Monday, May 13			
Time	***Session***	***Speaker***	***Title***
8:00 -- 9:00	**Breakfast**		
9:00 -- 9:15	**Welcome and Opening remarks**	Keith Webster, Dean, CMU Libraries Michael McQuade, Vice President for Research, CMU Beth Plale, CISE/OAC, National Science Foundation	
9:15 -- 10:15	**Keynote 1**	Tom Mitchell, Carnegie Mellon University	Discovery from Brain Image Data
10:15 -- 10:35	**Break**		
10:35--12:20	**Session 1: Automation in data curation and metadata generation (Chair: Paola Buitrago)**		
10:35 -- 11:05	Long Talk	Cornelia Caragea, University of Illinois at Chicago	Keyphrase extraction from scholarly documents for data discovery and reuse
11:05 -- 11:35	Long Talk	Natasa Miskov-Zivanov, University of Pittsburgh	Dynamic System Explanation, DySE, a framework that evolves to reason about complex systems.
11:35 -- 11:50	Short Talk	Matias Carrasco Kind, National Center for Supercomputing Applications (NCSA)	Searching for similarities and anomalies in a pool of galaxy images using Deep Learning
11:50 -- 12:05	Short Talk	Rema Padman, Carnegie Mellon University	Ask the Doctor if YouTube is Right for You: An Augmented-Intelligence Video Recommender System for Patient Education
12:05 -- 12:20	Short Talk	Claudia Engel, Stanford University	Image Recognition for Archaeological Research
12:20 --13:20	**Lunch**		
13:20 -- 15:05	**Session 2: Automation in data discovery (Chair: Huajin Wang)**		
13:20 -- 13:50	Invited Talk	Natasha Noy, Google AI	Google Dataset Search: An open ecosystem for data discovery
13:50 -- 14:05	Short Talk	Fernando Chirigati, NYU	A Dataset Search Engine for Data Augmentation
14:05 -- 14:20	Short Talk	Alexander New, Rensselaer Polytechnic Institute	A Semantalytic Approach to Accelerated Data Reuse for Reproducible Scientific Discovery
14:20 -- 14:35	Short Talk	Shenghui Wang, OCLC Research, Netherlands	An innovative approach to scalable semantic search
14:35 -- 14: 50	Short Talk	Cornelia Caragea, University of Illinois at Chicago	Building Specialized Collections from Web Archiving
14: 50 -- 15:05	Short Talk	Jian Wu, Old Dominion University	Reuse and Discovery for Scholarly Big Data
15:05 -- 15:25	**Break**		
15:25 -- 16:10	**Panel 1**	Keith Webster, Carnegie Mellon University (Moderator) Cliff Lynch, Coalition for Networked Information Natasha Noy, Google AI Casey Greene, University of Pennsylvania Alex London, Carnegie Mellon University Paola Buitrago, Pittsburgh Supercomputing Center	**Challenges and opportunities in data reuse using the power of AI**
16:10 -- 17:30	**Poster + Networking**		
17:30 -- 19:30	**Reception**		

