

Digital Total - Computing & Data Science an der Universität Hamburg und in der Wissenschaftsmetropole Hamburg

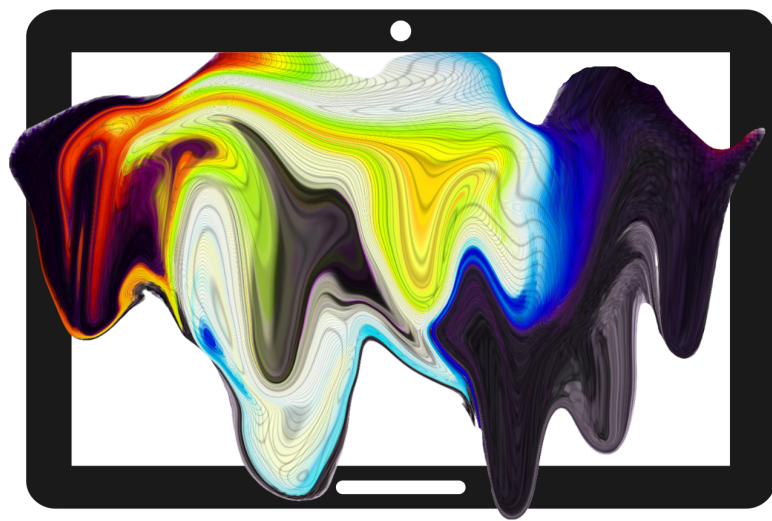
Edited by: Martin Semmann, Seid Muhie Yimam,
Katrín Schöning-Stierand and Chris Biemann

House of Computing and Data Science

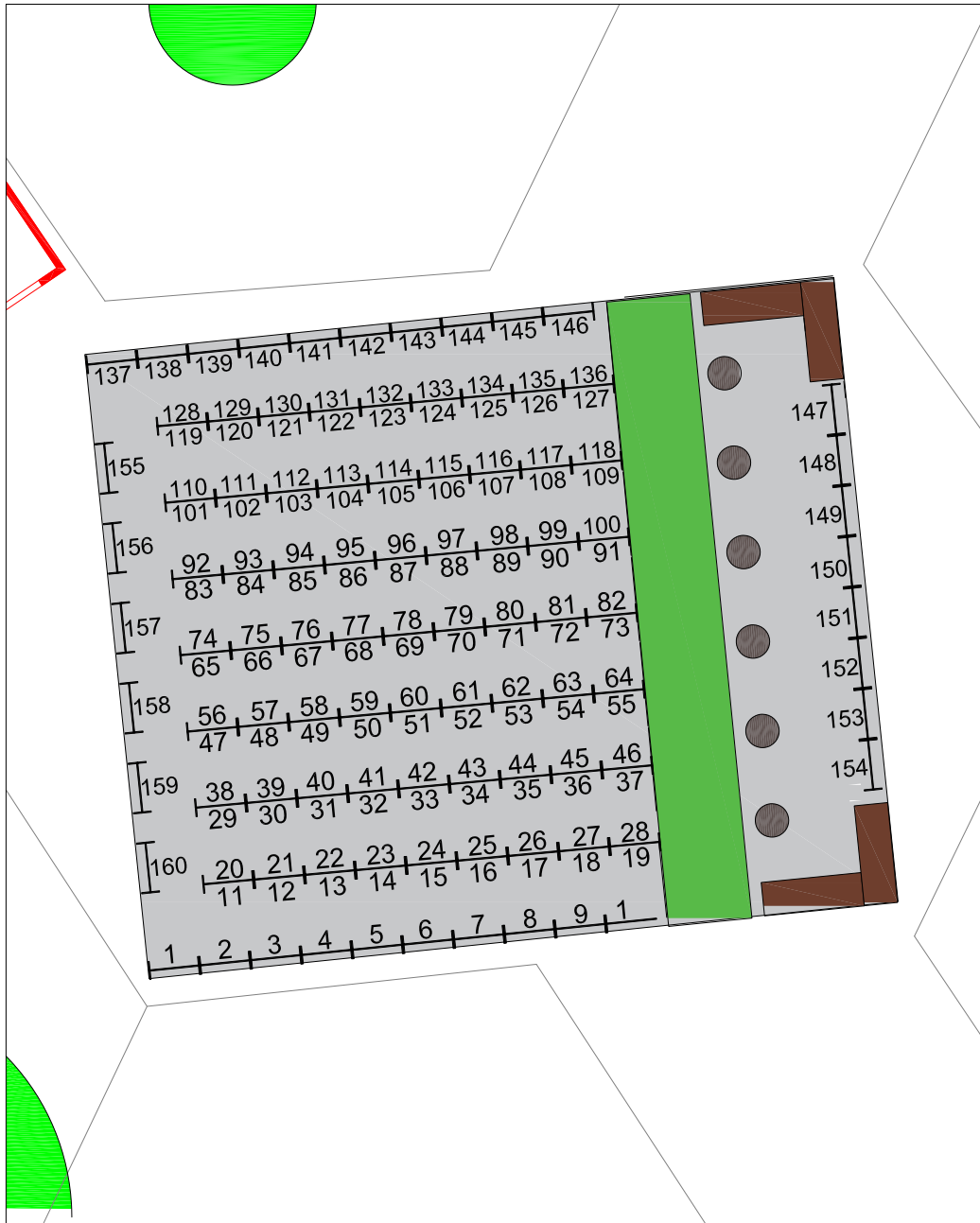
Monday, 9 October 2023 - Tuesday, 10
October 2023

Von-Melle-Park 4, Hamburg, Deutschland

Book of Abstracts



DIGITAL TOTAL



Poster Area | Tent Plan

Contents

Affiliations	1
Program	4
TentID-1: Using PEMT for pharmaceutical patent landscaping ⇒ Authors: Yojana Gadiya Andrea Zaliani, Matrin Hofmann-Apitius, Philip Gribbon	9
TentID-2: Elucidation of epigenetic mechanisms underlying normal and disease biology ⇒ Authors: Sheraz Gul Andrea Zaliani, Carsten Claussen, Ferdinand Seum, Jaleed Aslam, Olaf Schneider, Philip Groth, Sung-Min Pyo, Tara Beirne, Wilfried Blum	10
TentID-3: Reproducible Knowledge Graph workflows for embedding chemical entities and associated biology of diseases: A use case in MpoX ⇒ Authors: Reagon Karki, Andrea Zaliani, Yojana Gadiya, Philip Gribbon	11
TentID-4: Redirected Walks through the European XFEL ⇒ Authors: Eike Langbehn, Reinhard Bacher	12
TentID-5: Ein Intelligentes Fenster zum Hamburger Hafen: Entwicklung einer Augmented Reality Fensterscheibe für Schiffe und Infotainment-Installationen ⇒ Authors: Jan Jessen Eike Langbehn	13
TentID-6: Facilitating FAIR Data Practice: The Research Data Repository for the Universität Hamburg ⇒ Authors: Kai Joachim Wörner	14
TentID-7: CRIS at UHH: digital, structured, sustainable ⇒ Authors: Lisa Brase, Steffen Grünler	15

TentID-8: Pattern Analysis Software Tools ⇒ Authors: Hussein Mohammed	16
TentID-9: Computational Analysis of Handwriting Style ⇒ Authors: Hussein Mohammed	17
TentID-10: Digital Data Handling at UWA ⇒ Authors: Sylvia Melzer Stefan Thiemann, Ralf Möller	18
TentID-12: Few-shot learning for automated content analysis (FLACA) in the German media debate on arms deliveries to Ukrain ⇒ Authors: Jonas Rieger, Mattes Ruckdeschel, Kostiantyn Yanchenko, Gerret von Nordheim, Katharina Kleinen von Königslöw, Gregor Wiedemann	19
TentID-13: Social Media Observatory ⇒ Authors: Felix Victor Münch, Gregor Wiedemann, Jan Rau, Jan-Hinrik Schmidt, Philipp Kessling	20
TentID-14: Augmenting Social Media Actor Lists with Explorative Network Sampling: Mapping German Right-wing Counterpublics on Twitter ⇒ Authors: Felix Victor Münch, Philipp Kessling, Jan Rau	21
TentID-15: Supporting Sustainable Research: The Services of the Center for Sustainable Research Data Management ⇒ Authors: Alexander Heinrich, Hagen Peukert, Iris Vogel, Juliane Jacob, Kai Joachim Wörner, Lisa Brase, Stefan Thiemann, Steffen Grünler	22
TentID-16: Visual-Pattern Detection in Historical Documents ⇒ Authors: Hussein Mohammed	23
TentID-17: Palimpsests inpainting with Generative AI ⇒ Authors: Hussein Mohammed Jost Gippert, Mahdi Jampour	24
TentID-18: Profiling of written artefacts with bioinformatic and chemometric methods ⇒ Authors: Lucas F. Voges, Stephan Seifert	25
TentID-19: Introduction of the research group "Chemometrics of complex material systems" ⇒ Authors: Stephan Seifert	26
TentID-20: Smart and sustainable cities? Testing a Socio-Technical	

Transitions Model of the Interplay of Urban Digitalization and Sustainable Development ⇒ Authors: Laura Marie Edinger-Schons	27
TentID-21: The Role of Digitalization for Firm Financial and Ecological Performance: A Socio-technical Systems Perspective ⇒ Authors: Alicia Minnerup Caroline Ruiner, Frederik Maibaum, Johann Nils Foege, Johannes Kriebel, Laura Marie Edinger-Schons, Sascha Alavi	28
TentID-22: Assessment of antibiotic resistance from microbial genomic data ⇒ Authors: Lennart Bartels, Inken Wohlers	29
TentID-23: Biomolecular Data Science in Pneumology ⇒ Authors: Lennart Bartels, Inken Wohlers	30
TentID-24: Genome-based prediction of autoimmune diseases using relation neural networks ⇒ Authors: Lennart Bartels Marius Jahrens, Thomas Martinetz, Hauke Busch, Inken Wohlers	31
TentID-25: ProteinsPlus: On-The-Fly Structure-Based Design on the Web ⇒ Authors: Christiane Ehrt Thorben Schulze, Katrin Schöning-Stierand, Rainer Fährrolfes, Konrad Diedrich, Joel Graef, Emanuel Ehmki, Florian Flachsenberg, Jochen Sieg, Patrick Penner, Jonathan Pletzer-Zelgert, Thorben Reim, Matthias Rarey	32
TentID-26: Exploring Ultra-Large Chemical Spaces With Genetic Algorithms ⇒ Authors: Christian Meyenburg Uschi Dolfus, Matthias Rarey	33
TentID-27: DASHH: Data Science in Hamburg –Helmholtz Graduate School for the Structure of Matter -> Educating the Next Generation of Data Scientists ⇒ Authors: Christiane Ehrt, Heike Hufnagel Martinez	34
TentID-28: Calculating More Property Distributions of Chemical Fragment Spaces ⇒ Authors: Justin Lübbers Louis Bellmann, Matthias Rarey, Uta Lessel	35
TentID-29: BACI/4C: Next-Generation Multiphysics Simulation for Challenging Real-World Problems ⇒ Authors: Dirk Steglich Christian Cyron, Georg Hammerl, Hoang-Giang Bui, Ingo Scheider	36

TentID-30: High-throughput processing for diffraction data ⇒ Authors: Malte Storm Christina Krywka, Peter Staron	37
TentID-31: KI4D4E: Artificial Intelligence for Synchrotron-radiation 4D Tomography Data ⇒ Authors: Tak Ming Wong	38
TentID-32: Exploratory Clinical-Lab Data Analysis ⇒ Authors: Olaf Deppe	39
TentID-33: Entwicklung eines computergestützten Tools zur Bereitstellung von Informationen, Struktur und Entscheidungshilfen für das hausärztliche Management von Multimorbidität (gp-multitool.de) ⇒ Authors: Anshu Saini, Ingmar Schäfer, Julia Nothacker, Martin Scherer	40
TentID-34: The Prompt-a-thon: Designing a Format for Value Co-Creation with Generative AI for Research and Practice ⇒ Authors: Constantin von Brackel-Schmidt, Emir Kucevic, Tilo Böhmann	41
TentID-35: The Digital Causality Lab - New Ways to Teach Causality ⇒ Authors: Philipp Bach Gangli Tan, Martin Spindler	42
TentID-36: Predicting COVID-19 Vaccination Uptake from Public Discourse: A Machine Learning Approach ⇒ Authors: Esra Eren Bayindir, Jonas Schreyögg, Martin Spindler, Robert Fuchs	43
TentID-37: Causal AI mit DoubleM ⇒ Authors: Philipp Bach, Sven Klaassen, Martin Spindler	44
TentID-38: In the Pursuit of Eco-Friendly Wind Energy –Unmanned Aerial Systems and Computational Strategies ⇒ Authors: Tessa T. Taefi Anna Roswag, Gerald Peklar, Marc Roswag, Matthias Roswag, Sven Stölzel, Veit Dominik Kunz	45
TentID-39: Massively Parallel Hybrid Molecular-Continuum Simulations ⇒ Authors: Louis Viot	46
TentID-40: marFM®, A Multilingual Automatic Speech Recognizer (ASR) for VHF-Radio Communication ⇒ Authors: Emin Nakilcioglu Maximilian Reimann, Ole John	47
TentID-41: Enabling Secure Communication for Cyber-Physical	

Processes in Critical Infrastructures ⇒ Authors: Sibylle Fröschle Philipp Wolters	48
TentID-42: Presenting a multi-layered manuscript of a Byzantine Greek lexicon within a responsive digital edition ⇒ Authors: Alessandro Musino, Christian Brockmann, , Daniel Deckers, Eva Wöckener-Gade, Lena Hofmann, Louiza Argyriou, , Stefano Valente	49
TentID-43: Making Academies' long-term research projects ready for the NFDI ⇒ Authors: Timm Lehmberg	50
TentID-44: Decentralized Data Mining ⇒ Authors: Tim Schatto-Eckrodt	51
TentID-45: Herbie: ELN and research database from Hereon ⇒ Authors: Catriona Eschke Fabian Kirchner, Martin Held, Milosz Meller	52
TentID-46: Prediction of therapeutic targets to halt inflammation- induced neurodegeneration. ⇒ Authors: Manuel Friese, Marcel Seungsu Woo	53
TentID-47: CSL Poster Overview ⇒ Authors: Gesa Ziemer ines bigdeli issazadeh	54
TentID-48: Data analysis at the European XFEL ⇒ Authors: Ivette Bermudez Macias	55
TentID-49: BORGES: A database for the storage, search and retrieval of ICON data ⇒ Authors: Jairo Alonso Segura Bermudez, Luis Ko- rnblueh, Raphael Schlarb, Reinhard Budich, Sven Willner	56
TentID-50: The VoicesCarry App: Co-Designing Worker Engagement Technology with Supply Chain Workers ⇒ Authors: Emma van den Terrell, Laura Edinger-Schons	57
TentID-51: ML based process monitoring of semi-automatic drilling processes ⇒ Authors: Malte Flehmke	58
TentID-52: PUNCH4NFDI - Facilitating FAIR data in Germany ⇒ Au- thors: Christiane Schneide, Thomas Schörner	59

TentID-53: MRI Total: Towards an automated, standardized, reproducible, version-controlled, transparent, FAIR, quality-assured, privacy-compliant processing pipeline for MRI data ⇒ Authors: Lennart Wittkuhn	60
TentID-54: Helmholtz Imaging@DESY ⇒ Authors: David Schwartz, Martin Burger Tim Roith	61
TentID-55: Computational Imaging@DESY ⇒ Authors: Lukas Weigand Lorenz Kuger, Martin Burger, Samira Kabri, Tim Roith	62
TentID-56: Using computational methods to study niche audiences ⇒ Authors: Judith Möller Ellen Linnert	63
TentID-57: INEL –Grammar, Corpora and Language Technology for Indigenous Northern Eurasian Languages ⇒ Authors: Elena Lazarenko, Aleksandr Riaposov	64
TentID-58: Climate Data Operators - CDO ⇒ Authors: Uwe Schulzweida	65
TentID-59: Motion Editing tool for Reproducing Grammatical Elements of Japanese Sign Language Avatar Animation ⇒ Authors: Tsubasa Uchida Hiroyuki Kaneko, Masanori Sano, Naoki Nakatani, Taro Miyazaki	66
TentID-60: Interdisciplinary and International Collaborations: The MARS Research Group at Hamburg University of Applied Sciences (HAW) ⇒ Authors: Thomas Clemen Daniel Glake, Helen Haase, Ulfia Clemen	67
TentID-61: Conceptual foundation of machine learning in Engineering using example of tailored laser pulse forming and Optical Mirror Design ⇒ Authors: Utsa Chattopadhyay	68
TentID-62: Towards the Automatic Generation of Models for Prediction, Monitoring, and Testing of Cyber-Physical Systems ⇒ Authors: Eric MSP Veith, Görschwin Fey, Jakob Schyga, Jan Christian Wieck, Jochen Kreutzfeldt, Johannes Hinckeldeyn, Julian Kohlisch, Markus Knitt, Maximilian Schmidt, Stephan Balduin, Swantje Plambeck	69
TentID-63: Digitisation Services @ SUB HH ⇒ Authors: Kerstin	

Wendt, Michael Kubina	70
TentID-64: ahoi.digital - Die Allianz der Hamburger Hochschulen für Informatik ⇒ Authors: Janis-Marie Paul Jörg Rainer Noennig, Martin Semmann, Nihat Ay, Tilo Böhmann, Ulrike Steffens	71
TentID-65: Sind wir interessant für KI –oder: ist KI interessant für uns? Analysen des Textkorpus „Mehrsprachigkeitsentwicklung im Zeitverlauf (MEZ)“ ⇒ Authors: Ingrid Gogolin, Nora Dünkel, Thorsten Klinger, Kseniia Pershina, Anouk Ticheloven, Birger Schnoor, Irina Usinova	72
TentID-66: Spoki: A Reactive and Scalable Network Telescope ⇒ Authors: Raphael Hiesgen	73
TentID-67: bAlome Center for Biomedical AI- Towards Intelligent Health ⇒ Authors: Anna Reinicke-Vogt, Claus Christian Hilgetag, Frank Ückert, Jan Erik Gewehr, Lorenz Adlung, Marina Zimmermann, Rene Werner, Stefan Bonn, Stefano Panzeri, Vadim Ustinov	74
TentID-68: ExaOcean: Improving Performance of the ICON-O Ocean-model on heterogeneous Exascale-Supercomputers with Machine Learning ⇒ Authors: Daniel Ruprecht Christopher Kadow, Lars Hoffmann, Peter Korn	75
TentID-69: Combining numerical methods and machine learning for physical and engineering sciences ⇒ Authors: Sebastian Götschel Daniel Ruprecht	76
TentID-70: From Super-Resolution to Downscaling - An Image-Inpainting Deep Neural Network for High Resolution Weather and Climate Models ⇒ Authors: Christopher Kadow, Maximilian Witte Danai Filippou, David Hall, Johannes Meuer, Étienne Plésiat	77
TentID-71: DDLitLab ⇒ Authors: Carolin Scharfenberg, Julia Niemann-Lenz, Kai-Uwe Schnapp, Katharina Kleinen-von Königslöw	78
TentID-72: Does Journalism Reflect Polarized Online Networks? A Comparative Analysis of the Discourse on the "Last Generation"	

