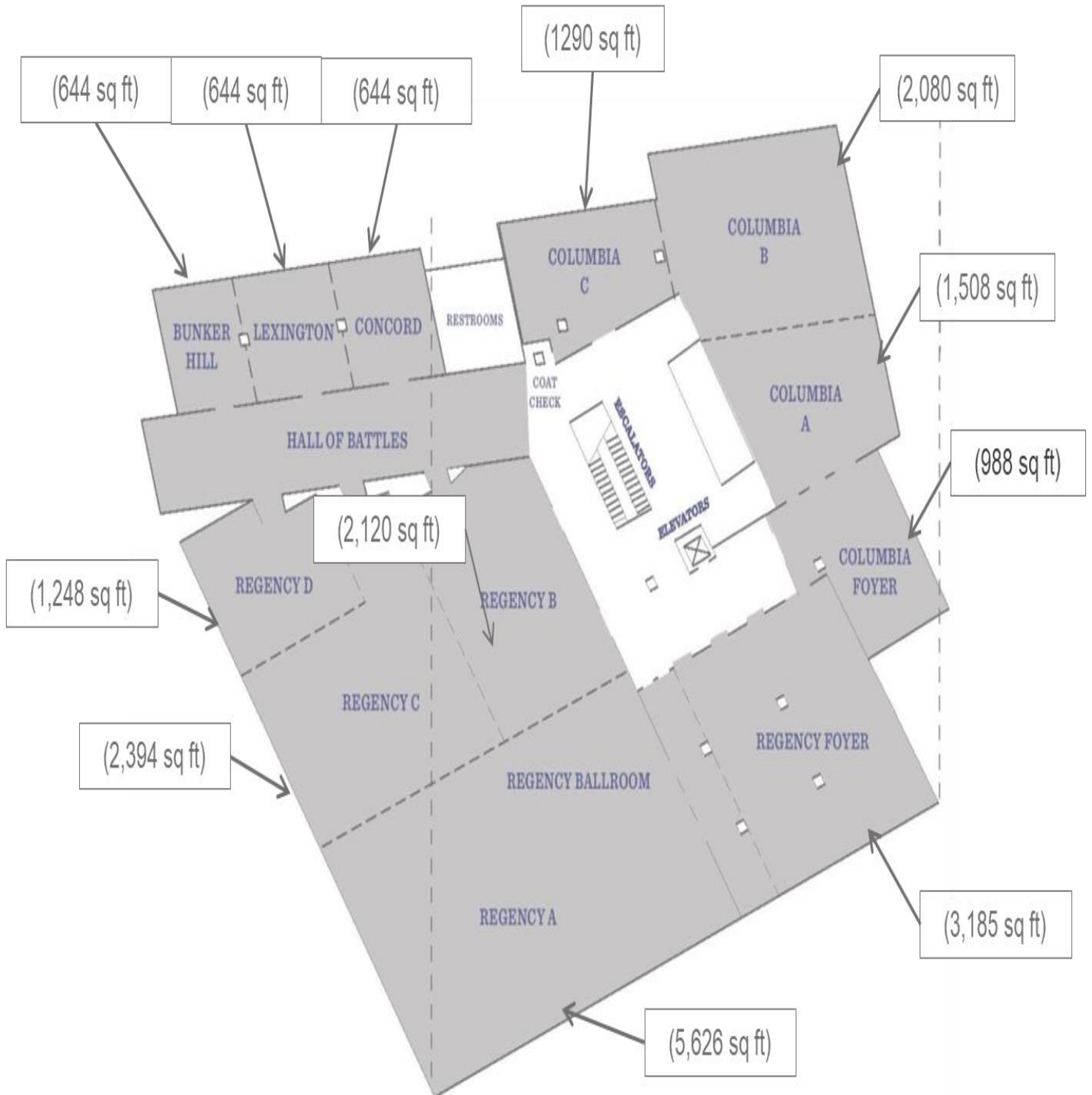
Online Workshop Sessions

Dec. 15, Sunday		
Workshop #11: The 7th International Workshop on “Big Data Engineering and Analytics in Cyber-Physical Systems” (BigEACPS’24)	Dec 15, Morning	Chair: Akbar Namin
Workshop #12: The 2nd International Workshop on Exploring Cutting-edge Data Mining: Applications of Infrared Target Detection and Large Language Models	Dec 15, Morning	Chair: Jinli Zhang
Combined Workshops: #19: Foundation Models for Science Big Data & #37: IEEE 5th Workshop on High Performance Computing, Big Data Analytics, and Integration for Multi-Omics Biomedical Data (HPC-BOD)	Dec 15, Morning	Chairs: Xuan Wang & Fahad Saeed
Workshop #25: IEEE Workshop on Big Data Analytics for Medical Imaging	Dec 15, Morning	Chair: Carmen Bisogni
Workshop #35: The IEEE International Workshop on Large Language Models for Finance	Dec 15, Morning	Chair: Quanzhi Li
Workshop #40: Workshop on Handling Resource constraints for/using Bigdata and AI	Dec 15, Morning	Chair: Manikandan Ravikiran
Workshop #45: 2nd International Workshop on Dataspaces and Digital Twins for Critical Entities and Smart Urban Communities (DSpaCES)	Dec 15, Morning	Chair: Marco Zappatore
Workshop #49: The 8th IEEE International Workshop on Big Spatial Data (BSD 2024)	Dec 15, Morning	Chair: Ashwin Shashidharan
Workshop #22: 1st Workshop on Robust Machine Learning for Distribution Shifts	Dec 15, Afternoon	Chair: Chen Zhao
Workshop #26: 8th International Workshop on Big Data Analytics for Cyber Intelligence and Defense (BDA4CID 2024)	Dec 15, Whole Day	Chair: Stephen McGough
Combined Workshops: #30 AI-Powered Renewable Energy Forecasting: Techniques and Challenges (AIPREF) & #32: AI-driven Agriculture: Opportunities and	Dec 15, Whole Day	Chairs: Luciano Caroprese, Ester Zumpano,