Tuesday, May 14			
	Session	***Speaker***	***Title***
8:00 -- 9:00	**Breakfast**		
9:00 -- 10:00	**Keynote 2**	Glen de Vries, Medidata Solutions	New Evidence Models: Clinical trials in the age of AI and Precision Medicine
10:00 --10:20	**Break**		
10:20 -- 12:05	**Session 3: Integrating datasets and enabling interoperability (Chair: Nick Nystrom)**		
10:20 -- 10:35	Short Talk	Evgeny Toropov, Carnegie Mellon University	Data reuse through domain adaptation AI algorithms for the self driving industry
10:35 -- 11:05	Long Talk	Jiacheng Zhu, Carnegie Mellon University	A self-organized Scenario-based Heterogeneous Traffic Database for Autonomous Vehicles
	~~Short Talk~~	~~Catherine Ordun, Booz Allen Hamilton~~	~~Model Tracking using the Keras API - Simple Metadata Management~~
11:05 -- 11:20	Short Talk	Daniel Clothiaux, Carnegie Mellon University	Visual and Statistical Analysis and Comparison of Handwritten and Font Datasets
11:20 -- 11:35	Short Talk	Xu Fei, Code Ocean	Lowering the barriers to experiment, data, and method reproducibility in AI research with a cloud-based computational reproducibility platform
11:35 -- 11:50	Short Talk	Rémi Mégret, University of Puerto Rico, Rio Piedras campus	LabelBee: a web platform for large-scale semi-automated analysis of honeybee behavior from video
11:50 -- 13:00	**Lunch**		
13:00 -- 17:00	**Session 4: Biomedical applications (Chairs: Sean Davis and Andreas Pfenning)**		
13:00 -- 13:30	Invited Talk	Casey Greene, University of Pennsylvania	Data reuse enables ML-based analysis of rare diseases
13:30 -- 13:45	Short Talk	Ben Busby, NCBI	Prototype ML Software for Several Distinct Classes of Biomedical Data Science Problems Developed in NIH-Hackathons!
13:45 -- 14:15	Invited Talk	Lisa Parker, University of Pittsburgh	Data Privacy: Control, Use, and Governance
15 minute break			
14:30 -- 15:00	Invited Talk	Sean Davis, National Cancer Institute, NIH	Data engineering: tools and approaches to facilitate data reuse and data science
15:00 -- 15:15	Invited Short Talk	Irene Kaplow, Carnegie Mellon University	Predicting Tissue-Specific cis-Regulatory Elements Across Mammals to Identify Potential Evolutionary Mechanisms
15:15 -- 15:45	Invited Talk	Fiona Nielsen , Repositive	Standards, incentives, tools – Which are the necessities for data discovery in academia vs industry?
15 minute break			
16:00 -- 16:30	Invited Talk	Alex London, Carnegie Mellon University	Understanding the Role of Explainaiblity and Verification in Medical AI
16:30 --16:45	Invited Talk	Nick Nystrom, PSC and Carnegie Mellon University	Enabling Data Discoverability in the Human BioMolecular Atlas Program (HuBMAP)
16:45 --17:15	Invited Talk	Bob Murphy, Carnegie Mellon University	AI for Biological Discovery: Data Integration and Self-Driving Instruments
17:15	**Dinner on your own**		

Wednesday, May 15			
	Session	***Speaker***	***Title***
8:00 -- 8:45	**Breakfast**		
8:45 -- 9:30	**Outcome and future planning meeting - all invited to participate (Moderator: Huajin Wang)**		
9:30 -- 10:15	**Panel 2**	Karen Lightman, Metro21: Smart Cities Institute (Moderator) Robet Tamburo, Carnegie Mellon University Bob Gradeck, Western Pennsylvania Regional Data Center Santi Garces, City of Pittsburgh	Enabling Smart and Safe Communities Through AI
10:15 -- 10:35	**Break**		
10:35 -- 11:55	**Session 5: Data security, privacy and algorithmic bias (Chair: Sayeed Choudhury)**		
10:35 -- 11:05	Invited Talk	Matt Fredrikson, Carnegie Mellon University	Finding Bias, Discrimination, and Private Data Leakage in Machine Learning Systems
11:05 -- 11:25	Long Talk	Lena Pons, Carnegie Mellon University	Sharable Cyber Threat Intelligence Using Weak Anonymization
11:25 -- 11:40	Short Talk	Andrew Yale, Rensselaer Polytechnic Institute	Privacy Preserving Synthetic Health Data
11:40 -- 11:55	Short Talk	Michael Ellis, University of North Carolina at Greensboro	Protecting fMRI data from unforeseen privacy attacks in a distributed machine learning environment
11:55 -- 12:10	**Closing remarks**	Keith Webster, Carnegie Mellon University	
12:10	**Ajourn. Boxed lunch to go**		

Post conference event:		
Software Engineering Institute (SEI) Research Services High Tea 2019		
Time: 14:00	Location: Jordan Auditorium	
Speaker:		**Topic**
Tom Longstaff, SEI Chief Technical Officer (Mater of Ceremonies)		
Keith Webster, Dean of Carnegie Mellon University Libraries		The Library's Role in the Fourth Industrial Revolution
Sarah Sheard, Principal engineer, Software Engineering Institute, Carnegie Mellon University		Systems at the SEI
Matt Burton, Lecturer, School of Computing and Information, University of Pittsburgh		AI and the Future of Library Science
Tom Corbett, Special Faculty, Entertainment Technology Center, Carnegie Mellon University		Forms Follow Functions: How AI and Spatial Computing Will Impact the Future of Video Games and Historic Preservation
English high tea reception		
Time: 15:00	Location: SEIber Café	