Climate Protest Movement in (Social) Media ⇒ Authors: Hendrik Meyer, Louisa Pröschel, Michael Brüggemann	79
TentID-73: When Recommender Systems Select Political News ⇒ Authors: Hendrik Meyer, Jessica Kunert, Juliane Lischka, Katharina Kleinen-von Königslöw, Laura Laugwitz, Lucien Heitz, Nadja Schaetz, Rana Abdullah	80
TentID-74: Federated Machine Learning for Big Data ⇒ Authors: Stefan Schulte	81
TentID-75: Beta maṣāḥəft: Die webbasierte Kollaborationsplattform für Handschriftenforschung ⇒ Authors: Alessandro Bausi Carsten Hoffmann, Denis Nosnitsin, Dorothea Reule, Eugenia Sokolinski	82
TentID-76: Leveraging Jupyter on Maxwell HPC: joyful, visual and green computing ⇒ Authors: Frank Schlünzen Arlena Mills-Marzoli, Axel Wichmann, Neele Rahmlow, Sven Sternberger, Yves Kemp	83
TentID-77: Digital Scholarship Services at State- and University Library Hamburg ⇒ Authors: Jonas Müller-Laackman	84
TentID-78: QLK –Queer-linguistisches Korpus ⇒ Authors: Lars Vorberger Miriam Hinterholzer	85
TentID-79: Institute for Data Science Foundations ⇒ Authors: Nihat Ay	86
TentID-80: Computational aspects of the active self ⇒ Authors: Frank Röder, Carlotta Langer Jan Dohmen, Manfred Eppe, Nihat Ay	87
TentID-81: Machine Learning based optimization of transverse beam emittance in Free Electron Lasers ⇒ Authors: Alexander Klemps, Nihat Ay	88
TentID-82: Natural Language Processing in the Legal Domain ⇒ Authors: Daniel Katz, Dirk Hartung, Lauritz Gerlach, Abhik Jana, Michael James Bommarito	89
TentID-83: Flimma: a federated and privacy-aware tool for differential gene expression analysis ⇒ Authors: Amir Abbasinejad, Daniel	

Rückert, David B. Blumenthal, Georgios Kaissis, Jan Baumbach, Julian Klemm, Julian Matschinske, Julian Späth, Markus List, Mohammad Bakhtiari, Nina Wenke, Olga Zolotareva, Paolo Tieri, Reihaneh Torkzadehmahani, Reza Nasirigerdeh, Tobias Frisch	90
TentID-84: Privacy-preserving federated prognostic model learning from multi-centric questionnaire data to predict pain changes in osteoarthritis ⇒ Authors: Supratim Das Mahdie Rafiei, Baumbach Jan, Baumbach Linda	91
TentID-85: Predicting multiple conformations of flexible proteins ⇒ Authors: Karen Manalastas-Cantos Kish Adoni, Matthias Pfeifer, Birgit Märten, Kay Grünewald, Konstantinos Thalassinou, Maya Topf	92
TentID-86: Shaking out Intrinsic Structural Variation in Femtosecond Protein Crystallography using Deep Learning ⇒ Authors: Florian Griese, José Licón-Saláiz, Samuel Flewett, Thomas White	93
TentID-87: Machine Learning for the Automated Selection and Reconstruction of Multi-Modal Nanotomography Data of Bone-Implant Interfaces ⇒ Authors: Benedict Schacht Berit Zeller-Plumhoff, Christian Wilms, Imke Greving, Simone Frintrop	94
TentID-88: Integrated Climate Data Center ⇒ Authors: Stefan Kern, Remon Sadikni	95
TentID-89: Open Access Publishing for Hamburg Research. Quality-Assured Scholarly Open Access Publication Services ⇒ Authors: Isabella Meinecke, Oliver Krüger, Tobias Buck	96
TentID-90: Scholarly Publishing in the European Research Area (ERA) - Supporting the Visibility of Hamburg Research Results ⇒ Authors: Isabella Anna Greiner, Isabella Meinecke, Oliver Krüger	97
TentID-91: Legal Hypergraphs ⇒ Authors: Corinna Coupette, Dirk Hartung, Daniel Katz	98
TentID-92: A Personalized Web Application to Predict Changes in Knee Pain among Patients with Osteoarthritis Participating at the GLA:D® program ⇒ Authors: Mahdie Rafiei Supratim Das, Jan	

Baumbach, Linda Baumbach	99
TentID-93: Privacy-Preserving Federated Differential Protein Expression Analysis with FedProt ⇒ Authors: Yuliya Burankova Olga Zolotareva, Miriam Abele, Fabian Gruhn, Christine von Törne, Teresa Barth, Lisa Schweizer, Pieter Giesbertz, Mohammad Bakhtiari, Anne Hartebrodt, Tobias Frisch, Chen Meng, Julian Matschinske, Julian Späth, Richard Röttger, Veit Schwämmle, Stefanie M. Hauck, Stefan Lichtenthaler, Axel Imhof, Matthias Mann, Christina Ludwig, Bernhard Küster, Jan Baumbach	100
TentID-94: Detection of alternative splicing: deep sequencing or deep learning ⇒ Authors: Lena Maria Hackl Jan Baumbach, Olga Tsoy	101
TentID-95: Identification of differentially expressed biclusters for unsupervised patient stratification ⇒ Authors: Michael Hartung, Andreas Maier, Fernando Miguel Delgado Chaves, Yuliya Burankova, Olga Isaeva, Daniel He, Katharina Kaufmann, Fábio Malta de Sá Patroni, Alexey Savchik, Zoe Chervontseva, Niklas Probul, Alexandra Abisheva, Evgenia Zotova, Olga Tsoy, David Blumenthal, Martin Ester, Olga Zolotareva, Jan Baumbach	102
TentID-96: Decentralized Batch Effect Correction in Multi-Center Biomedical Studies: Introducing FedComBat, a Federated Learning Approach ⇒ Authors: Mohammad Mahdi Kazemi Majdabadi Jan Baumbach, Olga Zolotareva, Yuliya Burankova .	104
TentID-97: The CDCS - A Collaborative Research Center for Data Science in Hamburg ⇒ Authors: Annett Ungethüm, Antonin Sulc, José Licón-Saláiz, Karen Manalastas-Cantos, Khalique Newaz, Lennart Rustige, Lennart Rustige, Patrick L.S. Connor, Samuel Flewett	105
TentID-98: ERIS: Hamburg Information System on the Representation of Greek and Roman Violence ⇒ Authors: Justine Diemke, Werner Rieß	106
TentID-99: FAIR-Based Strategy for Data Management and Handling in Large Climate Modeling Projects: Experiences from PalMod-II ⇒ Authors: Swati Gehlot Andrea Lammert, Hannes Thiemann, Karsten Peters-von Gehlen	107

TentID-100: World Data Center for Climate (WDCC) - Improvement of FAIRness of an established repository ⇒ Authors: Eileen Hertwig Andrea Lammert, Heinke Höck, Andrej Fast, Hannes Thiemann	108
TentID-101: DRaCOoN: Differential Regulation and CO-expression Networks ⇒ Authors: Fernando Miguel Delgado Chaves Ferdinand Spurny, Jan Baumbach, Olga Zolotareva	109
TentID-102: Automated symptom evaluation using sensor data from wearables in patients with Parkinson’s disease utilizing explainable AI ⇒ Authors: Alexander Wiederhold, Qi Rui Zhu, Sören Spiegel, Frank Ückert, Christopher Lukas Gundler	110
TentID-103: On the power and limits of structure-based exon-exon interaction prediction ⇒ Authors: Jeanine Liebold Aylin Del Moral Morales, Karen Manalastas-Cantos, Olga Tsoy, Stefan Kurtz, Jan Baumbach, Khalique Newaz	111
TentID-104: sPLINK: a hybrid federated tool as a robust alternative to meta- analysis in genome-wide association studies ⇒ Authors: Julian Matschinske, Mohammad Bakhtiari	112
TentID-105: Research data management at the German Institute for Global and Area Studies (GIGA) ⇒ Authors: Catharina Hibler	112
TentID-106: Hyperbolic Learning on Omics Data ⇒ Authors: Ilya Petrov, Kate Trifonova, Semyon Karaban, Ariana Migutskaya, Jackie Khodzhaeva, Mikhail Gelfand, Zoe Chervontseva	114
TentID-107: Exploring supervised machine learning algorithms for predicting pitting corrosion damage of AA1050 exposed to ethanol-containing fuels ⇒ Authors: Lukas Jarren Christian Feiler, Daniel Höche, Eugen Gazenbiller, Matthias Oechsner, Mikhail Zheludkevich, Rüdiger Reitz, Visheet Arya	115
TentID-108: Selection of Effective Corrosion Inhibitors for Aluminium Alloys based on Data-Driven Techniques ⇒ Authors: Lisa Sahlmann Can Özkan, Christian Feiler, Daniel Höche, J.M.C. Mol, Mikhail Zheludkevich, Natalia Konchakova, Peyman Taheri, Sviatlana Lamaka	116
TentID-109: A Computational Approach to Magnesium Corrosion En-	

gineering ⇒ Authors: Tim Würger Bahram Vaghefinazari, Christian Feiler, David Winkler, Di Mei, Mikhail Zheludkevich, Robert Meißner, Sviatlana Lamaka	117
TentID-110: Leveraging Semantic Information for Efficient Self-Supervised Emotion Recognition with Audio-Textual Distilled Models ⇒ Authors: Danilo de Oliveira, Navin Raj Prabhu, Timo Gerkmann	118
TentID-111: Feature Reduction Improves Drug Response Prediction ⇒ Authors: Farzaneh Firoozbakht Behnam YOUSEFI, Olga Tsoy, Jan Baumbach, Benno Schwikowski	119
TentID-112: Studying protein 3D structures using network science ⇒ Authors: Khalique Newaz Tijana Milenkovic	120
TentID-113: NeDRex - an integrative and interactive network medicine platform for drug repurposing ⇒ Authors: James Skelton, Sepideh Sadegh Andreas Maier, Jan Baumbach	121
TentID-114: The AIME registry for artificial intelligence in biomedical research ⇒ Authors: Julian Matschinske David Blumenthal, Jan Baumbach, Nina Wenke, Olga Zolotareva	122
TentID-115: Inferring Tissue-Specific Alternative Splicing from Blood Samples ⇒ Authors: Jan Baumbach, Olga Tsoy, Silke Van Elferen	123
TentID-116: Visualization and Sign Representation in the Digital Dictionary DGS (DW-DGS) ⇒ Authors: Anke Müller, Felicitas Otte, Gabriele Langer, Sabrina Wähl, Thomas Hanke	124
TentID-117: MY DGS –ANNIS: ANNIS and the Public DGS Corpus ⇒ Authors: Amy Isard, Reiner Konrad	125
TentID-118: The DW-DGS: A corpus-based dictionary of German Sign Language is being compiled ⇒ Authors: Anke Müller, Gabriele Langer, Sabrina Wähl, Thomas Hanke	126
TentID-119: Diffusion Models for Audio-Visual Speech Enhancement ⇒ Authors: Julius Richter Timo Gerkmann	127
TentID-120: Multimodal Modelling of Cultural Artefacts in Digital Space ⇒ Authors: Frank Steinicke, Julia Nantke, Qianqi Huang .	128

TentID-121: In-the-wild Speech Emotion Conversion Using Disentangled Self-Supervised Representations and Neural Vocoder-based Resynthesis ⇒ Authors: Navin Raj Prabhu Nale Lehmann-Willenbrock, Timo Gerkmann	129
TentID-122: Comparative Annotation to Explore and Explain Text Similarities (CompAnno) ⇒ Authors: Janis Pagel, Julia Nantke, Marie Flüh, Nils Reiter	130
TentID-123: 123 - DisKo –Das Diversitäts Korpus ⇒ Authors: Mareike Schumacher, Marie Flüh	131
TentID-124: Extending the Public DGS Corpus in Size and Depth ⇒ Authors: Elena Jahn, Marc Schulder, Reiner Konrad, Thomas Hanke	132
TentID-125: Semi-Automatic Subtitle Alignment for Sign Languages ⇒ Authors: Maren Brumm Sam Bigeard	133
TentID-126: The Multilingual Sign Languages Wordnet ⇒ Authors: Sam Bigeard Marc Schulder, Maria Kopf, Thomas Hanke	134
TentID-127: The Sign Language Dataset Compendium: Creating an Overview of Digital Sign Language Resources ⇒ Authors: Maria Kopf, Marc Schulder, Thomas Hanke	135
TentID-128: KI für die Internationale Raumstation ⇒ Authors: Alexander Diedrich, Henrik Steude, Lukas Moddemann, Mark Tappe, Oliver Niggemann, Stephan Myschik	136
TentID-129: Akidu - Automatische, durch KI-integrierte Disposition für Universalterminals ⇒ Authors: Lothar Hotz, Matthias Kerzel, Pascal Rost, Rainer Herzog	137
TentID-130: DaFne - Platform Data Fusion Generator ⇒ Authors: Ulrike Steffens, Jörg Rainer Noennig, Pamela Kunert, Ayse Glass	138
TentID-131: AI for Happiness: Pedestrian Path Generation through Agent-Based Simulations with Deep Reinforcement Learning ⇒ Authors: Ayse Glass, Burak Bek, Eylül Kibele Mengeş, Iryna Okhrin, Jörg Rainer Noennig, Mariela Rossana Sanchez, Pramod Baddam, rene jaekel, roman glass	139

TentID-132: Generative Machine Learning for Particle Physics ⇒ Authors: Gregor Kasieczka, Thorsten Buss	140
TentID-133: Anomaly detection for discovery in particle physics ⇒ Authors: Manuel Sommerhalder Gregor Kasieczka, Louis Moureaux, Tobias Quadfasel	141
TentID-134: Fast machine learning in particle physics ⇒ Authors: Gregor Kasieczka, Karim El Morabit, Sven Bollweg	142
TentID-135: Inflation Narratives from a Machine Learning Perspective ⇒ Authors: Cedric Möller, Junbo Huang, Max Weinig, Ricardo Usbeck, Ulrich Fritsche	143
TentID-136: Inflation Narratives from an Economics Perspective ⇒ Authors: Cedric Möller, Junbo Huang, Max Weinig, Ricardo Usbeck, Ulrich Fritsche	144
TentID-137: SmartShip: KI für Echtzeit-Anomalieerkennung und automatischer Objekterkennung in maritimen Such- und Rettungsaktionen ⇒ Authors: Alexander Windmann, Lukas Augustin Jonas Schöttler, Juan Zhang, Oliver Niggemann, Pavle Ivanovic, Philipp Neumann	145
TentID-138: Smart Assembly im Labor für intelligente Leichtbauproduktion ⇒ Authors: Phillip Johann Overlöper, Sven Karras Constantin Hildebrandt, Henrik Eschen, Jonas Ehrhardt, Oliver Niggemann	146
TentID-139: hpc.bw –Competence Platform for Software Efficiency and Supercomputing (dtec.bw) ⇒ Authors: Alexander Kolling, Hauke Preuß, Marie Rathmann, Philipp Neumann, Piet Jarmatz .	147
TentID-140: EvalSpek-ML ⇒ Authors: Daniel Boschmann Oliver Niggemann, Willi Großmann	148
TentID-141: Data-based tool try out in sheet metal forming ⇒ Authors: Michael Hohmann Adili Yiming, Lars Penter, Oliver Niggemann, Steffen Ihlenfeldt	149
TentID-142: Morphology-based Molecular Classification of Spinal Cord Ependymomas Using Deep Neural Networks ⇒ Authors: Yannis Schumann Julia Neumann, Leonille Schweizer, Maja Krech,	

Matthias Dottermusch, Philipp Neumann, Tasja Lempertz, Ulrich Schüller	150
TentID-143: Understanding and Improving Large Language Models ⇒ Authors: Michaela Regneri Alhassan Abdelhalim, Merle Mosch, Sören Laue	152
TentID-144: Machine Learning Group at Universität Hamburg ⇒ Au- thors: Sören Laue Michaela Regneri, Tomislav Prusina	153
TentID-145: Multilingual Racial Hate Speech Detection Using Trans- fer Learning ⇒ Authors: Abinew Ali Ayele, Chris Biemann, Seid Muhie Yimam, Skadi Dinter	154
TentID-146: German Large Language Legal Model (GLLLM) ⇒ Au- thors: Chris Biemann, Fereidoun Rashidi, Martin Semmann, Seid Muhie Yimam	155
TentID-147: ARDIAS: AI-Enhanced Research Discovery and Advisory System ⇒ Authors: Aditya Jogalekar, Chris Biemann, Debayan Banerjee, Martin Semmann, Seid Muhie Yimam	156
TentID-148: House of Computing and Data Science ⇒ Authors: Chris Biemann, Katrin Schöning-Stierand, Martin Semmann, Seid Muhie Yimam	157
TentID-149: Transformation of ERIS Information System to Advanced Search ⇒ Authors: Haruka Inoba, Seid Muhie Yimam, Martin Sem- mann, Chris Biemann Werner Rieß, Justine Diemke, Ahmet Yilmaz	158
TentID-150: D-WISE - Digitale Wissenssoziologische Diskursanalyse ⇒ Authors: Liv Ohlsen, Tim Fischer Chris Biemann, Florian Schneider, Fynn Petersen-Frey, Gertraud Koch, Isabel Eiser	159
TentID-151: FUNDus meets HCDS - Novel avenues to explore scien- tific collections ⇒ Authors: Florian Schneider, Iris Vogel, Martin Semmann	160
TentID-152: RESCUE-MATE - Rethinking Innovation in Natural Disas- ters ⇒ Authors: Heiko Bornholdt, Julia Bräker, Martin Semmann, Mathias Fischer	161

TentID-153: Mikropolis - Orientierung in der digitalen Transformation ⇒ Authors: Arno Rolf, Janick Edinger, Larissa Gebken, Lothar Hotz, Martin Semmann	162
TentID-154: InteGreatDrones: New Perspectives on Ports Through Drones and AI ⇒ Authors: André Kelm, Helen Schloh, Janick Edinger, Liviana Franke, Simone Frintrop	163
TentID-155: Group dynamics in the Metaverse: A conceptual model and first empirical insights ⇒ Authors: Marvin Grabowski Frank Steinicke, Anita Blanchard, Sebastian Rings, Nale Lehmann-Willenbrock	164
TentID-156: Advanced mathematical solutions for circulating cardiovascular biomolecules ⇒ Authors: Benedikt Schrage, Christoph Al-Taie, Julius Obergassel, Olga Tsoy, Olga Zolotareva, Silke Van Elferen	165
TentID-157: FLIGHT: Federated LearnIng Guided digital Health ⇒ Authors: Jens Lohmann, Olga Zolotareva, Guido Sauter, Stefan Bonn, Jan Baumbach	166
TentID-158: Sustainable data management @ Tamilex ⇒ Authors: Charles Li, Eva Wilden	167
TentID-159: Research in Geoinformatics and Geovisualization at HCU ⇒ Authors: Jochen Schiewe	168
Index	171