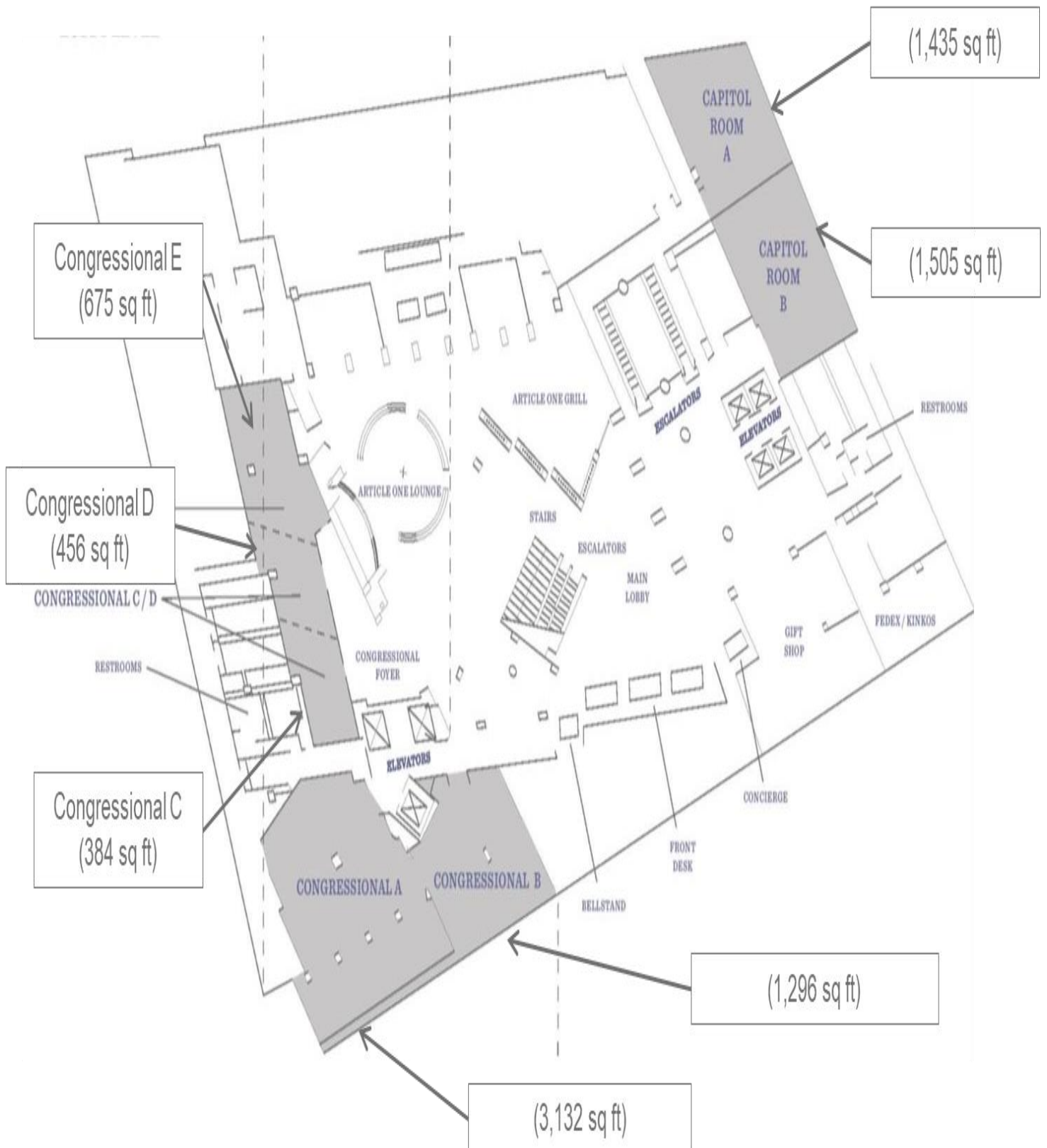
Challenges & #33: Cultural Heritage & BigData: Opportunities and Challenges		
Workshop: #31: International workshop on Scalable and Deep Graph Learning and Mining	Dec 15, Whole Day	Wissem Inoubli
Dec. 16, Monday		
Workshop #16: 12th International Workshop on Distributed Storage and Blockchain Technologies for Big Data	Dec 16, Morning	Chair: Hong Tan
Workshop #41: Fourth International Workshop on Data science for equality, inclusion and well-being challenges (DS4EIW 2024)	Dec 16, Morning	Chair: Tania Cerquitelli
Workshop #50: 2nd International Workshop on Creating Big Data for morals, ethics, and values in natural language in an automatically inferable format	Dec 16, Morning	Chair: Keisuke Nakamura
Workshop #27: 3rd International Workshop on Big Data Analytics for Health and Medicine (BDA4HM 2024)	Dec 16, Whole Day	Chair: Stephen McGough
Dec. 17, Tuesday		
Workshop #21: 5th International Workshop on User Understanding from Big Data	Dec 17, Morning	Chair: Wutao Wei
Workshop #29: Trustworthy Machine Learning for Healthcare	Dec 17, Morning	Chair: Xiao Shou
Workshop #28: 4th International Workshop on Big Data Analytics for Sustainability (BDA4S 2024)	Dec 17, Whole Day	Chair: Stephen McGough

Hotel Floor Layout

Ballroom Level



Lobby Level



2nd Floor Conference