Affiliations

AHOI.DIGITAL || ZMNH || WINDSTROM ERNEUERBARE ENERGIEN
|| VESPERTILLO FAUNISTISCHE UNTERSUCHUNGEN || VANCOUVER
PROSTATE CENTRE || UNIVERSITÄTSKLINIKUM HAMBURG-EPPENDORF
|| UNIVERSITÄT ZU LÜBECK || UNIVERSITÄT ZÜRICH || UNIVERSITÄT
OSNABRÜCK || UNIVERSITÄT OLDENBURG || UNIVERSITÄT MÜNSTER
|| UNIVERSITÄT MANNHEIM || UNIVERSITÄT MAINZ || UNIVERSITÄT
KÖLN || UNIVERSITÄT HOHENHEIM || UNIVERSITÄT HAMBURG,
INSTITUT FÜR EXPERIMENTALPHYSIK || UNIVERSITÄT HAMBURG,
HAMBURG CENTER FOR HEALTH ECONOMICS || UNIVERSITÄT HAMBURG,
FAKULTÄT BETRIEBSWIRTSCHAFT, LEHRSTUHL FÜR STATISTIK MIT
ANWENDUNGEN IN DER BWL || UNIVERSITÄT HAMBURG, DEPARTMENT
OF ENGLISH LANGUAGE AND LITERATURE/INSTITUT FÜR ANGLISTIK
UND AMERIKANISTIK || UNIVERSITÄT HAMBURG / THE CLUSTER
OF EXCELLENCE: UWA || UNIVERSITÄT HAMBURG - ZENTRUM FÜR
NACHHALTIGES FORSCHUNGSDATENMANAGEMENT || UNIVERSITÄT
HAMBURG & LEIBNIZ INSTITUT FÜR MEDIENFORSCHUNG (HBI ||
UNIVERSITÄT HAMBURG || UNIVERSITÄT ERLANGEN || UNIVERSITY OF
SOUTHERN DENMARK || UNIVERSITY OF NOTRE DAME || UNIVERSITY
OF NORTH CAROLINA AT CHARLOTTE || UNIVERSITY OF MANNHEIM
|| UNIVERSITY OF IRELAND, GALWAY || UNIVERSITY OF HAMBURG ||
UNIVERSITY OF CAMPINAS || UNIVERSITY OF BRITISH COLUMBIA
|| UNIVERSITY MEDICAL CENTER HAMBURG-EPPENDORF (UKE ||
UNIVERSITY HOSPITAL FRANKFURT, GOETHE UNIVERSITY - INSTITUTE FOR
NEUROLOGY || UNIVERSITY GRENOBLE ALPES, GRENOBLE, FRANCE ||
UNIVERSIDAD AUTÓNOMA METROPOLITANA-CUAJIMALPA (UAM-C ||
UNI HAMBURG || UHH, IDGS || UHH - ZFDM || THE NETHERLANDS
CANCER INSTITUTE || TECHNISCHE UNIVERSITÄT HAMBURG ||
TECHNICAL UNIVERSITY OF MUNICH, UNIVERSITY OF HAMBURG, GERMANY
|| TECHNICAL UNIVERSITY OF MUNICH, UNIVERSITY OF HAMBURG ||
TECHNICAL UNIVERSITY OF MUNICH || TUHH/DESY || TUHH
INSTITUTE FOR DATA SCIENCE FOUNDATIONS (E-21 || TUHH || TU
HAMBURG || TU DRESDEN || STAATS- UND UNIVERSITÄTSBIBLIOTHEK
HAMBURG || SIMON FRASER UNIVERSITY || SIGNAL PROCESSING (SP ||
SIGNAL PROCESSING || SCHOOL OF COMPUTING, NEWCASTLE UNIVERSITY,
NEWCASTLE UPON TYNE, UK || STILL GMBH || RESEARCH GROUP
FOR COMPUTATIONAL MOLECULAR DESIGN (AMD || PYO LABS LTD
|| PROFESSUR FÜR NACHHALTIGES WIRTSCHAFTEN || PHD STUDENT
|| PHD CANDIDATE, INDUSTRIAL & ORGANIZATIONAL PSYCHOLOGY
|| OFFIS - INSTITUTE FOR INFORMATION TECHNOLOGY || NXP

SEMICONDUCTORS HAMBURG || NHK SCIENCE & TECHNOLOGY RESEARCH LABORATORIES, UNIVERSITÄT HAMBURG, INSTITUT FÜR DEUTSCHE GEBÄRDENSPRACHE UND KOMMUNIKATION GEHÖRLOSER || NHK SCIENCE & TECHNOLOGY RESEARCH LABORATORIES || MAX-PLANCK-INSTITUT FÜR METEOROLOGIE || MAX PLANCK INSTITUTE OF BIOCHEMISTRY || LUDWIG MAXIMILIAN UNIVERSITY OF MUNICH || LOGISTIK-INITIATIVE HAMBURG || LEUPHANA UNIVERSITÄT LÜNEBURG || LEIBNIZ-INSTITUT FÜR VIROLOGIE (LIV || LEIBNIZ-INSTITUT FÜR MEDIENFORSCHUNG | HANS-BREDOW-INSTITUT || LEIBNIZ UNIVERSITY HANNOVER || LEIBNIZ INSTITUTE FOR MEDIA RESEARCH | HANS-BREDOW-INSTITUT (HBI || LANGUAGE TECHNOLOGY GROUP, UNIVERSITY OF HAMBURG, HAMBURG, GERMANY || LA TROBE UNIVERSITY || JÜLICH SUPERCOMPUTING CENTRE || JOURNALISTIK UND KOMMUNIKATIONSWISSENSCHAFT, UNIVERSITÄT HAMBURG || INSTITUTION FOR COMPUTATIONAL SYSTEMS BIOLOGY | UNIVERSITY OF HAMBURG || INSTITUTE OF STRUCTURAL AND MOLECULAR BIOLOGY, DIVISION OF BIOSCIENCES, UNIVERSITY COLLEGE LONDON || INSTITUTE OF PSYCHOLOGY, UNIVERSITY OF HAMBURG || INSTITUTE OF PATHOLOGY || INSTITUTE OF MEDICAL SYSTEMS BIOLOGY || INSTITUTE FOR DATA SCIENCE FOUNDATIONS - HAMBURG UNIVERSITY OF TECHNOLOGY || INSTITUTE FOR COMPUTATIONAL SYSTEMS BIOLOGY, UNIVERSITY OF HAMBURG, HAMBURG, GERMANY || INSTITUTE FOR COMPUTATIONAL SYSTEMS BIOLOGY || INSTITUT FÜR STATISTIK || INSTITUT FÜR PRODUKTIONSMANAGEMENT UND -TECHNIK || INSTITUT FÜR POLIKLINIK FÜR ALLGEMEINMEDIZIN. UNIVERSITÄTSKLINIKUM HAMBURG EPPENDORF || INSTITUT FÜR NEUROIMMUNOLOGIE UND MULTIPLE SKLEROSE (INIMS || INSTITUT FÜR EXPERIMENTALPHYSIK, UNIVERSITÄT HAMBURG || INSTITUT PASTEUR || INDUSTRIAL AND ORGANIZATIONAL PSYCHOLOGY || IT ANKER GMBH, HAMBURG || IDGS, UNIVERSITÄT HAMBURG || IDGS UNIVERSITÄT HAMBURG || IDGS / UNIVERSITÄT HAMBURG || HOUSE OF COMPUTING AND DATA SCIENCE, UNIVERSITÄT HAMBURG || HOUSE OF COMPUTING AND DATA SCIENCE (HCDS || HOUSE OF COMPUTING AND DATA SCIENCE || HOUSE OF COMPUTING AND DATA SCIENCE || HOUSE OF COMPUTING & DATA SCIENCE || HOCHSCHULE FÜR ANGEWANDTE WISSENSCHAFTEN HAMBURG || HELMUT-SCHMIDT-UNIVERSITÄT/UNIVERSITÄT DER BUNDESWEHR HAMBURG - CHAIR FOR HIGH PERFORMANCE COMPUTING || HELMUT-SCHMIDT-UNIVERSITÄT || HELMUT SCHMIDT UNIVERSITY HAMBURG / UNIVERSITY OF THE FEDERAL ARMED FORCES HAMBURG || HELMHOLTZ-ZENTRUM HEREON || HELMHOLTZ CENTER MUNICH || HAMBURG UNIVERSITÄT || HAMBURG UNIVERSITY || HAMBURG UNIV. OF APPLIED SCIENCES || HAMBURG SCHOOL OF FOOD SCIENCE

|| HAFENCITY UNIVERSITÄT HAMBURG || HAFENCITY UNIVERSITY
HAMBURG || HSU/UNI BW HAMBURG || HSU/UNI BW H || HSU/UNI BW
HAMBURG || HITEC E.V. || HCU || HAFENCITY UNIVERSITÄT
HAMBURG || HAW HAMBURG || GERMAN CLIMATE COMPUTING CENTRE
|| GERMAN CENTER FOR NEURODEGENERATIVE DISEASES || FRAUNHOFER
-CENTER FOR MARITIME LOGISTICS AND SERVICES CML || FRAUNHOFER
SCAI || FRAUNHOFER INSTITUTE OF TRANSLATION MEDICINE AND
PHARMACOLOGY || FRAUNHOFER INSTITUTE || FRAUNHOFER ITMP
SCREENINGPORT || FRAUNHOFER CML || FORSCHUNGSZENTRUM BOSTEL
|| FORSCHUNGSZENTRUM BORSTEL || FACHBEREICH PHYSIK, INSTITUT
FÜR EXPERIMENTALPHYSIK || EUROPEAN XFEL || DIRECTOR OF THE
INSTITUTE FOR COMPUTATIONAL SYSTEMS BIOLOGY UNIVERSITY OF
HAMBURG || DEPARTMENT OF HEALTH ECONOMICS AND HEALTH SER-
VICES RESEARCH, UNIVERSITY MEDICAL CENTER HAMBURG-EPPENDORF,
GERMANY || DKRZ GMBH || DKRZ || DESY/IT || DESY, FS-CI AND
HELMHOLTZ IMAGING || DESY || CHARITÉ BERLIN - INSTITUTE FOR
NEUROPATHOLOGY || CERASCREEN GMBH || CENTRE FOR STRUCTURAL
SYSTEMS BIOLOGY || CENTER FOR INFORMATION SERVICES AND HIGH
PERFORMANCE COMPUTING (ZIH || CENTER FOR DATA AND COMPUTING
IN NATURAL SCIENCE, UNIVERSITÄT HAMBURG || CTC GMBH (AN
AIRBUS COMPANY || CSSB || CSMC || BUCERIUS LAW SCHOOL GMBH
|| ALTIUS INSTITUTE FOR BIOMEDICAL SCIENCES || AKADEMIE DER
WISSENSCHAFTEN IN HAMBURG || ACMETRIC

Program

Program Day 1		
Time	Title	Room
08:30 - 09:00	Registration Welcome Hauke Heekeren, Presidnet of the University of Hamburg Rolf Greve, Head of Department for Science and Research, BWFGB, FHH Christian Pfromm, Chief Digital Officer, Senate Chancellery, FHH Chair: Chris Biemann, House of Computing & Data Science, UHH	Foyer
09:00 - 09:45	Language: DE	Audimax 1
09:45 - 10:15	Coffee Break / Poster Session Panel discussion Andreas Timm-Giel, Technical University of Hamburg Annette Hamann, ITEC e.V. and Beiersdorf Blanche Schwappach-Pignataro, University Hospital Hamburg-Eppendorf Cord Jakobeit, Academy of Sciences in Hamburg Philipp Moll, Max Planck Institute for the Structure and Dynamics of Matter Rolf Greve, Head of Department for Science and Research, BWFGB, FHH Wolfgang Schulz, Leibniz Institute for Media Research Hans-Bredow-Institut Chair: Katharina Schneider, Tagesspiegel	Tent
10:15 - 11:45	Language: DE	Audimax 1
11:45 - 13:15	Lunch Break PIER PLUS - Institutions and Digital Transformation (1/2) University of Hamburg Speaker: Tilo Böhmann Title: Empowering the Metropolitan Region - the Role of the University of Hamburg in the Digital Transformation of Research Language: EN Leibniz Institute of Virology Speaker: Wolfram Brune Title: "Viral Data Science at LIV" Language: EN Bernhard Nocht Institute of Tropical Medicine Speaker:: Jürgen May Title: "Tropical Medicine in the Digital Age" Language: EN University Hospital Hamburg-Eppendorf Speaker:: Frank Ückert Title: "From Digitization to AI: Challenges of and Opportunities at the UKE" Language: DE Fraunhofer IAPT Speaker: Frank Beckmann Title: "AI and Big Data as an Enabler for the Industrialization of Additive Production"	Campus
	Title: "AI and Big Data as an Enabler for the Industrialization of Additive Production" Language: DE Fraunhofer IAP/CAN Speaker: Christoph Gimmler Title: "CAN - Digital Roadmap: Nanotechnology, Research Data Management and AI" Language: DE Fraunhofer ITMP Speaker: Reagon Karki and Yojana Gadiya Title: "FAIR by Design with Knowledge Graphs" Language: EN Fraunhofer CML Speaker: Anisa Rizvanolli Title: "Data, digitization and optimisation in maritime logistics" and "Industrial Application Center Quantum Computing Hamburg" Language: DE	
13:15 - 15:15	Chair: Elisabeth Hettig, PIER PLUS, & Marion Stange, PIER Cross-Institutional Initiatives ahoi.digital - The Alliance of Hamburg Universities of Informatics Speaker: Joerg Noennig Language: DE The Center for Data and Computing in Natural Sciences (CDCS)	Audimax 1

	<p>Speaker: Annett Ungethüm Language: DE</p> <p>DASHH - Data Science in Hamburg - Helmholtz Graduate School for the Structure of Matter Speaker: Christiane Ehrt Language: EN</p> <p>MLE-TUHH - Machine Learning in Engineering Speaker: Gregor Vonbun-Feldbauer et al. Language: DE / EN</p> <p>Network of Labs Speaker: Janick Edinger, Martin Kohler Language: DE</p> <p>HITeC - Hamburger Informatik Technologie-Center e.V. Speaker: Lothar Hotz Language: DE</p> <p>ARIC - Artificial Intelligence Center Hamburg e.V. Speaker: Florian Vogt Language: DE</p> <p>Hamburg Innovation GmbH Speaker: Philipp Walter Language: DE</p>	
	<p>IKS Hamburg Innovations Contact Speaker: Florian Vogt Language: DE</p> <p>KIEZ of computing & data science networking platform Speaker: Anna Reinicke-Vogt Language: EN</p>	
13:15 - 15:15	Chair: Lothar Hotz, HITeC	Audimax 2
15:15 - 16:15	Coffee Break / Poster Session	Tent
	<p>PIER PLUS - Institutions and the Digital Transformation (2/2)</p> <p>German Electron Synchrotron DESY Speaker: Beate Heinemann Title: "Digital DESY" Language: EN</p> <p>European X-Ray Free-Electron Laser Facility Speaker: Steve Aplin Title: "Data Science at the Petabyte Scale for X-Ray Free Electron Lasers" Language: EN</p> <p>Max Planck Institute for the Structure and Dynamics of Matter Speaker: Philipp Moll Title: "Electronics beyond silicon: quantum materials research at the MPI for the Structure and Dynamics of Matter" Language: EN</p>	

	<p>Hamburg University of Technology Speaker: Nihat Ay Title: "Cyber-Physical & Cognitive Systems " Language: EN</p> <p>Helmut Schmidt University of Hamburg Speaker: Oliver Niggemann Title: "AI Research at Helmut Schmidt University" Language: DE</p> <p>Helmholtz Centre Hereon Speaker: Matthias Rehann Title: "Digitalization at the Hereon" Language: DE</p> <p>Max Planck Institute for Meteorology Speaker: Daniel Klocke Title: "Understanding the climate system using frontier compute technology" Language: EN</p> <p>German Institute for Global and Area Studies Speaker: Peter Peetz Title: "Digitalisation and the Global South: Research and Knowledge Transfer at the German Institute for Global and Area Studies (GIGA)" Language: DE</p> <p>Leibniz Institute for Media Research</p>	
--	--	--

16:15 - 18:15	<p>Speaker: Gregor Wiedemann Title: "Surveying Digital Media Ecosystems - Computational Communication Science at the Leibniz Institute for Media Research Hans-Bredow-Institut" Language: EN</p> <p>Chair: Marion Stange, PIER, & Elisabeth Hettig, PIER PLUS</p>	Audimax 1
16:15 - 18:15	<p>Transfer and Practice -</p> <p>Foundation ecosystem Hamburg</p> <p>Title: "Startup Factory Hamburg - a beacon for the Hamburg startup ecosystem?" Format: Interactive dialogue format with the audience</p> <p>Speaker: of Hamburg Bakr Fadl, Managing Director of the Transfer Agency of the University of Hamburg Jonas Wagner, Project staff at Startup Factory, Hamburg Innovation (HI)</p> <p>Language: DE</p>	Audimax 2
18:15 - 22:15	<p>Chair: Bakr Fadl, UHH, & Jonas Wagner, HI</p> <p>Social event</p>	Foyer
Program Day 2		
Time	Title	Room
08:30 - 09:30	Registration	Foyer

09:30 - 10:30	<p>Welcome and Keynote</p> <p>of Hamburg Laura Edinger-Schons, Chief Sustainability Officer of the University of Hamburg Sebastian Gerling, Chief Digital Officer of the University of Hamburg Keynote entitled "Digital Research Units" by Dirk Nowotka, Professor for Reliable Systems & Director Digital Science Center, Christian-Albrechts-Universität zu Kiel</p> <p>Chair: Chris Biemann, House of Computing & Data Science, UHH</p>	Audimax 1
10:30 - 11:15	<p>languages: DE</p> <p>Coffee Break / Poster Session</p> <p>Cross-Disciplinary Labs</p> <p>Maximizing the efficacy of early rhythm-control therapy for atrial fibrillation in patients with and without heart failure Speaker: Benedikt Schrage Language: EN</p> <p>Computational Human Dynamics (CHD) Speaker: Navin Navin Laxminarayanan Raj Prabhu & Marvin Grabowski Language: EN</p> <p>FLIGHT - Federated Learning-Guided digital Health Speaker: Michael Brehler Language: EN</p>	Tent

11:15 - 12:45	<p>In Search of Inflation Narratives: Combining Approaches from Economics and Computer Science. Speaker: Ulrich Fritsche Language: EN</p> <p>Predicting COVID-19 Vaccination Uptake from Public Discourse: A Machine Learning Approach Speaker: Esra Eren Bayindir Language: EN</p> <p>MuMokA - Multimodal modelling of cultural artifacts in digital space Speaker: Julia Nantke Language: DE</p> <p>Chair: Katrin Schöning-Stierand, House of Computing & Data Science, UHH</p> <p>Associated Cross-Disciplinary Labs</p> <p>ITMC-SAIL: Co-creating Knowledge on AI in Practice Speaker: Tom Lewandowski Language: DE</p> <p>Digital Humanities Lab Speaker: Heike Zinsmeister</p> <p>AI in the loop: Analysing Digital Discourses with the D-Wise Tool Suite Speaker: Gertraud Koch</p> <p>bAlome: shaping biomedical AI research at the UKE Speaker: René Werner</p>	Audimax 1
---------------	---	-----------

<p>11:15 - 12:45 12:45 - 14:00</p>	<p>Digital discovery machines: Machine learning in astrophysics and particle physics Speaker: Marcus Brüggem</p> <p>CDL2: Cross-disciplinary research with photon science applications Speaker: Anton Barty</p> <p>Proteins in silico: biology in the age of deep learning Speaker: Karen Malanastas-Cantos</p> <p>Dependable CPS: Computational Controls of Accelerators - one Asset in CDCS Speaker: Goerschwin Fey</p> <p>Language: EN</p> <p>Chair: Martin Semmann, House of Computing & Data Science, UHH</p> <p>Lunch Break</p> <p>Cluster of Excellence (1/2)</p> <p>Cluster of Excellence CLICCS Climate + AI Speaker: Christopher Kadow Climate Visualization Speaker: Marc Rautenhaus Climate in Digital Media Speaker: tba Organizational Fields on the WWW Speaker: Achim Oberg</p>	<p>Audimax 2 Campus</p>
<p>14:00 - 15:30</p>	<p>Cluster of Excellence AIM</p> <p>Data Science for the Advanced Imaging of Matter Speaker: Henry Chapman, Kartik Ayyer</p> <p>Chair: Jessica Tiedke, UHH</p> <p>Language: EN</p> <p>Computing & Data Science Offerings</p> <p>The House of Computing and Data Science - Digitalization of Research through Research Speaker: Martin Semman Language: DE</p> <p>Informatics - The core discipline of digitalisation Speaker: Hannes Federrath Language: DE</p> <p>Digital Total, Digital Transformation in the WISO Research Laboratory Speaker: Olaf Bock Language: DE</p> <p>The eScience Office - Digital Research at the Faculty of Educational Science Speaker: Michael Wuppermann Language: DE</p> <p>High-performance infrastructures in the RRZ for excellent research</p>	<p>Audimax 1</p>
<p>14:00 - 15:30 15:30 - 16:00</p>	<p>Speaker: Markus Böttger Language: DE</p> <p>Supporting Open Research: The Center for Sustainable Research Data Management Speaker: Kai Wörner Language: EN</p> <p>Science as distributed open-source knowledge development Speaker: Lennart Wittkuhn Language: EN</p> <p>Chair: Katrin Schöning-Stierand, House of Computing & Data Science, UHH</p> <p>Coffee Break / Poster Session</p>	<p>Audimax 2 Tent</p>
	<p>Cluster of Excellence (2/2)</p> <p>Cluster of Excellence QU Digitalization and AI in a Quantum Universe Speaker: Timo Weigand/ Gregor Kasieczka</p> <p>Cluster of Excellence UWA</p> <p>Speaker: Kaja Harter-Uibopuu/ Hussein Mohammed</p> <p>Material, Interaction and Transmission in Manuscript Cultures Digital Data Handling at UWA Chemometrics / Bioinformatics Computed Tomography</p>	

	Immersive Incribed Spaces Visual-Pattern Analysis	
	Chair: Jessica Tiedke, UHH	
16:00 - 17:30	Language: EN	Audimax 1
	Digital Teaching, Digital and Data Literacy & Data Science at UHH	
	(HUL).	
	"Digital and Data Literacy at UHH": Katharina Kleinen-von Königsłow, Julia Niemann-Lenz, Moritz Kreinsen and Stephan Leible (DDLitLab).	
	"Data Science at UHH": Timo Gerkmann (Institute of Computer Science), Gregor Kasieczka (Institute of Experimental Physics), Katharina Kleinen-von Königsłow (WISO Faculty).	
16:00 - 17:30	Chair: Katharina Kleinen-von Königsłow, DDLitLab	Audimax 2
	Awards	
	Sustainability Award, Laura Edinger-Schons, Sustainability Officer of the Univerisität Hamburg	
	Audience Award, Chris Biemann, House of Computing & Data Science, UHH	
17:30 - 17:45		Audimax 1
17:45 - 18:00	Closing	Audimax 1

Using PEMT for pharmaceutical patent landscaping

Author: Yojana Gadiya¹

Co-authors: Andrea Zaliani¹; Matrin Hofmann-Apitius²; Philip Gribbon¹

¹ *Fraunhofer ITMP ScreeningPort*

² *Fraunhofer SCAI*

Corresponding Authors: philip.gribbon@itmp.fraunhofer.de, yojana.gadiya@itmp.fraunhofer.de, andrea.zaliani@itmp.fraunhofer.de, martin.hofmann-apitius@scai.fraunhofer.de

Patents serve as critical catalysts in drug discovery, assuring legal safeguards to innovations while stimulating investments. By discerning patterns within patent data, researchers glean invaluable insights into pharmaceutical industry trends and priorities. Our study, powered by the Patent Enrichment Tool (PEMT), delves into patent literature concerning rare diseases (RD) and Alzheimer's disease (AD). We commence with an exploration of organizations engaged and gained insights into their significance from distinct perspectives. We subsequently analyze historical patent focus on therapeutic targets, correlating them with market dynamics to identify notable disease targets. Finally, we identify repurposed drugs using patent information. This study underscores patent documents' broad applicability, transcending legality to enrich drug discovery, design, and research—a valuable resource for future endeavors.

Keywords:

patent landscaping

PEMT

patent profiling

patent trend annotation

Link to contribution: [Using PEMT for pharmaceutical patent landscaping](#)

Elucidation of epigenetic mechanisms underlying normal and disease biology

Author: Sheraz Gul¹

Co-authors: Andrea Zaliani ¹; Carsten Claussen ¹; Ferdinand Seum ¹; Jaleed Aslam ²; Olaf Schneider ²; Philip Groth ²; Sung-Min Pyo ³; Tara Beirne ⁴; Wilfried Blum ²

¹ *Fraunhofer Institute*

² *Cerascreen GmbH*

³ *Pyo Labs Ltd*

⁴ *University of Ireland, Galway*

Corresponding Authors: jaleed.aslam@cerascreen.de, wilfried.blum@cerascreen.de, olaf.schneider@cerascreen.de, tarabeirne6@gmail.com, ferdinand.seum@itmp.fraunhofer.de, sheraz.gul@itmp.fraunhofer.de, carsten.claussen@itmp.fraunhofer.de, mail@pyo-labs.com, philip.groth@cerascreen.de, andrea.zaliani@itmp.fraunhofer.de

Epigenetics is defined as the study heritable alterations in gene function that occur without underlying changes in DNA sequence. Biologically, aging is linked with a gradual increase in molecular and cellular damage eventually leading to a decline in physiological reserves and an increased risk of developing diseases. It is now well established that a vast number of epigenetic changes, notably DNA methylation occur during the aging process. Epigenetic changes offer the potential for maintaining health or reversing diseases by modulating the function of druggable proteins coded by genes with dynamic and reversible methylation changes.

Keywords:

Epigenetics

Aging

Drug discovery

Disease

Drug targets

Link to contribution: [Elucidation of epigenetic mechanisms underlying normal and disease biology](#)

Reproducible Knowledge Graph workflows for embedding chemical entities and associated biology of diseases: A use case in Mpox

Authors: Reagon Karki¹; Andrea Zaliani²; Yojana Gadiya²; Philip Gribbon²

¹ *Fraunhofer Institute of Translation Medicine and Pharmacology*

² *Fraunhofer ITMP ScreeningPort*

Corresponding Authors: yojana.gadiya@itmp.fraunhofer.de, an-
drea.zaliani@itmp.fraunhofer.de, philip.gribbon@itmp.fraunhofer.de,
reagon.karki@itmp.fraunhofer.de

The outbreak of Mpox virus (MPXV) infection in May 2022 is declared a global health emergency by WHO. The MPXV pathophysiology and its underlying mechanisms are not yet understood. Likewise, the knowledge of biochemicals and drugs used against MPXV and their downstream effects is sparse. Using systems biology approaches, we used Knowledge Graph (KG) representations to depict chemical and biological aspects of MPXV. We collected and rationally assembled biological study results, assays, drug candidates, and pre-clinical evidence to form a dynamic and comprehensive network. The KG is compliant with FAIR principles allowing seamless transformation and integration to/with other formats and infrastructures. Its implementation in Mpox was the first usecase of this workflow and can be used for other infectious diseases.

Keywords:

knowledge graph
mpox
infectious disease

Link to contribution: [Reproducible Knowledge Graph workflows for embedding chemical entities and associated biology of diseases: A use case in Mpox](#)

Redirected Walks through the European XFEL

Authors: Eike Langbehn¹; Reinhard Bacher²

¹ *Hochschule für Angewandte Wissenschaften Hamburg*

² *DESY*

Corresponding Author: eike.langbehn@haw-hamburg.de

At modern particle accelerators, robots are playing an increasingly important role in facility inspection and maintenance to significantly improve the working conditions of technicians and scientists. Robot teleoperation using virtual reality (VR) technology offers a faster, more accurate, and more secure way to teleoperate robots in hazardous environments. Therefore, VR interaction techniques are required that support natural and direct movement controls. Redirected walking (RDW) is a VR locomotion technique that enables free natural walking in confined real world tracking spaces by imperceptibly guiding the user on paths in the real world that differ from those perceived in the virtual environment.

We integrated RDW algorithms into a robotic platform and evaluated this setup in a maintenance task at DESY's accelerator tunnels.

Keywords:

Virtual Reality
Human-Robot-Interaction
Teleoperation

Link to contribution: [Redirected Walks through the European XFEL](#)

Ein Intelligentes Fenster zum Hamburger Hafen: Entwicklung einer Augmented Reality Fensterscheibe für Schiffe und Infotainment-Installationen

Author: Jan Jessen¹

Co-author: Eike Langbehn¹

¹ *Hochschule für Angewandte Wissenschaften Hamburg*

Corresponding Authors: eike.langbehn@haw-hamburg.de, jan.jessen@haw-hamburg.de

Wir haben eine Augmented Reality Fensterscheibe entwickelt, die als Visualisierungsassistentz auf Schiffsbrücken oder zu Infotainment-Zwecken als Outdoor-Ausstellungsstück eingesetzt werden kann. Die Umsetzung umfasst AIS-Datenempfang, Tiefenkamera zur Betrachtererfassung und ein T-OLED-Display. Durch die Anzeige von perspektivkorrekten Inhalten im Hintergrund auf einem transparenten Display im Vordergrund soll eine immersive Erfahrung geschaffen werden. Die AR-Fensterscheibe ermöglicht es den Nutzern, den eigentlichen Inhalt zu betrachten, ohne auf einen externen Bildschirm (wie z.B. ein Smartphone) angewiesen zu sein. Beispielhaft entwickeln wir eine Anwendung für die AR-Fensterscheibe, die die Bewegungen der Schiffe (und Zusatzinformationen wie Herkunftsland, Zielort, usw.) im Hamburger Hafen abbildet und die für eine Outdoor-Ausstellung geeignet ist.

Keywords:

Augmented Reality
Smart Window
Harbour

Link to contribution: [Ein Intelligentes Fenster zum Hamburger Hafen: Entwicklung einer Augmented Reality Fensterscheibe für Schiffe und Infotainment-Installationen](#)

Facilitating FAIR Data Practice: The Research Data Repository for the Universität Hamburg

Author: Kai Joachim Wörner¹

¹ *Universität Hamburg*

Corresponding Author: kai.woerner@uni-hamburg.de

Following the FAIR principles helps make data findable by humans and machines, interoperable through standard formats and metadata, accessible through open access when possible, and reusable through clear licensing and richness of metadata, supporting improved understanding, reuse, and reproducibility of scientific data.

While following these principles requires the willingness to embrace such a paradigm, it also relies on the technical infrastructure to support researchers willing to do so.

For researchers of the Universität Hamburg, the Center for Sustainable Research Data management has established the Research Data Repository that is built with future-proof technology and based on recommended data handling practices to support researchers to preserve and share their data in a FAIR way.

Keywords:

fair data
research data management
open science
repositories
research data

Link to contribution: [Facilitating FAIR Data Practice: The Research Data Repository for the Universität Hamburg](#)

CRIS at UHH: digital, structured, sustainable

Authors: Lisa Brase¹; Steffen Grünler²

¹ *UHH - ZFDM*

² *Universität Hamburg - Zentrum für nachhaltiges Forschungsdatenmanagement*

Corresponding Authors: lisa.brase@uni-hamburg.de, steffen.gruenler@uni-hamburg.de

The current research information system (CRIS) of the UHH is the digital system where metadata of research activities is collected and linked in a structured way. The combined metadata in this one system creates a digital network of the actual research landscape at UHH. This network provides a fundamental base and an important tool for improving the findability, management and evaluation of research information. The widespread use of the system by the researchers and administrative staff leads to a steady increase in the quality of the university's data. The usability of the system also ensures the sustainable use of research information. The presentation of activities in the institutional 'FIS Portal' increases the visibility of researchers and university's research itself.

Keywords:

CRIS
FIS

Link to contribution: [CRIS at UHH: digital, structured, sustainable](#)

Pattern Analysis Software Tools

Author: Hussein Mohammed¹

¹ *Universität Hamburg / The Cluster of Excellence: UWA*

Corresponding Author: hussein.adnan.mohammed@uni-hamburg.de

The research of ancient written artefacts results in an ever-increasing amount of digital data in various forms, ranging from raw images of artefacts to automatically generated data from advanced acquisition techniques. The manual analysis of this data is typically time-consuming and can be subject to human error and bias. Therefore, a set of Pattern Analysis Software Tools (PAST) has been developed to automate this process. These software tools have been developed to facilitate a more efficient study of written artefacts and to help scholars benefit from the potential of pattern analysis. Each tool is developed and tested in close collaboration with experts from relevant fields of research to ensure its usability and applicability to actual research questions.

Keywords:

Pattern Analysis

Software Tools

Historical Documents Analysis

Link to contribution: [Pattern Analysis Software Tools](#)

Computational Analysis of Handwriting Style

Author: Hussein Mohammed¹

¹ *Universität Hamburg / The Cluster of Excellence: UWA*

Corresponding Author: hussein.adnan.mohammed@uni-hamburg.de

Images of digitised historical written artefacts contain much more information than the mere textual content found in their transcriptions, such as the texture of writing support, the general visual layout, and the non-textual visual elements. Among these aspects, the style of handwriting itself can tell us about the scribe, the approximate place of origin, and perhaps even help date the artefact itself. Therefore, handwriting style analysis is one of the very important topics in our research. To meet the different requirements of our research questions, we developed a novel training-free classifier, as well as a similarity measurement approach for the tasks of classifying the handwriting style, scribe identification, and measuring the similarity between handwritten samples.

Keywords:

Visual-Pattern Classification
Handwriting Style Analysis
Writer Identification
Historical Document Analysis

Link to contribution: [Computational Analysis of Handwriting Style](#)

Digital Data Handling at UWA

Author: Sylvia Melzer¹

Co-authors: Stefan Thiemann ¹; Ralf Möller ²

¹ *Universität Hamburg*

² *Universität zu Lübeck*

Corresponding Authors: sylvia.melzer@uni-hamburg.de, moeller@uni-luebeck.de, stefan.thiemann@uni-hamburg.de

Around 150 researcher, 40 academic disciplines, 60 research projects, and 11 research fields exist at the Cluster of Excellence UWA (Understanding Written Artefact). The size of the research field shows, it is a challenge to reconcile the many researchers with their different needs. The Research Field Data Linking has taken on the task of establishing research data management in the humanities in a sustainable, FAIR and CAREful way; developing IT approaches to support the reusability of research data and; making newly developed AI algorithms and other computer science approaches easy to use for humanities researchers; and adapting these algorithms to the research questions in the humanities.

Keywords:

Digital Data Handling
Sustainable Research Data Management
Written Artefacts

Link to contribution: [Digital Data Handling at UWA](#)

Few-shot learning for automated content analysis (FLACA) in the German media debate on arms deliveries to Ukraine

Authors: Jonas Rieger¹; Mattes Ruckdeschel¹; Kostiantyn Yanchenko²; Gerret von Nordheim²; Katharina Kleinen von Königslöw²; Gregor Wiedemann¹

¹ *Leibniz Institute for Media Research | Hans-Bredow-Institut (HBI)*

² *Universität Hamburg*

Corresponding Authors: kostiantyn.yanchenko@uni-hamburg.de,
rieger@statistik.tu-dortmund.de, katharina.kleinen@uni-hamburg.de,
g.wiedemann@leibniz-hbi.de, gerret.vonnordheim@uni-hamburg.de,
m.ruckdeschel@leibniz-hbi.de

The use of pre-trained language models based on transformer neural networks has significantly advanced the field of NLP and offers considerable potential for improving automatic content analysis, e.g., in communication science, where their widespread adoption is still limited. In our poster, we highlight challenges and promises by employing transformer models combined with parameter-efficient few-shot fine-tuning to need less labeled data in complex annotation tasks using automated procedures. The results indicate a noteworthy zero-shot understanding of ChatGPT of our definitions of claims and arguments while our tailormade few-shot methods outperform it using a medium number of human annotations as training data.

Keywords:

natural language processing
large language models
few-shot
annotation
content analysis

Link to contribution: [Few-shot learning for automated content analysis \(FLACA\) in the German media debate on arms deliveries to Ukraine](#)

Social Media Observatory

Authors: Felix Victor Münch¹; Gregor Wiedemann¹; Jan Rau; Jan-Hinrik Schmidt; Philipp Kessling

¹ *Leibniz-Institut für Medienforschung | Hans-Bredow-Institut*

Corresponding Authors: f.muench@leibniz-hbi.de, g.wiedemann@leibniz-hbi.de, p.kessling@leibniz-hbi.de, j.rau@leibniz-hbi.de

In unserem Poster stellen wir das Social Media Observatory (SMO) am Leibniz-Institut für Medienforschung | Hans-Bredow-Institut vor. Das SMO wird seit 2020 als Open-Science-Forschungsinfrastruktur innerhalb des Forschungsinstitut Gesellschaftlicher Zusammenhalt (FGZ) entwickelt. Es konzentriert sich auf die langfristige Beobachtung der öffentlichen Kommunikation auf ausgewählten Plattformen und Online-Nachrichtenmedien zur Beantwortung sozialwissenschaftlicher Forschungsfragen. Auf der Grundlage systematisch zusammengestellter Listen öffentlicher Sprecherinnen-Kategorien, wie beispielsweise Parlamentarierinnen oder Medienorganisationen, werden sowohl statistische als auch inhaltliche Daten erhoben, um den deutschen Social Media-Diskurs im Vergleich zu redaktionellen Massenmedien zu untersuchen. Aggregierte Ergebnisse werden über interaktive Dashboards veröffentlicht. Einem Do-it-yourself-Ansatz folgend stellt das SMO außerdem verschiedene Tools, kuratierte Datensätze und dokumentierte Workflows zur Verfügung, um beispielsweise thematische Ad-hoc-Datensammlungen durchzuführen.

Keywords:

Soziale Medien
Digitale Spurendaten
Forschungsethik
Big Data
Forschungsinfrastruktur

Link to contribution: [Social Media Observatory](#)

Augmenting Social Media Actor Lists with Explorative Network Sampling: Mapping German Right-wing Counterpublics on Twitter

Authors: Felix Victor Münch¹; Philipp Kessling¹; Jan Rau¹

¹ *Leibniz-Institut für Medienforschung | Hans-Bredow-Institut*

Corresponding Authors: p.kessling@leibniz-hbi.de, j.rau@leibniz-hbi.de, f.muench@leibniz-hbi.de

Our research addresses the challenge of accurately classifying far-right extremist accounts on Twitter/X or comparable platforms on a large scale to provide a comprehensive view of the digital far-right ecosystem. Traditional approaches rely on precise but limited account lists, leading to incomplete insights, especially for new or less popular accounts. In response, we present a mixed method approach using network sampling, utilising our self-developed tool, Ponyexpress. We apply a variety of topology-based network sampling methods to sample accounts within right-wing counterpublics in the German Twittersphere. The study evaluates the precision and recall of these methods and maps the results to the overall network structure. Our findings indicate the effectiveness of specific sampling strategies such as combining k-cores with snowball sampling.

Keywords:

data mining
digital traces
social networks
network sampling
right-wing counterpublics

Link to contribution: [Augmenting Social Media Actor Lists with Explorative Network Sampling: Mapping German Right-wing Counterpublics on Twitter](#)

Supporting Sustainable Research: The Services of the Center for Sustainable Research Data Management

Authors: Alexander Heinrich¹; Hagen Peukert¹; Iris Vogel¹; Juliane Jacob¹; Kai Joachim Wörner¹; Lisa Brase²; Stefan Thiemann¹; Steffen Grünler³

¹ *Universität Hamburg*

² *UHH - ZFDM*

³ *Universität Hamburg - Zentrum für nachhaltiges Forschungsdatenmanagement*

Corresponding Authors: alexander.heinrich@uni-hamburg.de, kai.woerner@uni-hamburg.de, lisa.braser@uni-hamburg.de, steffen.gruenler@uni-hamburg.de, iris.vogel@uni-hamburg.de, hagen.peukert@uni-hamburg.de, juliane.jacob@uni-hamburg.de, stefan.thiemann@uni-hamburg.de

The Center for Sustainable Research Data Management (RDM Center) is a central institution of the Universität Hamburg that bundles infrastructures, competencies, and tasks to offer sustainable services to researchers.

The poster will provide an overview of what these services are and how they can be helpful in the everyday tasks researchers are faced with.

The services are, among others, the Research Data Repository for the FAIR handling of research data, the FUNDus! Collections Portal with which scientific collection objects of the Universität Hamburg are generally accessible, an extensive portfolio of training offerings as well as the research information system (RIS) in which metadata about research projects, publications, and references to research data are collected and offered for research.

Keywords:

research data
research information
sustainability
research support
services

Link to contribution: [Supporting Sustainable Research: The Services of the Center for Sustainable Research Data Management](#)

Visual-Pattern Detection in Historical Documents

Author: Hussein Mohammed¹

¹ *Universität Hamburg / The Cluster of Excellence: UWA*

Corresponding Author: hussein.adnan.mohammed@uni-hamburg.de

Automatic pattern detection has become increasingly important for scholars in the humanities as the number of manuscripts that have been digitised has grown. Nevertheless, this task can be challenging for state-of-the-art computer vision approaches due to the lack of annotations and the small size of many patterns, which can be smaller than 0.1% of the image size. Therefore, we developed a training-free approach in order to overcome the first challenge, and a training-based approach in order to overcome the second challenge. Both approaches have been evaluated using research data from digitised manuscripts, and both yielded state-of-the-art results.

Keywords:

Visual-Pattern Detection

Object Detection

Historical Documents Analysis

Link to contribution: [Visual-Pattern Detection in Historical Documents](#)

Palimpsests inpainting with Generative AI

Author: Hussein Mohammed¹

Co-authors: Jost Gippert ²; Mahdi Jampour ²

¹ *Universität Hamburg / The Cluster of Excellence: UWA*

² *Universität Hamburg*

Corresponding Authors: mahdi.jampour@uni-hamburg.de, jost.gippert@uni-hamburg.de, hussein.adnan.mohammed@uni-hamburg.de

A palimpsest is a manuscript page that has been scraped or washed off in preparation for reuse, in the form of another document. When working with palimpsest manuscripts, even with the employment of Multispectral Imaging (MSI), the under-text often remains challenging to read. In some cases, it might not be readable at all. This difficulty arises due to various factors, including parts of the under-text covered by the over-text and other visual elements. However, addressing these challenges opens the door to innovative solutions, one of which is the technique of inpainting using generative AI approaches to restore the lost parts of handwriting from the undertext of palimpsest manuscripts.

Keywords:

Inpainting

Generative AI

Palimpsests

Historical Documents Analysis

Link to contribution: [Palimpsests inpainting with Generative AI](#)

Profiling of written artefacts with bioinformatic and chemometric methods

Authors: Lucas F. Voges¹; Stephan Seifert¹

¹ CSMC

Corresponding Authors: stephan.seifert@uni-hamburg.de, lucas.voges@uni-hamburg.de

Written artefacts are studied at the CSMC with the participation of scientists from a variety of disciplines using a wide range of material analysis methods. We contribute to this collaboration with various data analysis techniques, including machine learning and statistical approaches. They are aiming at the classification of the written artefacts based on different properties, e.g. production materials, preservation states and environmental conditions to contribute to the characterization of their biological identity and historical background.

Here we introduce the analysis of DNA data with bioinformatic approaches and the analysis of data obtained by infrared spectroscopy, for which chemometric approaches are applied.

Keywords:

Chemometrics
Bioinformatics
Machine Learning
Written Artefacts

Link to contribution: [Profiling of written artefacts with bioinformatic and chemometric methods](#)

Introduction of the research group "Chemometrics of complex material systems"

Author: Stephan Seifert¹

¹ *Hamburg School of Food Science*

Corresponding Author: stephan.seifert@uni-hamburg.de

In Stephan Seifert's research group "Chemometrics of Complex Material Systems" at the Hamburg School of Food Science, analytical data corresponding to fingerprints of biological samples are analysed. These fingerprints come, for example, from food or written artefacts, which are classified using machine learning methods based on various properties, such as their geographical origin. The methodological focus of the group is on the development of random forest approaches, e.g. for variable selection and the analysis of the underlying mechanisms, as well as on the fusion of different analytical data.

Keywords:

chemometrics
bioinformatics
machine learning
random forest

Link to contribution: [Introduction of the research group "Chemometrics of complex material systems"](#)

Smart and sustainable cities? Testing a Socio-Technical Transitions Model of the Interplay of Urban Digitalization and Sustainable Development

Author: Laura Marie Edinger-Schons¹

¹ *Professur für Nachhaltiges Wirtschaften*

Corresponding Author: laura.marie.edinger-schons@uni-hamburg.de

Facing challenges of climate change and urbanization cities worldwide turn to new means of digitalization aiming to increase city sustainability. However, the way urban digitalization affects sustainability is less clear. Drawing on socio-technical transitions theory we introduce a model of the dynamic interplay of digital and sustainable transformation. We aim to test that model using longitudinal data on European cities from Eurostat and OECD. Our research aims to enrich the current scholarly debate on the interplay of digitalization and sustainable development and guide practitioners aiming to meaningfully align digital and sustainable transformation in their cities. . By presenting our research approach in a poster format we hope to receive valuable feedback on our theoretical model and the envisioned methodological approach.

Keywords:

Twin Transformation, Socio-Technical Transitions Theory, Smart and Sustainable Cities

Link to contribution: [Smart and sustainable cities? Testing a Socio-Technical Transitions Model of the Interplay of Urban Digitalization and Sustainable Development](#)

The Role of Digitalization for Firm Financial and Ecological Performance: A Socio-technical Systems Perspective

Author: Alicia Minnerup¹

Co-authors: Caroline Ruiner ²; Frederik Maibaum ³; Johann Nils Foege ³; Johannes Kriebel ⁴; Laura Marie Edinger-Schons ⁵; Sascha Alavi ⁶

¹ *University of Mannheim*

² *Universität Hohenheim*

³ *Leibniz University Hannover*

⁴ *Universität Münster*

⁵ *Universität Mannheim*

⁶ *Universität Oldenburg*

Corresponding Authors: frederik.maibaum@wa.uni-hannover.de,
laura.marie.edinger-schons@uni-hamburg.de, caroline.ruiner@uni-
hohenheim.de, nils.foege@wa.uni-hannover.de, sascha.alavi@uni-oldenburg.de,
johannes.kriebel@wiwi.uni-muenster.de, alicia.minnerup@mail.uni-
mannheim.de

Through the lens of Socio-Technical Systems theory, this study scrutinizes the impact of digitalization on both the financial and the ecological aspect of firm performance. We propose that digitalization, i.e., a significant modification in the technical subsystem of a firm, can amplify a firm's financial and ecological performance only if it is accompanied with an equivalent shift in its social subsystem, i.e., implementing measures for organizational support. We measure digitalization of companies through an approach based on the textual content of annual reports and natural language processing. Results of our panel regression analyses support our theorizing that the positive influence of digitalization on a firm's financial and ecological performance is contingent on the provision of sufficient organizational support for employees.

Keywords:

Digitalization, Financial Performance, Ecological Performance, Organizational Support, Socio-Technical Systems Theory

Link to contribution: [The Role of Digitalization for Firm Financial and Ecological Performance: A Socio-technical Systems Perspective](#)

Assessment of antibiotic resistance from microbial genomic data

Authors: Lennart Bartels¹; Inken Wohlers²

¹ *Forschungszentrum Borstel*

² *Forschungszentrum Bostel*

Corresponding Authors: lbartels@fz-borstel.de, iwohlers@fz-borstel.de

Increasing antibiotic resistance of disease-causing microbes poses a major public health problem. Resistance is caused by evolutionary changes in the microbial genome that inactivate the antibiotic's molecular mechanism of action. For diseases such as tuberculosis, single genetic variants that confer resistance are known and considered in treatment regimes. For other pathogens, relationships between variants and resistance are often not known and assessed by time-consuming lab-experimental assays. Here we focus on infection-causing *Haemophilus influenzae*. We perform regression analyses to identify genetic variants that are associated with lab-based resistance measurements. We find that single variants are not sufficient to explain resistance. Thus, we are currently identifying associated variant combinations and investigate suitable machine learning approaches for predicting resistance from the entire genome.

Keywords:

microbial genome-wide association study
antibiotic resistance
genetic variants
prediction

Link to contribution: [Assessment of antibiotic resistance from microbial genomic data](#)

Biomolecular Data Science in Pneumology

Authors: Lennart Bartels¹; Inken Wohlers²

¹ *Forschungszentrum Borstel*

² *Forschungszentrum Bostel*

Corresponding Authors: iwohlers@fz-borstel.de, lbartels@fz-borstel.de

The Data Science group at Research Center Borstel, Leibniz Lung Center performs cutting-edge research on data generated by novel biomolecular technologies to address questions in pneumology. Towards this, we use computational methods on biological high-throughput data, so called OMICS data, to gain insights into molecular relationships underlying phenotypic traits and diseases. Focus thereby lies on bioinformatic analyses of next-generation sequencing-based genomics and transcriptomics data. Data, methodology and their applications in pneumology are combined through interdisciplinary collaboration with computer scientists, mathematicians, biologists and physicians. Further, the group harmonizes, administrates and extends a center-wide high-performance computing infrastructure and is currently setting up an accompanying center-wide FAIR (Findable, Accessible, Interoperable, Reusable) OMICS research data management.

Keywords:

group profile
biomolecular data
bioinformatics
lung
research data management

[Link to contribution: Biomolecular Data Science in Pneumology](#)

Genome-based prediction of autoimmune diseases using relation neural networks

Author: Lennart Bartels¹

Co-authors: Marius Jahrens²; Thomas Martinetz²; Hauke Busch²; Inken Wohlers¹

¹ *Forschungszentrum Borstel*

² *Universität zu Lübeck*

Corresponding Author: lbartels@fz-borstel.de

The development of autoimmune diseases arises from a complex interplay of genetic predisposition and environmental influences. Deep learning based approaches could boost the predictive performance by capturing non-linear relationships between genetic variants and the phenotype. In this project, I present and evaluate a deep neural network architecture based on Relation Neural Networks, a network module specifically designed for relational reasoning tasks. The model receives genomic variants as input and represents them with the help of embedding layers. I have performed a successful proof-of concept classification task, predicting population affiliation for individuals from the 1000 Genomes Project. The proposed model is subsequently tested on two disease prediction tasks, predicting rheumatoid arthritis and inflammatory bowel disease for individuals in the UK Biobank.

Keywords:

Deep learning,
Machine learning,
Genetics,
Autoimmune diseases,
Complex diseases

Link to contribution: [Genome-based prediction of autoimmune diseases using relation neural networks](#)

ProteinsPlus: On-The-Fly Structure-Based Design on the Web

Author: Christiane Ehart¹

Co-authors: Thorben Schulze¹; Katrin Schöning-Stierand²; Rainer Fährrolfes¹; Konrad Diedrich¹; Joel Graef¹; Emanuel Ehmki¹; Florian Flachsenberg¹; Jochen Sieg¹; Patrick Penner¹; Jonathan Pletzer-Zelgert¹; Thorben Reim¹; Matthias Rarey¹

¹ *Research Group for Computational Molecular Design (AMD), ZBH - Center for Bioinformatics, Universität Hamburg, Hamburg, Germany*

² *House of Computing and Data Science (HCDS), Universität Hamburg*

Corresponding Authors: matthias.rarey@uni-hamburg.de, christiane.ehart@uni-hamburg.de

The ProteinsPlus web server (<https://proteins.plus>) ¹ offers modelling support for in-depth investigation of biomolecules. Users can perform computational studies for experimental protein structures from the Protein Data Bank (PDB) ² and predicted models from the AlphaFold Protein Structure Database ³.

The services include structure quality analyses, structure preparation, geometric analyses, pocket prediction and characterization, binding site comparison, automated molecular docking, 2D interaction visualization, protein-protein interface classification, and mutation analyses.

In this contribution, we will present the ProteinsPlus web server in a nutshell with implications on their potential application domains in structure-based design.

¹ Schöning-Stierand *et al.* (2020). *Nucleic Acids Res* 48, W48-W53.

² Berman *et al.* (2000). *Nucleic Acids Res* 28, 235-242.

³ Jumper *et al.* (2021). *Nature* 596, 583-589.

Keywords:

ComputationalMolecularDesign

ProteinStructures

DrugDesign

WebServer

DataScience

Link to contribution: [ProteinsPlus: On-The-Fly Structure-Based Design on the Web](#)

Exploring Ultra-Large Chemical Spaces With Genetic Algorithms

Author: Christian Meyenburg¹

Co-authors: Uschi Dolfus ¹; Matthias Rarey ¹

¹ *Universität Hamburg*

Corresponding Authors: christian.meyenburg@uni-hamburg.de, matthias.rarey@uni-hamburg.de, uschi.dolfus@uni-hamburg.de

A major challenge in modern drug design is the vast number of possible molecules that have to be navigated to find a few molecules of interest for a particular project.

Robust search heuristics like genetic algorithms can elevate established methods in the realm of cheminformatics to find this figurative needle in a haystack. Our approach, Galileo, finds promising hit compounds in ultra-large chemical fragment spaces that contain several trillion molecules. These hits both have the desired properties for a drug development project and are most likely synthetically available.

We showcase an application of Galileo in a search for molecules that fulfill a given pharmacophore, i.e., the collection of properties that a molecule needs to possess for a desired biological activity.

Keywords:

Cheminformatics

Drug Design

Evolutionary Algorithm

Molecular Optimization

Chemical Spaces

Link to contribution: [Exploring Ultra-Large Chemical Spaces With Genetic Algorithms](#)

DASHH: Data Science in Hamburg –Helmholtz Graduate School for the Structure of Matter -> Educating the Next Generation of Data Scientists

Authors: Christiane Ehrt¹; Heike Hufnagel Martinez²

¹ *Research Group for Computational Molecular Design (AMD), ZBH - Center for Bioinformatics, Universität Hamburg, Hamburg, Germany*

² *DESY*

Corresponding Authors: heike.hufnagel.martinez@desy.de, christiane.ehrt@uni-hamburg.de

The new generation of analytical instruments produces enormous amounts of data to be analyzed and reconstructed. This development demands innovative methods at the interface of natural sciences and applied mathematics/computer science and a new generation of data scientists well-trained for interdisciplinary research.

At DASHH (<https://www.dashh.org>), doctoral researchers are supervised by one principal investigator from the natural sciences and one principal investigator from mathematics/computer science and benefit from a comprehensive graduate program. Our interdisciplinary data science projects cover the areas Photon Science, Accelerator Science, Particle Physics, Structural Biology, Materials Science and Reproducible Data Science.

Since its foundation in 2019, DASHH has become a cornerstone in the quickly developing Data Science community in Hamburg and beyond.

Keywords:

Data Science,
Natural Sciences,
Informatics,
Mathematics,
Graduate School,
interdisciplinary research

Link to contribution: [DASHH: Data Science in Hamburg –Helmholtz Graduate School for the Structure of Matter -> Educating the Next Generation of Data Scientists](#)

Calculating More Property Distributions of Chemical Fragment Spaces

Author: Justin Lübbers

Co-authors: Louis Bellmann ; Matthias Rarey ; Uta Lessel

Corresponding Authors: bellmann@zbh.uni-hamburg.de, matthias.rarey@uni-hamburg.de, uta.lessel@boehringer-ingelheim.com, justin.luebbers@uni-hamburg.de

In the early stages of drug discovery projects, it is a common task to search digital molecule collections to find promising lead structures that serve as a starting point for developing new drugs. Recently developed combinatorial compound catalogs are orders of magnitude larger than traditional databases, promising higher potential for finding relevant lead structures. Many cheminformatics tasks require new algorithms dealing with these chemical fragment spaces as traditional methods are incompatible. Here, we address the problem of calculating property distributions in the form of histograms for entire catalogs.

For this purpose, Bellmann et al. developed the SpaceProp algorithm. We introduce SpaceProp2, an extension of SpaceProp that supports further molecular properties, namely the TPSA, rotatable bond counts, and SMARTS patterns.

Keywords:

drug discovery
chemical fragment spaces
chemoinformatics

Link to contribution: [Calculating More Property Distributions of Chemical Fragment Spaces](#)

BACI/4C: Next-Generation Multiphysics Simulation for Challenging Real-World Problems

Author: Dirk Steglich

Co-authors: Christian Cyron ; Georg Hammerl ; Hoang-Giang Bui ; Ingo Scheider

Corresponding Authors: ingo.scheider@hereon.de, georg.hammerl@hereon.de, giang.bui@hereon.de, christian.cyron@hereon.de, dirk.steglich@hereon.de

BACI/4C is a massively-parallel multi-physics research code to analyze and solve a plethora of physical real-world problems by means of advanced computational mechanics. BACI/4C provides simulation capabilities for a variety of physical models, including single fields such as solids and structures, fluids, scalar transport, or porous media, and multi-physics coupling and interactions between several physical fields. The capabilities of BACI/4C are sufficient to address problems in aerospace, civil, chemical, or process engineering. Moreover, BACI/4C is also used in medical applications and biophysics. It is soon available as open-source and we expect a fast increasing number of groups worldwide to participate in its development beyond Hereon.

Keywords:

Multiphysics Simulation

Link to contribution: [BACI/4C: Next-Generation Multiphysics Simulation for Challenging Real-World Problems](#)

High-throughput processing for diffraction data

Author: Malte Storm

Co-authors: Christina Krywka ; Peter Staron

Corresponding Authors: christina.krywka@hereon.de, malte.storm@hereon.de, peter.staron@hereon.de

Synchrotron X-ray diffraction (XRD) experiments are a versatile tool in understanding material properties and processes, for example light-weight materials or additive manufacturing and for generating data to create digital twin models.

The analysis of XRD data is often still the domain of experts because software tools were designed for flexibility with numerous parameters which makes them difficult to use for non-experts. Our new software pydidax is a framework designed to open up the user-base of XRD experiments by delivering a user-friendly processing tool capable of handling large datasets of potentially tens of thousands of diffraction images. Processing workflows are modular and based on plugins which can incorporate any computational methodology.

Keywords:

High-throughput processing
Diffraction

Link to contribution: [High-throughput processing for diffraction data](#)

KI4D4E: Artificial Intelligence for Synchrotron-radiation 4D Tomography Data

Author: Tak Ming Wong

Corresponding Author: tak.wong@hereon.de

The Helmholtz-Zentrum Hereon is operating imaging beamlines for X-ray tomography (P05 IBL, P07 HEMS) for academic and industrial users at the synchrotron radiation source PETRA III at DESY in Hamburg, Germany. The high X-ray flux density and coherence of synchrotron radiation enable high-resolution in situ/operando/vivo tomography experiments. Here, large amounts of 4D data are collected from a wide variety of samples, which is challenging to reconstruct, process, and analyze. In this multi-disciplinary project - KI4D4E, we utilize modern machine learning methods for the data processing of synchrotron-radiation tomography experiments, such as segmentation, denoising, multi-modal imaging, phase retrieval, and digital volume correlation, which are applied to the data analysis of biodegradable implant materials.

Keywords:

Artificial Intelligence
Synchrotron-radiation

Link to contribution: [KI4D4E: Artificial Intelligence for Synchrotron-radiation 4D Tomography Data](#)

Exploratory Clinical-Lab Data Analysis

Author: Olaf Deppe¹

¹ *Universitätsklinikum Hamburg-Eppendorf*

Corresponding Author: o.deppe@uke.de

Laboratory test results play a significant role in clinical decisions for individual patients. Analysing these results over large populations and extended periods could offer additional insights. The UKE has collected lab results for 24 years, and hundreds of thousands of patients; however, this wealth of data has yet to be explored. Decomposing the factors that contribute to the observed variance beyond the specifics of the individual case presents a formidable task. Data science techniques are essential in identifying long-term trends and technical discontinuities on top of circadian and seasonal variations so that reference ranges, which allow for tagging cases outside the typical variance, can be determined from the data itself.

Keywords:

Clinical Lab Data
(Partial) Autocorrelation
Reference Limit Estimator

Link to contribution: [Exploratory Clinical-Lab Data Analysis](#)

Entwicklung eines computergestützten Tools zur Bereitstellung von Informationen, Struktur und Entscheidungshilfen für das hausärztliche Management von Multimorbidität (gp-multitool.de)

Authors: Anshu Saini¹; Ingmar Schäfer¹; Julia Nothacker¹; Martin Scherer¹

¹ *Institut für Poliklinik für Allgemeinmedizin. Universitätsklinikum Hamburg Eppendorf*

Corresponding Authors: a.saini@uke.de, m.scherer@uke.de, j.nothacker@uke.de, in.schaefer@uke.de

Hintergrund

Um das hausärztliche Management älterer Patienten mit Multimorbidität zu unterstützen, wurde ein digitales Tool entwickelt, das ermöglicht, behandlungsrelevante Informationen zu erfassen und zu dokumentieren.

Methoden

Funktionalitäten und Anwenderfreundlichkeit des Tools wurden in Fokusgruppen mit Hausärzten und Patienten diskutiert. Das Tool wurde als Webapplikation designt und von einem externen Dienstleister programmiert.

Ergebnisse

Zentrale Elemente des Tools sind standardisierte Untersuchungsbögen. Zugangslinks werden von Hausärzten verschickt und die Untersuchungsbögen werden eigenverantwortlich von Patienten auf elektronischen Endgeräten ausgefüllt. Ergebnisse werden über eine Cloud-Lösung inkl. Verlaufsdarstellung für den Hausarzt verfügbar gemacht.

Diskussion

Studienteilnehmer waren sich weitgehend einig, dass ein digitales Tool auch bei älteren Patienten eine sinnvolle Begleitung der hausärztlichen Versorgung sein kann. Auf die Verständlichkeit der Fragen musste ein besonderer Fokus gelegt werden.

Keywords:

Digitalisierung in der Hausarztpraxis

Multimorbidität

Patientenzentrierung

Versorgungsforschung

Link to contribution: [Entwicklung eines computergestützten Tools zur Bereitstellung von Informationen, Struktur und Entscheidungshilfen für das hausärztliche Management von Multimorbidität \(gp-multitool.de\)](https://www.gp-multitool.de)

The Prompt-a-thon: Designing a Format for Value Co-Creation with Generative AI for Research and Practice

Authors: Constantin von Brackel-Schmidt¹; Emir Kucevic¹; Tilo Böhmann¹

¹ *Universität Hamburg*

Corresponding Authors: constantin.schmidt@uni-hamburg.de, emir.kucevic@uni-hamburg.de, tilo.boehmann@uni-hamburg.de

Generative Artificial Intelligence (GenAI) has immense potential for innovation and problem solving, but there remains a significant knowledge gap in effectively harnessing its capabilities for value creation. To bridge this gap, a Design Science Research project was undertaken to develop a practical tool for researchers and practitioners to explore first insights into the usage of conversational GenAI such as ChatGPT. The result is the Prompt-a-thon (PaT), a human-AI co-creation format tested in five instances. The evaluation showed increasing GenAI knowledge among participants and satisfaction with the PaT. Besides, GenAI input-prompt parameters were identified, serving as potential foundation for prompt-engineering or prompt-strategies. The findings provide a replicable approach for value co-creation that addresses GenAI usage in the rapidly evolving landscape.

Keywords:

artificial intelligence
generative AI
human-AI collaboration
design science research

Link to contribution: [The Prompt-a-thon: Designing a Format for Value Co-Creation with Generative AI for Research and Practice](#)

The Digital Causality Lab - New Ways to Teach Causality

Author: Philipp Bach¹

Co-authors: Gangli Tan ; Martin Spindler

¹ *Universität Hamburg, Fakultät Betriebswirtschaft, Lehrstuhl für Statistik mit Anwendungen in der BWL*

Corresponding Authors: christianana_tan@outlook.com, philipp.bach@uni-hamburg.de, martin.spindler@uni-hamburg.de

The **Digital Causality Lab (DCL)** is an innovative teaching project at the University of Hamburg which focuses on the topics Causal Inference and Data Literacy. It is one of the teaching projects that have been funded in the DDLitLab Project (Stiftung Innovation in der Hochschullehre). In the Digital Causality Lab, we use interactive apps to complement a new MOOC on causal inference, which is open to students from all fields (studium generale). Participants acquire and apply data literacy skills in the context of causal data products (case studies).

Keywords:

Causality
Causal Inference
Data Literacy
Open Source
Teaching

Link to contribution: [The Digital Causality Lab - New Ways to Teach Causality](#)

Predicting COVID-19 Vaccination Uptake from Public Discourse: A Machine Learning Approach

Authors: Esra Eren Bayindir¹; Jonas Schreyögg¹; Martin Spindler¹; Robert Fuchs²

¹ *Universität Hamburg, Hamburg Center for Health Economics*

² *Universität Hamburg, Department of English Language and Literature/Institut für Anglistik und Amerikanistik*

Corresponding Author: esra.eren.bayindir@uni-hamburg.de

This project explores the relationship between public discourse and COVID-19 vaccination uptake and how to use real world data to identify public opinion on COVID-19 vaccination. The analysis will apply machine learning techniques to Twitter data and will link these to data on vaccination rates. In Machine Learning the focus is mainly on providing point forecasts. But in many real-world applications it is also key to address the uncertainty of the estimates provided by the machine learning methods. In recent years so-called conform prediction has been developed to quantify the uncertainty of machine learning methods. Within this project this new methodology will be further developed and applied to the problem at hand.

Keywords:

Machine learning

COVID-19

Twitter

Link to contribution: [Predicting COVID-19 Vaccination Uptake from Public Discourse: A Machine Learning Approach](#)

Causal AI mit DoubleML

Authors: Philipp Bach; Sven Klaassen; Martin Spindler¹

¹ *Institut für Statistik*

Corresponding Authors: martin.spindler@uni-hamburg.de, philipp.bach@uni-hamburg.de, sven.klaassen@uni-hamburg.de

Während KI aktuell hauptsächlich für Vorhersagen verwendet werden, die auf Korrelationen basieren, sind viele Fragestellungen in der Industrie und Forschung kausaler Natur. Das neue Gebiet der Causal AI / Causal ML kombiniert Methoden der kausalen Inferenz mit Methoden des künstlichen Intelligenz. Ein Ansatz dafür ist das sogenannte Double Machine Learning. Es soll dieser Ansatz methodisch vorgeschlagen werden als auch das open source Paket DoubleML, das eins der führenden Frameworks für Causal AI ist und weite Verbreitung in Industrie und Forschung hat.

Keywords:

Machine Learning

AI

Causal AI

Causal ML

Link to contribution: [Causal AI mit DoubleML](#)

In the Pursuit of Eco-Friendly Wind Energy – Unmanned Aerial Systems and Computational Strategies

Author: Tessa T. Taefi¹

Co-authors: Anna Roswag²; Gerald Peklar³; Marc Roswag¹; Matthias Roswag²; Sven Stölzel⁴; Veit Dominik Kunz¹

¹ *Hochschule für Angewandte Wissenschaften Hamburg*

² *Vespertillio Faunistische Untersuchungen*

³ *NXP Semiconductors Hamburg*

⁴ *WindStrom Erneuerbare Energien*

Corresponding Authors: tessa.taefi@haw-hamburg.de, marc.roswag@haw-hamburg.de, veitdominik.kunz@haw-hamburg.de

The project Drones4Bats, funded by the BMWK, employs computational methodologies with autonomous unmanned aerial systems (UAS) to enhance bat-friendly wind turbine systems (WTS) operation and increase wind energy output. It involves developing and comparing various UAS technologies, including Multicopters, specialized bat-friendly UAS, and lighter-than-air UAS. Key computational methods include advanced algorithms for wind turbine blade detection using 77GHz radar, secure ranging and autonomous landing with Ultra-Wideband for recharging, integrated into diverse flight stacks. The project systematically examines UAV impacts on bats, explores economic and ecological potential, and utilizes machine learning algorithms for cloud-based data evaluation. Field-testing shall optimize WTS siting, minimize shutdowns, reduce bat fatalities, and accelerate renewable energy production, addressing the green-green dilemma of climate protection and species conservation.

Keywords:

Sensor Integration

Unmanned Aerial Systems

Cloud-Based Data Evaluation

Eco-Friendly Technology

Link to contribution: [In the Pursuit of Eco-Friendly Wind Energy –Unmanned Aerial Systems and Computational Strategies](#)

Massively Parallel Hybrid Molecular-Continuum Simulations

Author: Louis Viot

Corresponding Author: louis.viot@hsu-hh.de

Micro- and nanofluid flow simulations require considering some effects at the molecular scale. Molecular-continuum coupled flow simulations can perform computationally intensive molecular dynamics (MD) simulations in localized regions of a geometry under consideration, and employ classical, computationally cheaper computational fluid dynamics (CFD) solvers for the remaining larger computational domain. This approach reduces computational efforts tremendously while still considering molecular effects in the flow solution. In this poster, we present our methods used in the context of hybrid MD-CFD simulations, such as massively parallel software concepts, parallel-in-time algorithms or multi-MD instance sampling, and their impact on computational time, precision and energy efficiency in comparison to full-MD simulations.

Keywords:

molecular dynamics, cfd, hpc, multiscale

Link to contribution: [Massively Parallel Hybrid Molecular-Continuum Simulations](#)

marFM®, A Multilingual Automatic Speech Recognizer (ASR) for VHF-Radio Communication

Author: Emin Nakilcioglu¹

Co-authors: Maximilian Reimann ¹; Ole John ¹

¹ *Fraunhofer CML*

Corresponding Authors: maximilian.reimann@cml.fraunhofer.de,
ole.john@cml.fraunhofer.de, emin.nakilcioglu@cml.fraunhofer.de,
sophia.lichtenberg@cml.fraunhofer.de

Radio communication is a central task and safety-critical for international shipping. Despite being based on resilient and robust technology, the quality of transmitted audio signals varies depending on antenna height and prevailing weather conditions. In addition, shipping is characterized by a high degree of crew multi-nationality, so that communications over VHF radio contain a variety of different accents and dialects that further complicate the understanding of radio messages. To address these challenges in maritime communication, we have been developing a transcription solution for marine radio that automatically converts received VHF radio signals into text. We introduce marFM®, a multilingual automatic speech recognizer (ASR) for maritime radio communication.

Keywords:

radio communication, deep learning, asr, maritime

Link to contribution: [marFM®, A Multilingual Automatic Speech Recognizer \(ASR\) for VHF-Radio Communication](#)

Enabling Secure Communication for Cyber-Physical Processes in Critical Infrastructures

Author: Sibylle Fröschle¹

Co-author: Philipp Wolters¹

¹ *Technische Universität Hamburg*

Corresponding Authors: sibylle.froeschle@tuhh.de, philipp.wolters@tuhh.de

Machine-to-machine communication over wireless networks is increasingly adopted to improve service and maintenance processes at airports, ports, and manufacturing plants. This brings with it the challenge of how to bootstrap a secure communication channel between the machines involved. Building on the idea of secure device pairing we research novel schemes for key establishment that exploit the proximity of the machines, the physical presence of a human or robotic operator, and/or the physical characteristics of the process. This approach allows us to engineer for post-quantum resistance as well as resilience against multi-instance attacks. To meet current safety and security norms we also provide the methodology to formally capture and verify the corresponding requirements.

Keywords:

Cyber-physical systems,
Secure communication,
Post-quantum resistance,
Critical infrastructures

Link to contribution: [Enabling Secure Communication for Cyber-Physical Processes in Critical Infrastructures](#)

Presenting a multi-layered manuscript of a Byzantine Greek lexicon within a responsive digital edition

Authors: Alessandro Musino; Christian Brockmann, ; Daniel Deckers; Eva Wöckener-Gade; Lena Hofmann; Louiza Argyriou, ; Stefano Valente

Our poster illustrates how the features of the *Vaticanus Barberinianus graecus 70* (11th c. CE, from Southern Italy) can be reflected in a responsive digital edition currently developed within the *Etymologika* project. This remarkable written artefact's complexity and multi-layered nature put high demands on data modelling and GUI development.

The web interface includes a view of the digitised manuscript, simultaneously illustrating relevant elements of the critical edition on the zoomable manuscript images. Its responsive behaviour helps visualise how the layout corresponds to the process of creation and revision of the manuscript.

The project *Etymologika* is carried out under the auspices of the Academy of Sciences and Humanities in Hamburg with funding from the Academies' Programme.

Keywords:

digital humanities, digital edition, manuscript

Link to contribution: [Presenting a multi-layered manuscript of a Byzantine Greek lexicon within a responsive digital edition](#)

Making Academies' long-term research projects ready for the NFDI

Author: Timm Lehmberg¹

¹ *Akademie der Wissenschaften in Hamburg*

Corresponding Author: tim.m.lehmberg@awhamburg.de

The poster presentation will provide an overview of the digital infrastructure and strategy developed at the Academy of Science and Humanities in Hamburg as part of the activities in the framework of the NFDI cluster *Text+*. To this end, a brief presentation of the six projects funded by the German Academies' long-term program in Hamburg as well as their respective research goals and requirements for digital infrastructure services, is given.

Based on this, an environment is presented that aims to provide services that can be used synergistically by the long-term projects and, at the same time, enable connection to national and supranational research data consortia.

Keywords:

NFDI, digital infrastructure, digital humanities

Link to contribution: [Making Academies' long-term research projects ready for the NFDI](#)

Decentralized Data Mining

Author: Tim Schatto-Eckrodt¹

¹ *Uni Hamburg*

Corresponding Author: tim.schatto-eckrodt@uni-hamburg.de

Obwohl die populären Social Media Plattformen sich teils stark unterscheiden, haben sie eine Eigenschaft gemeinsam: Sie sind zentral organisiert. Im Gegensatz zu diesem zentralisierten Modell steht das dezentralisierte Web (DW). Angebote des DW sollen sich jedoch nicht nur auf technischer Ebene von den etablierten Social Media Plattformen unterscheiden, sondern auch eine Alternative zur Machtkonzentration der zentralisierten Plattformen bieten. Einer der aktuell populärsten DW-Dienste ist die an Twitter angelehnte Social Media Plattform Mastodon. Mit diesem Poster wird eine schematische Darstellung des technischen Aufbaus von Mastodon und den Funktionsweisen und Einschränkungen der Mastodon API für wissenschaftliche Zwecke vermittelt werden. Zusätzlich werden die Ergebnisse einer explorativen Analyse von 1.244 Mastodon-Instanzen und insgesamt 249.251 Nutzer*innen der Plattform vorgestellt.

Keywords:

mastodon
decentralisation
data mining

Link to contribution: [Decentralized Data Mining](#)

Herbie: ELN and research database from Hereon

Author: Catriona Eschke

Co-authors: Fabian Kirchner ; Martin Held ; Milosz Meller

Corresponding Authors: anke-lisa.hoehme@hereon.de, fabian.kirchner@hereon.de, milosz.meller@hereon.de, catriona.eschke@hereon.de, martin.held@hereon.de

Herbie is an Electronic Lab Notebook (ELN) developed at Helmholtz-Zentrum Hereon. Highly structured, yet flexible protocols for parameters and results of individual manufacturing and characterization processes are described with machine-readable ontologies in RDF using OWL and SHACL. By interlinking the protocols and putting them into context with involved staff, samples, equipment, projects, free-text comments, and the actual data files, Herbie enables scientists and technicians to conveniently find all data they need in one place, thus facilitating efficient data evaluation and analysis. Originating from the field of metallic biomaterials, Herbie supports cross-discipline research including but not limited to materials science, engineering, biology, and medicine. Through active participation in various standardization and collaboration initiatives, Herbie can easily connect to other software systems and ELNs.

Keywords:

Electronic Lab Notebook (ELN)
Research Data Management
Ontology
Metallic Biomaterials

Link to contribution: [Herbie: ELN and research database from Hereon](#)

Prediction of therapeutic targets to halt inflammation-induced neurodegeneration.

Authors: Manuel Friese¹; Marcel Seungso Woo¹

¹ *Institut für Neuroimmunologie und Multiple Sklerose (INIMS)*

Corresponding Authors: marcel.seungso-woo@zmnh.uni-hamburg.de,
manuel.friese@zmnh.uni-hamburg.de

Inflammation-induced neurodegeneration poses a clinical challenge in multiple sclerosis and aging. Despite intense research neuroprotective therapies are not available. Applying neuronal networks on “Omics” data holds the promise to identify novel therapies. We aim to establish a platform to find and prioritize genes that determine neurodegeneration. First, we will use generative neuronal networks of scSeq data with genetic perturbations, from different organs and neuronal perturbations to predict neuronal responses. Second, we will add gene-survival dependencies from cell lines to predict modulators of neuronal vulnerability. These genes will be iteratively deleted in neurons to measure viability which will feed further training. Thus, our platform will generalize transcriptomic changes between cell types, predict neuronal viability of genetic alterations, and prioritizes treatment targets.

Keywords:

Neurodegeneration
Multiple Sclerosis
Omics
Treatment targets

Link to contribution: [Prediction of therapeutic targets to halt inflammation-induced neurodegeneration.](#)

CSL Poster Overview

Author: Gesa Ziemer¹

Co-author: ines bigdeli issazadeh²

¹ *HCU*

² *HafenCity University Hamburg*

Corresponding Authors: ines.issazadeh@hcu-hamburg.de, gesa.ziemer@hcu-hamburg.de

This poster shall provide an overview of the history and existing projects within our research institute at HafenCity University Hamburg. Our research projects employ innovative tools and digital city models to visualize and simulate data-driven urban developments, collaborating with a diverse range of stakeholders, including academia, policymakers, civil society, business, non-governmental organizations, and local communities. Our aim is to provide a brief visualisation of our work and how data can be valuable for progressive and participative urban development. We also added two selected projects which focus on climate sustainability.

Keywords:

data, climate, sustainability, digital tools, visualisation

Link to contribution: [CSL Poster Overview](#)

Data analysis at the European XFEL

Author: Ivette Bermudez Macias¹

¹ *European XFEL*

Corresponding Author: ivette.bermudez@xfel.eu

Our poster presents an overview of the activities within the Data Analysis group at European XFEL. These encompass the development of diverse software tools designed to support the steering and interpretation of experimental data during beamtimes, as well as the post-experiment explanation of their outcomes. We introduce **DAMNIT**, a tool engineered to automatically generate experiment overviews, effectively supplanting the labor-intensive task of manual spreadsheet creation. This advancement streamlines the experimental workflow, promoting efficiency, accuracy and reproducibility. Additionally, we show the Python package **EXtra**, a compendium of libraries tailored to assist the analysis of data generated at EuXFEL. Lastly, along the twin transformation initiative, we present the data reductions methods that are already implemented at EuXFEL and its future.

Keywords:

data analysis
data reduction

Link to contribution: [Data analysis at the European XFEL](#)

BORGES: A database for the storage, search and retrieval of ICON data

Authors: Jairo Alonso Segura Bermudez¹; Luis Kornblueh¹; Raphael Schlarb¹; Reinhard Budich¹; Sven Willner¹

¹ *Max-Planck-Institut für Meteorologie*

Corresponding Authors: reinhard.budich@mpimet.mpg.de, sven.willner@mpimet.mpg.de,
raphael.schlarb@mpimet.mpg.de, luis.kornblueh@mpimet.mpg.de,
jairo.segura@mpimet.mpg.de

BORGES tackles the handling of the ever-growing climate model data produced. It is a semantic database for the storage, search and retrieval of model data. It serves mainly demands of the up-to-date, competitive Earth system model ICON, developed and used at the MPI-M, Hamburg. BORGES excels for both large ensembles of model experiments as well as very high-res, large volume, „storm-resolving“ model configurations. In a classical HPC context, the system retrieves large data streams and the associated metadata from the model experiments using semantic data management.

For storing the data stream, BORGES builds on the domain-specific object storage Field Database, developed by the European Centre for Medium-Range Weather Forecast. This ensures metadata storage consistent with stream data in classical databases.

Keywords:

Semantic data management
HPC
ICON
fdb

Link to contribution: [BORGES: A database for the storage, search and retrieval of ICON data](#)

The VoicesCarry App: Co-Designing Worker Engagement Technology with Supply Chain Workers

Authors: Emma van den Terrell¹; Laura Edinger-Schons²

¹ *University of Mannheim*

² *University of Hamburg*

Corresponding Authors: evadent@mail.uni-mannheim.de, laura.marie.edinger-schons@uni-hamburg.de

The Problem: businesses are digitally collecting information on workers' wellbeing and labor condition in their supply chains, but little is known about the challenges that emerge from digitally-assisted auditing technology. Our poster describes the development of the VoicesCarry app, a web-based survey application that focuses on collecting self-reported information on supply chain workers' wellbeing at work. The poster describes the app's participatory co-creation that emerged from a multi-stakeholder project group (academic teams from computer science and business ethics, business representatives such as Vaude and factory management, representatives of civil society, and garment industry employees). The poster will report on the qualitative results from the app's development process that navigates the social side of sustainability and digital technology development and implementation.

Keywords:

Worker Wellbeing
App Development
Co-creation
Participatory Design

Link to contribution: [The VoicesCarry App: Co-Designing Worker Engagement Technology with Supply Chain Workers](#)

ML based process monitoring of semi-automatic drilling processes

Author: Malte Flehmke¹

¹ *Institut für Produktionsmanagement und -technik*

Corresponding Author: malte.flehmke@tuhh.de

Since one third of rivet holes during aircraft assembly are produced with semi-automatic drilling units, in this work reliable and efficient methods for process state prediction using Machine Learning (ML) classification methods were developed for this application. Process states were holistically varied in the experiments, gathering motor current and machine vibration data. These data were used as input to identify the optimal combination of five data feature preparation and nine ML methods for process state prediction. K-nearest-neighbour, decision tree and artificial neural network models provided reliable predictions of the process states: workpiece material, rotational speed, feed, peck-feed amplitude and lubrication state. Data preprocessing through sequential feature selection and principal components analysis proved to be favourably for these applications.

Keywords:

Process Monitoring, Machine Learning, Drilling, Aircraft Assembly

Link to contribution: [ML based process monitoring of semi-automatic drilling processes](#)

PUNCH4NFDI - Facilitating FAIR data in Germany

Authors: Christiane Schneide¹; Thomas Schörner¹

¹ *DESY*

Corresponding Authors: christiane.schneide@desy.de, thomas.schoerner@desy.de

PUNCH4NFDI (the NFDI consortium of particle, astro-, astroparticle, hadron and nuclear physics) has as its central deliverable a community-overarching science data platform (SDP), in which complex workflows can be executed on digital research products (DRPs) in a transparent, automatised and FAIR (Findable, Accessible, Interoperable, Reusable) way. The SDP consists of several ingredients, not least the Compute4PUNCH and Storage4PUNCH federated infrastructures, that are more or less far advanced in their development and the interplay of which is now being implemented.

Furthermore, PUNCH4NFDI is working on coherent metadata for PUNCH sciences, and providing training in research data management.

In this poster, we will give an overview of ongoing and planned activities by showcasing two example workflows to be executed on the SDP.

Keywords:

PUNCH4NFDI
Research Data Management
FAIR principles
physics

Link to contribution: [PUNCH4NFDI - Facilitating FAIR data in Germany](#)

MRI Total: Towards an automated, standardized, reproducible, version-controlled, transparent, FAIR, quality-assured, privacy-compliant processing pipeline for MRI data

Author: Lennart Wittkuhn¹

¹ *Institute of Psychology, University of Hamburg*

Corresponding Author: lennart.wittkuhn@uni-hamburg.de

Neuroscientists at University of Hamburg (UHH) collect thousands of hours of human brain imaging data per year. However, this data treasure is not systematically standardized, stored, or controlled for quality. We propose an institution-wide processing pipeline for human brain imaging data that connects state-of-the-art open-source neuroinformatics software tools and leverages UHH computational infrastructure, including the object storage and high-performance computing cluster. Immediately following collection, data are transformed to the Brain Imaging Data Structure (BIDS) standard and submitted to in-depth automated quality assurance (MRIQC). In addition, an interactive dashboard allows browsing metadata of available data sets and MRI quality metrics, fostering transparency and scientific exchange. This project promises high-quality reproducible research outputs and boost the efficiency of scientific workflows.

Keywords:

neuroimaging
reproducibility
open data
quality assurance
FAIR

Link to contribution: [MRI Total: Towards an automated, standardized, reproducible, version-controlled, transparent, FAIR, quality-assured, privacy-compliant processing pipeline for MRI data](#)

Helmholtz Imaging@DESY

Authors: David Schwartz¹; Martin Burger²

Co-author: Tim Roith ²

¹ *DESY/IT*

² *DESY, FS-CI and Helmholtz Imaging*

Corresponding Authors: tim.roith@desy.de, martin.burger@desy.de,
david.schwartz@desy.de

This poster will provide an overview of the Helmholtz Imaging Platform at DESY and its major activities.

Helmholtz Imaging involves both service for practitioners on different tasks in digital imaging as well as cutting edge research. The group at DESY focuses at the early part of imaging, i.e. the image formation part and related tasks such as image reconstruction, denoising, motion estimation, and registration.

Keywords:

Image Reconstruction, Image Analysis, AI in Imaging

Link to contribution: [Helmholtz Imaging@DESY](#)

Computational Imaging@DESY

Author: Lukas Weigand¹

Co-authors: Lorenz Kuger ¹; Martin Burger ¹; Samira Kabri ¹; Tim Roith ¹

¹ *DESY, FS-CI and Helmholtz Imaging*

Corresponding Authors: lukas.weigand@desy.de, martin.burger@desy.de, samira.kabri@desy.de, tim.roith@desy.de, lorenz.kuger@desy

We will present the work and expertise of the newly established Computational Imaging Group at DESY. Being founded in mathematics, we provide expertise in inverse problems as well as theoretical foundations of machine learning methods.

We give an overview of mathematical methods in imaging and applications thereof, in particular algorithms for reducing computational cost in large scale problems. Moreover, we give insight into theoretically founded training of robust or sparse neural networks.

Keywords:

Inverse problems, image reconstruction, machine learning, neural networks, robust training.

Link to contribution: [Computational Imaging@DESY](#)

Using computational methods to study niche audiences

Author: Judith Möller¹

Co-author: Ellen Linnert

¹ *Universität Hamburg & Leibniz Institut für Medienforschung (HBI)*

Corresponding Authors: judith.woude@uni-hamburg.de, ellen.linnert@student.uva.nl

This study employs computational methods, utilizing YouTube user data donations and survey data, to investigate the effects of fringe bubbles on social media. These niche communities, often situated at the margins of the public sphere, are shaped by algorithmically curated biased content, potentially distorting users' perceptions of public discourse and amplifying non-mainstream voices.

Through supervised machine learning and large language models, we automatically classify the video data and identify potential fringe bubbles. We analyze the effects of exposure to fringe issues on users' perceptions of the public sphere, participatory behavior, and democratic values. The findings contribute to understanding the role of AI-based systems in the development and societal impact of fringe movements, such as the anti-vaccination movement.

Keywords:

computational methods

media effects

filter bubbles

Link to contribution: [Using computational methods to study niche audiences](#)

INEL –Grammar, Corpora and Language Technology for Indigenous Northern Eurasian Languages

Authors: Elena Lazarenko¹; Aleksandr Riaposov¹

¹ *Universität Hamburg*

Corresponding Authors: elena.lazarenko@uni-hamburg.de, aleksandr.riaposov@uni-hamburg.de

Ever since 2016, the spanning 18 years long-term project INEL has been generating deeply annotated language corpora and accompanying digital resources using existing and acquiring new language material from a number of heavily endangered languages of the Northern Eurasian Area. The core aim is not only to make these resources sustainably available after their publication but also provide continuous data curation and analysis already during their preparation. This puts high demands on the digital workflows that involve various data formats, tools, and approaches at all levels from data preprocessing to the publication of the corpus. The tools range from well-established software for linguistic work to frameworks and tools developed and maintained specifically for the project needs.

Keywords:

language corpora,
workflow management,
quality assurance,
data sustainability

Link to contribution: [INEL –Grammar, Corpora and Language Technology for Indigenous Northern Eurasian Languages](#)

Climate Data Operators - CDO

Author: Uwe Schulzweida¹

¹ *Max-Planck-Institut für Meteorologie*

Corresponding Author: uwe.schulzweida@mpimet.mpg.de

The Climate Data Operators (CDO) software is a collection of many operators for processing of climate and weather model data. It includes simple statistical and arithmetic, data selection, interpolation functions and many other.

The extremely large amount of data produced by high-resolution climate simulations make such problems essentially I/O bound. CDO is optimised to cope with this.

Computationally intensive operators are OpenMP parallelised and prepared to use GPUs if necessary. Many operators can be combined arbitrarily by operator chaining, providing another parallel layer. This is realised using POSIX threads.

Due to its ease of use, reliability, constant evolution and high efficiency, CDO has become a very popular tool in the climate community over the last 20 years.

Keywords:

Climate Data Operators
High performance computing
Data processing
GRIB
NetCDF

[Link to contribution: Climate Data Operators - CDO](#)

Motion Editing tool for Reproducing Grammatical Elements of Japanese Sign Language Avatar Animation

Author: Tsubasa Uchida¹

Co-authors: Hiroyuki Kaneko ²; Masanori Sano ²; Naoki Nakatani ²; Taro Miyazaki ²

¹ *NHK Science & Technology Research Laboratories, Universität Hamburg, Institut für Deutsche Gebärdensprache und Kommunikation Gehörloser*

² *NHK Science & Technology Research Laboratories*

Corresponding Authors: sano.m-fo@nhk.or.jp, miyazaki.t-jw@nhk.or.jp, nakatani.n-ge@nhk.or.jp, kaneko.h-dk@nhk.or.jp, tsubasa.uchida@uni-hamburg.de

Avatar animation is one of the ways to produce sign language content. Using motion capture data allows for fluent avatar movements closer to human signer. However, the generated animation usually cannot adjust grammatical features such as intonation, signing space, and facial expression to match a specific context. This is because all motion data is in a fixed form at the time it is captured.

Therefore, we propose a motion editing tool that can reproduce grammatical elements of Japanese Sign Language (JSL) by editing each motion data. Evaluation experiment shows that editing the speed and blend span of multiple words corresponding to the delimitation of phrases and clauses can reduce the error rate of understanding JSL avatar animations.

Keywords:

sign language
avatar animation
motion capture
motion editing

Link to contribution: [Motion Editing tool for Reproducing Grammatical Elements of Japanese Sign Language Avatar Animation](#)

Interdisciplinary and International Collaborations: The MARS Research Group at Hamburg University of Applied Sciences (HAW)

Author: Thomas Clemen¹

Co-authors: Daniel Glake²; Helen Haase³; Ulfia Clemen⁴

¹ *Hamburg Univ. of Applied Sciences*

² *IT Anker GmbH, Hamburg*

³ *HAW Hamburg*

⁴ *Logistik-Initiative Hamburg*

Corresponding Authors: helen.haase@haw-hamburg.de, thomas.clemen@haw-hamburg.de, uc@hamburg-logistik.net, daniel.glake@itanker.de

The MARS Research Group at Hamburg University of Applied Sciences fosters interdisciplinary and international collaboration. Our projects address diverse demands, from city planning and logistics to global change mitigation. We have developed an extensive framework for large-scale agent-based simulations, extensively used in both our research and global university courses. Our current work explores how swarms of robots and drones can learn complex tasks, such as car assembly, from human interaction. We also demonstrate using multi-agent systems, IoT sensors, and spatiotemporal databases to create a Digital Twin of multi-modal mobility in Hamburg. This poster showcases the MARS Group's commitment to addressing multifaceted challenges through technology and interdisciplinary teamwork, promoting innovative solutions for real-world problems.

Keywords:

multi-agent systems
collaborative learning
spatio-temporal data
digital twin

Link to contribution: [Interdisciplinary and International Collaborations: The MARS Research Group at Hamburg University of Applied Sciences \(HAW\)](#)

Conceptual foundation of machine learning in Engineering using example of tailored laser pulse forming and Optical Mirror Design

Author: Utsa Chattopadhyay¹

¹ *TUHH/DESY*

Corresponding Author: utsa.chattopadhyay@tuhh.de

Next generation plasma accelerators require ultrashort laser pulses with tailored pulse characteristics, provided at high-peak and average power. Advancing the recent method of versatile tuning of wavelength of Ultrashort pulses further promises to open entirely new parameter regimes in particular for high-power ultrafast lasers. The approach relies on nonlinear light-matter interaction of temporally shaped laser pulses. The underlying processes can be optimized by a multitude of control parameters including pulse energy, system size, nonlinear medium characteristics, etc. In particular, pulse-shape characteristics and dispersion properties of the nonlinear system open key control parameter spaces determining the output pulse characteristics. It employs a cavity-like arrangement in which state-of-the-art dispersion-engineered multi-layer-mirror technology can be used as additional tuning knobs, enabling multi-dimensional dispersion tuning for pulse shaping optimization.

Keywords:

Ultrashort Laser
Optimization
Optical Mirror Design
Pulse Shaping

Link to contribution: [Conceptual foundation of machine learning in Engineering using example of tailored laser pulse forming and Optical Mirror Design](#)

Towards the Automatic Generation of Models for Prediction, Monitoring, and Testing of Cyber-Physical Systems

Authors: Eric MSP Veith¹; Görschwin Fey²; Jakob Schyga²; Jan Christian Wieck³; Jochen Kreutzfeldt²; Johannes Hinckeldeyn⁴; Julian Kohlisch¹; Markus Knitt²; Maximilian Schmidt²; Stephan Balduin¹; Swantje Plambeck²

¹ *OFFIS - Institute for Information Technology*

² *TUHH*

³ *Fraunhofer –Center for Maritime Logistics and Services CML*

⁴ *STILL GmbH*

Corresponding Authors: eric.veith@offis.de, jochen.kreutzfeldt@tuhh.de, johannes.hinckeldeyn@kiongroup.com, julian.kohlisch@offis.de, stephan.balduin@offis.de, swantje.plambeck@tuhh.de, goerschwin.fey@tuhh.de, jakob.schyga@tuhh.de, maximilian.schmidt@tuhh.de, markus.knitt@tuhh.de, jan.christian.wieck@cml.fraunhofer.de

Modeling Cyber-Physical Systems (CPS) is a challenging task as they are composed of heterogeneous components that interact with each other and with the physical environment. This inherent complexity requires a high level of knowledge about the system and its environment when modeling CPS. We develop a framework for the automatic generation of models for CPS, which integrate into tools for prediction, monitoring, and testing. This framework employs machine learning techniques to learn models from data or simulation. To learn models of arbitrary application areas and different purposes, we use a modular approach that allows to support various learning techniques. This approach reduces the effort for designing, testing, and maintaining CPS in both research and industry settings.

Keywords:

modeling
prediction
monitoring
testing
CPS

Link to contribution: [Towards the Automatic Generation of Models for Prediction, Monitoring, and Testing of Cyber-Physical Systems](#)

Digitisation Services @ SUB HH

Authors: Kerstin Wendt¹; Michael Kubina¹

¹ *Staats- und Universitätsbibliothek Hamburg*

Corresponding Authors: michael.kubina@sub.uni-hamburg.de,
kerstin.wendt@sub.uni-hamburg.de

Over the past 10 years, the Hamburg State and University Library has built up a comprehensive technical and organisational infrastructure for the digitisation of 2D objects for digital work with texts and images of a wide variety of materials (newspapers, journals, manuscripts, estate material, old prints, maps, copperplate engravings, etc.). We also support other institutions in Hamburg in the systematic digitisation of their cultural heritage holdings. As a result, high-quality master TIFFs (8.1 million pages) are available for research, most of them as full texts (5.1 million pages). In the case of the library's own holdings, the content focus of the digitisation was on the relevant historical newspapers up to 1945.

Keywords:

digitisation
cultural heritage
research data
OCR
OLR

Link to contribution: [Digitisation Services @ SUB HH](#)

ahoi.digital - Die Allianz der Hamburger Hochschulen für Informatik

Author: Janis-Marie Paul¹

Co-authors: Jörg Rainer Noennig²; Martin Semmann³; Nihat Ay ; Tilo Böhmann⁴; Ulrike Steffens⁵

¹ *ahoi.digital*

² *HafenCity Universität Hamburg*

³ *House of Computing and Data Science, Universität Hamburg*

⁴ *Universität Hamburg*

⁵ *HAW Hamburg*

Corresponding Author: janis-marie.paul@uni-hamburg.de

ahoi.digital is a program conceived based on the specific recommendations of the Wissenschaftsrat (WR) following their 2016 evaluation of computer science departments in Hamburg. The WR advised that Hamburg's higher education institutions should establish a closer, institutionalized collaboration among their computer science divisions through a cooperative platform. ahoi.digital was hence established to promote cooperation and coordination in computer science in Hamburg, ensuring national and international competitiveness. By creating a cross-university coordination committee, it laid the foundation for harmonizing and focusing on computer science efforts. ahoi.digital consolidates the computer science departments of HafenCity University (HCU), Hochschule für Angewandte Wissenschaften Hamburg (HAW), Technische Universität Hamburg (TUHH), and Universität Hamburg (UHH). This aggregation has established a framework to strategically shape future directions and maximize identified potentials. ahoi.digital focuses on existing priority areas like Cognitive Systems, Smart & Cyber-Physical Systems, and Information Governance Technologies, with the addition of Data Science based on the WR's recommendation. Additionally, the graduate school "sharing.city" promotes cross-institutional co-supervision of dissertations, primarily to foster networking among new colleagues in the region.

Keywords:

computer science
collaboration

Link to contribution: [ahoi.digital - Die Allianz der Hamburger Hochschulen für Informatik](#)

Sind wir interessant für KI –oder: ist KI interessant für uns? Analysen des Textkorpus „Mehrsprachigkeitsentwicklung im Zeitverlauf (MEZ)“

Authors: Ingrid Gogolin; Nora Dünkel; Thorsten Klinger; Kseniia Pershina; Anouk Ticheloven; Birger Schnoor; Irina Usinova

Corresponding Author: gogolin@uni-hamburg.de

Die MEZ-Studie untersucht die Entwicklung mehrsprachiger Literalität bei n=2103 Schüler(inne)n der Sekundarstufe (Anfangsstichprobe) aus deutsch-russischen, deutsch-türkischen oder einsprachig deutschen Familien. Sie lernten Englisch als erste, Französisch oder Russisch als zweite Fremdsprachen. Schriftsprachliche Fähigkeiten wurden in vier Messzeitpunkten getestet. Bewertet wurden die Fähigkeiten in Deutsch, Russisch, Türkisch, Englisch sowie den zweiten Fremdsprachen. Das Korpus enthält 16.859 handgeschriebene Texte als Bilddateien und Transkriptionen, die die handschriftliche Schreibweise –also originelle Orthographie –genau wiedergeben.

Die Auswertung des komplexen Korpus ist sehr aufwendig. Das Verfahren wäre aber weltweit einsetzbar, da es den Qualitätsanforderungen an Diagnostik einer mehrsprachigen Schülerschaft standhält. Wir nutzen ein theoretisches Modell der Schreibentwicklung, das für die verschiedenen Sprachen identisch ist, also vergleichende Aussagen erlaubt. Eine interessante Anwendungsmöglichkeit für KI?

Keywords:

text corpus analysis, multilingualism, educational science

Link to contribution: [Sind wir interessant für KI –oder: ist KI interessant für uns? Analysen des Textkorpus „Mehrsprachigkeitsentwicklung im Zeitverlauf \(MEZ\)“](#)

Spoki: A Reactive and Scalable Network Telescope

Author: Raphael Hiesgen¹

¹ *HAW Hamburg*

Corresponding Author: raphael.hiesgen@haw-hamburg.de

Internet-wide scans are cheaply and quickly performed in IPv4. They are not only used to analyze the Internet ecosystem but abused to find vulnerable systems. We developed Spoki, a reactive-network telescope built on top of native actors in C++. It accepts TCP connections and collects payloads to look beyond the source addresses and get deeper insight into scanners.

Spoki is deployed at four prefixes in two regions, which helps us to study topological and regional differences. Designed as a long-term project, Spoki has collected TBs of data. This large-scale collection allows us to analyze unforeseen events, such as the Log4Shell incident. Clustering Log4Shell scanners by their infrastructure revealed a large-scale campaign responsible for more than 50% of events in 2022.

Keywords:

Internet Measurement
Scanners
Scalable Systems
Data Analysis
Security

Link to contribution: [Spoki: A Reactive and Scalable Network Telescope](#)

bAIome Center for Biomedical AI- Towards Intelligent Health

Authors: Anna Reinicke-Vogt¹; Claus Christian Hilgetag¹; Frank Ückert¹; Jan Erik Gewehr¹; Lorenz Adlung¹; Marina Zimmermann¹; Rene Werner¹; Stefan Bonn¹; Stefano Panzeri¹; Vadim Ustinov¹

¹ *Universitätsklinikum Hamburg-Eppendorf*

bAIome is the center for biomedical AI at University Medical Center Hamburg-Eppendorf (UKE). It consists of faculty members and staff from different institutes and departments within UKE engaged in research and education in AI relevant to biomedicine. bAIome promotes basic and applied AI research that has the possibility to translate into innovative solutions that will integrate into clinical practice.

Keywords:

biomedicine, machine learning, AI, Omics, Image processing, EHR

[Link to contribution: bAIome Center for Biomedical AI- Towards Intelligent Health](#)

ExaOcean: Improving Performance of the ICON-O Oceanmodel on heterogeneous Exascale-Supercomputers with Machine Learning

Author: Daniel Ruprecht¹

Co-authors: Christopher Kadow ²; Lars Hoffmann ³; Peter Korn

¹ *Technische Universität Hamburg*

² *German Climate Computing Centre*

³ *Jülich Supercomputing Centre*

Corresponding Authors: kadow@dkrz.de, daniel.ruprecht@tuhh.de, peter.korn@mpimet.mpg.de, l.hoffmann@fz-juelich.de

Ocean models are a key component of every weather or climate model. Despite the computing power of modern supercomputers, however, important dynamical features can so only be resolved in simulations over a few weeks.

ExaOcean will deliver modern mathematical algorithms to achieve better parallel scaling and faster runtimes in highly resolved simulations on new supercomputers. We will integrate techniques from machine learning into mesh based algorithms, using data from highly resolved short term simulations to train a correction term for long term simulations we lower mesh resolution. This will enable “effectively sub-mesoscale resolving simulations”, where the effect of the sub-mesoscale vortices on the larger scale dynamics is represented via the ML correction term.

Keywords:

ocean modeling
super-resolution
high-performance computing
machine learning
mesh-based methods

Link to contribution: [ExaOcean: Improving Performance of the ICON-O Oceanmodel on heterogeneous Exascale-Supercomputers with Machine Learning](#)

Combining numerical methods and machine learning for physical and engineering sciences

Author: Sebastian Götschel¹

Co-author: Daniel Ruprecht²

¹ *TU Hamburg*

² *Technische Universität Hamburg*

Corresponding Authors: daniel.ruprecht@tuhh.de, sebastian.goetschel@tuhh.de

Many areas in the physical or engineering sciences rely on computational models to some extent. These models can be based on fundamental physics processes, typically leading to a set of differential equations. Alternatively, machine learning techniques can be used to infer input-output relations out of very large sets of data. Both approaches come with different strengths and weaknesses but rely on mathematical algorithms to function reliably and efficiently. In the last years, we are also increasingly seeing synergies between both worlds, e.g., when ML is used as part of a numerical algorithm for solving differential equations of a physics-based model. Our poster will present case studies for combining ML and numerical methods, with applications including medical imaging and high-performance computing.

Keywords:

physics-based modeling
machine learning
high-performance computing
numerical methods
medical imaging

Link to contribution: [Combining numerical methods and machine learning for physical and engineering sciences](#)

From Super-Resolution to Downscaling - An Image-Inpainting Deep Neural Network for High Resolution Weather and Climate Models

Authors: Christopher Kadow; Maximilian Witte

Co-authors: Danai Filippou ; David Hall ; Johannes Meuer ; Étienne Plésiat

Corresponding Authors: meuer@dkrz.de, kadow@dkrz.de, filippou@dkrz.de, witte@dkrz.de, plesiat@dkrz.de, dhall@nvidia.com

High resolution in weather and climate was always a common and ongoing goal of the community. In this regards, machine learning techniques accompanied numerical and statistical methods in recent years. Here we demonstrate that artificial intelligence can skilfully downscale low resolution climate model data when combined with numerical climate model data. We show that recently developed image inpainting technique perform accurate super-resolution via transfer learning using the HighResMIP of CMIP6 (Coupled Model Intercomparison Project Phase 6) experiments. Its huge data base offers a unique training opportunity for machine learning approaches. The transfer learning purpose allows also to downscale other CMIP6 experiments and models, as well as observational data like HadCRUT5.

Keywords:

Artificial Intelligence, Super Resolution, Inpainting, Climate Modeling, Observations

Link to contribution: [From Super-Resolution to Downscaling - An Image-Inpainting Deep Neural Network for High Resolution Weather and Climate Models](#)

DDLitLab

Authors: Carolin Scharfenberg¹; Julia Niemann-Lenz¹; Kai-Uwe Schnapp¹; Katharina Kleinen-von Königslöw¹

¹ *Universität Hamburg*

Corresponding Authors: katharina.kleinen@uni-hamburg.de, kai-uwe.schnapp@uni-hamburg.de, carolin.scharfenberg@uni-hamburg.de, julia.niemann-lenz@uni-hamburg.de

Das von der Stiftung Innovationen in der Hochschullehre geförderte Projekt „Digital and Data Literacy in Teaching Lab“ (DDLitLab) fördert digitale Lehrinnovationen mit Bezug zur „Data Literacy Education“ durch Lehrprojekte an allen Fakultäten und im fakultätsübergreifenden Studium Generale. DDLitLab verfolgt zwei Ziele:

1. Die Weiterentwicklung der „Digital University Teaching Literacy“ (DUTy) der Lehrenden. Wir wollen gemeinsam mit dem Hamburger Zentrum für Universitäres Lehren und Lernen Strukturen weiterentwickeln und schaffen, in denen Lehrende ihre digital-didaktischen Kompetenzen ausbauen können, um digitale und hybride Lehrformate noch souveräner zu beherrschen und zu gestalten.
2. Die Förderung der „Data Literacy“ von Studierenden. Wir wollen Studierende im Rahmen von fachspezifischen, fachübergreifenden und transferorientierten Lehrveranstaltungen zu einem kritischen Verständnis von Datenpraktiken und einem mündigen Umgang mit Daten befähigen.

Keywords:

Digital Literacy
Data Literacy
Lehrinnovationen
Datenprojekte

Link to contribution: [DDLitLab](#)

Does Journalism Reflect Polarized Online Networks? A Comparative Analysis of the Discourse on the "Last Generation" Climate Protest Movement in (Social) Media

Authors: Hendrik Meyer; Louisa Pröschel; Michael Brüggemann

Corresponding Authors: hendrik.meyer-1@uni-hamburg.de, louis-
aproschel@gmail.com, michael.brueggemann@uni-hamburg.de

While science emphasizes the urgency of addressing the climate crisis, political actions to combat climate change lag. Numerous climate protest movements have emerged in recent years. However, "The Last Generation, " stood out in recent German media coverage.

Our research focuses on the discourse around this climate protest group, which gained significant attention due to their disruptive, potentially polarizing protest methods.

Our project analyzes German Twitter discourse (1, 400, 000 tweets) and media coverage (7, 000 articles) about the movement. We employ automated network analyses, manual codings, and automated methods such as transformer-based classifiers. Preliminary findings reveal polarization of community structures and content in (online) media, with extreme frames being propagated by right-wing actors and media like AfD politicians, the Bild, or Junge Freiheit.

Keywords:

network analysis
social media
climate protest
content analysis
journalism

Link to contribution: [Does Journalism Reflect Polarized Online Networks? A Comparative Analysis of the Discourse on the "Last Generation" Climate Protest Movement in \(Social\) Media](#)

When Recommender Systems Select Political News

Authors: Hendrik Meyer¹; Jessica Kunert²; Juliane Lischka¹; Katharina Kleinen-von Königslöw¹; Laura Laugwitz¹; Lucien Heitz³; Nadja Schaezt¹; Rana Abdullah⁴

¹ *Journalistik und Kommunikationswissenschaft, Universität Hamburg*

² *Universität Mainz*

³ *Universität Zürich*

⁴ *Universität Hamburg*

Corresponding Authors: laura.laugwitz@uni-hamburg.de, nadja.schaetz@uni-hamburg.de, juliane.lischka@uni-hamburg.de, katharina.kleinen@uni-hamburg.de

News recommender systems (NRS) determine news exposure for digital publics. However, the effects of NRSs on users remain understudied. This study examines how diversity of political news within NRSs can influence political news use, knowledge, and attitudes. We implemented an NRS that gathers news from a variety of news outlets, identifies political party representations in the articles, and recommends them based on a diversity model in a smartphone app. Results are based on a field experiment conducted during the state elections in Lower Saxony, Germany. While NRSs can increase knowledge about minority parties, they do not change voting behavior. These findings underscore the complex nature of NRSs in the political context and the need for interdisciplinary collaboration in their design.

Keywords:

Computational methods

Journalism

News recommender systems

Voting behavior

Link to contribution: [When Recommender Systems Select Political News](#)

Federated Machine Learning for Big Data

Author: Stefan Schulte¹

¹ *Technische Universität Hamburg*

Corresponding Author: stefan.schulte@tuhh.de

On this poster, we are going to present the research of the Institute for Data Engineering at TU Hamburg on the processing of very large data streams as well Federated Learning (FL). Federated Learning (FL) offers an alternative approach to centralized Machine Learning, by distributing the model generation across different entities. Thus, learning can be conducted close to the data sources, and only the learned model is shared with other entities. This leads to benefits both with regard to data privacy and communication overhead. In addition, we will present our work on Data Stream Processing, i.e., how to handle very large amounts of streaming data coming, e.g., from Internet of Things devices, in a resource-efficient way.

Keywords:

Big Data
Machine Learning
Data Stream Processing
Federated Learning

Link to contribution: [Federated Machine Learning for Big Data](#)

Beta maṣāḥəft: Die webbasierte Kollaborationsplattform für Handschriftenforschung

Author: Alessandro Bausi¹

Co-authors: Carsten Hoffmann¹; Denis Nosnitsin¹; Dorothea Reule¹; Eugenia Sokolinski¹

¹ *Universität Hamburg*

Corresponding Authors: eugenia.sokolinski@uni-hamburg.de, dorothea.reule@uni-hamburg.de, alessandro.bausi@uni-hamburg.de

Das Langzeitvorhaben der Akademie der Wissenschaften in Hamburg „Die Schriftkultur des christlichen Äthiopiens und Eritreas“(2016-2040) entwickelt eine virtuelle Forschungsumgebung, in der detaillierte Beschreibungen von Handschriften mit Informationen zu Literaturwerken, Personen (Schreibern, Besitzern und Autoren) und Orten verknüpft werden (<https://betamasahft.eu/>). Die computergestützte Analyse wird neue Fragestellungen zu Handschriftenkunde, Geschichte, historischer Geographie, Linguistik und Literaturwissenschaft ermöglichen.

Die für das Projekt notwendigen Daten können dank einem Git-basierten Workflow gemeinsam erstellt, kontrolliert und revidiert werden, wobei die Änderungen sofort auf das Produktionsserver übertragen werden. Auf diese Weise kann das Projekt von den Kompetenzen von vielen Kollegen weltweit leben und die Ergebnisse in Realzeit mit der Forschungsgemeinschaft teilen.

Keywords:

GitHub

TEI XML

XSLT

Collaborative Research

Link to contribution: [Beta maṣāḥəft: Die webbasierte Kollaborationsplattform für Handschriftenforschung](#)

Leveraging Jupyter on Maxwell HPC: joyful, visual and green computing

Author: Frank Schlünzen¹

Co-authors: Arlena Mills-Marzoli ¹; Axel Wichmann ¹; Neele Rahmlow ¹; Sven Sternberger ¹; Yves Kemp ¹

¹ *DESY*

Corresponding Authors: neele.rahmlow@desy.de, gesa.arlena.mills-marzoli@desy.de, sven.sternberger@desy.de, frank.schlunzen@desy.de, axel.wichmann@desy.de, yves.kemp@desy.de

Jupyter notebooks are great tools to mitigate the complexities of (heterogeneous) HPC systems - like the Maxwell cluster at DESY which serves the computational needs of all user facilities on campus, as well as a wide variety of applications ranging from plasma accelerators to quantum chemistry. We aim to expand the Jupyter ecosystem utilizing python application frameworks to provide application environments with real-time visualization capabilities and tailored to the needs of (less experienced) users. On this basis we are implementing for example Jupyter-driven remote desktops, user-friendly dashboards to compose or monitor batch-jobs, and visual frontends for data catalogues like SciCat. The Jupyter extensions are accompanied by visual tools for resource utilization and CO2 footprints suitable for both users and administrators.

Keywords:

HPC
Jupyter

Link to contribution: [Leveraging Jupyter on Maxwell HPC: joyful, visual and green computing](#)

Digital Scholarship Services at State- and University Library Hamburg

Author: Jonas Müller-Laackman¹

¹ *Staats- und Universitätsbibliothek Hamburg*

Corresponding Author: jonas.mueller-laackman@sub.uni-hamburg.de

The Department of Digital Scholarship Services of the State and University Library Hamburg is a partner for the conception, planning and implementation of projects in the field of digital humanities. We offer advice and support for project proposals and research projects, and act as a communication interface between the library infrastructure and Hamburg's academic institutions. We work closely with the departments of the SUB, Hamburg University and other academic stakeholders, such as the HCDS. From the end of 2023, we will co-run the Crossdisciplinary Lab for Digital Humanities in the Humanities Library of Hamburg University, which will become a place of exchange and collaboration in the digital humanities.

Keywords:

Digital Humanities
Digital Scholarship Services
Digital Literacy
Digital Research
Library

Link to contribution: [Digital Scholarship Services at State- and University Library Hamburg](#)

QLK –Queer-linguistisches Korpus

Author: Lars Vorberger

Co-author: Miriam Hinterholzer

Corresponding Authors: lars.vorberger@uni-hamburg.de, miriam.hinterholzer@studium.uni-hamburg.de

The poster presents the project QLK –Queer Linguistic Corpus. The goal of the project is to build a queer linguistic corpus. The combination of queer linguistics, corpus linguistics, and data science represents a promising and challenging approach that is largely a desideratum for German. The project is applied and transfer-oriented and aims to build a research resource with a variety of tools to enable new ways of digitally exploring the relationship between language and sexuality. Both quantitative and qualitative methods will be used (e.g., crawling & manual text search). QLK launched in July in collaboration with the House of Computing and Data Science. The poster will present the first implementation approaches and an evaluation example besides the conception.

Keywords:

Queer Linguistics
Corpus Linguistics
Data Science
Research Resource
Mixed Methods

Link to contribution: [QLK –Queer-linguistisches Korpus](#)

Institute for Data Science Foundations

Author: Nihat Ay

Corresponding Author: nihat.ay@tuhh.de

This poster outlines the research and teaching profile of the Institute for Data Science Foundations at the Hamburg University of Technology.

Keywords:

Data Science, Machine Learning, Information Geometry, Embodied Intelligence, Causal Inference

Link to contribution: [Institute for Data Science Foundations](#)

Computational aspects of the active self

Authors: Frank Röder¹; Carlotta Langer¹

Co-authors: Jan Dohmen ¹; Manfred Epe ¹; Nihat Ay ¹

¹ *Institute for Data Science Foundations - Hamburg University of Technology*

Corresponding Authors: manfred.epe@tuhh.de, frank.roeder@tuhh.de, nihat.ay@tuhh.de, jan.dohmen@tuhh.de, carlotta.langer@tuhh.de

We, as humans, routinely talk about ourselves, but what is this “self”, how does it arise, and what influences it? What are the underlying mechanisms that help humans perceive themselves and act in the world? Recent work studies the so-called minimal self and shows how the sense of body-ownership, agency and control contribute to it. In our research, as part of the DFG SPP “the active self”, we investigate the minimal self and its emergence through sensorimotor experiences of embodied agents by utilizing information theory and reinforcement learning. Focusing on artificial agents enables us to extensively experiment and analyze their internal representations to further understand the minimal self.

Keywords:

minimal self
reinforcement learning
information theory
embodied intelligence
machine learning

Link to contribution: [Computational aspects of the active self](#)

Machine Learning based optimization of transverse beam emittance in Free Electron Lasers

Authors: Alexander Klemps¹; Nihat Ay

¹ *TUHH Institute for Data Science Foundations (E-21)*

Corresponding Authors: alexander.klemps@tuhh.de, nihat.ay@tuhh.de

In the context of Free Electron Lasers like the European XFEL at DESY Hamburg, it is essential to optimize the transverse emittance of a charged particle beam. Within the project OPAL-FEL we therefore aim to implement a Machine Learning based online optimal control process to minimize emittance.

The process will consist of a forward prediction model and an inverse feedback model.

Their implementations will incorporate experimental as well as artificial data, sampled by high fidelity digital twin simulations taking into account non linear space charge effects and other physical phenomena.

Additionally we are interested in studying the impact of different methodologies to be used such as Helmholtz machines, variational autoencoders and physics-informed neural networks.

Keywords:

free-electron laser
machine learning
transverse emittance
beam control
digital twin

Link to contribution: [Machine Learning based optimization of transverse beam emittance in Free Electron Lasers](#)

Natural Language Processing in the Legal Domain

Authors: Daniel Katz; Dirk Hartung¹; Lauritz Gerlach; Abhik Jana; Michael James Bommarito

¹ *Bucerius Law School gGmbH*

Corresponding Author: dirk.hartung@law-school.de

We summarize the current state of the field of NLP and Law with a specific focus on recent technical and substantive developments. To support our analysis, we construct and analyze a corpus of more than six hundred NLP and Law related papers published over 10 years. Our analysis highlights several major trends. Namely, an increasing number of papers written, tasks undertaken, and languages covered over the course of the past decade. We observe an increase in the sophistication of the methods which researchers deployed in this applied context. Legal NLP is beginning to match the methodological sophistication of general NLP. However, many questions in both the academic and commercial sphere remain open making Law an interesting domain for NLP research.

Keywords:

Legal Data Science
Natural Language Processing
Evaluation
Law

Link to contribution: [Natural Language Processing in the Legal Domain](#)

Flimma: a federated and privacy-aware tool for differential gene expression analysis

Authors: Amir Abbasinejad; Daniel Rückert; David B. Blumenthal; Georgios Kaissis; Jan Baumbach; Julian Klemm; Julian Matschinske; Julian Späth; Markus List; Mohammad Bakhtiari; Nina Wenke; Olga Zolotareva; Paolo Tieri; Reihaneh Torkzadehmahani; Reza Nasirigerdeh; Tobias Frisch

Aggregating transcriptomics data across hospitals can increase sensitivity and robustness of differential expression analyses, yielding deeper clinical insights. As data exchange is often restricted by privacy legislation, meta-analyses are frequently employed to pool local results. However, the accuracy might drop if class labels are inhomogeneously distributed among cohorts. Flimma (<https://featurecloud.ai/app/flimma>) addresses this issue by implementing the state-of-the-art workflow limma voom in a federated manner, i.e., patient data never leaves its source site. Flimma results are identical to those generated by limma voom on aggregated datasets even in imbalanced scenarios where meta-analysis approaches fail

Keywords:

Federated Learning
Differential Gene Expression
Privacy-Preserving learning
SMPC

Link to contribution: [Flimma: a federated and privacy-aware tool for differential gene expression analysis](#)

Privacy-preserving federated prognostic model learning from multi-centric questionnaire data to predict pain changes in osteoarthritis

Author: Supratim Das¹

Co-authors: Mahdie Rafiei ; Baumbach Jan ; Baumbach Linda

¹ *Institution for Computational Systems Biology | University of Hamburg*

Corresponding Authors: l.baumbach@uke.de, supratim.das@uni-hamburg.de, jan.baumbach@uni-hamburg.de, mahdie.rafeei@uni-hamburg.de

Multi-centric patient-derived data typically is decentralized and can not be integrated into one datacenter for privacy reasons. We address this problem using federated learning to train logistic, linear, and random forest regression models in a privacy-preserving fashion across data from multiple centers. We utilize data from GLA:D® osteoarthritis registries and compare federated models against models trained locally and models trained on the full data in a centralized fashion. We demonstrate that models trained in a federated, privacy-preserving fashion achieve performances comparable to those trained in a centralized fashion on the full data, but outperform models learned from smaller, local datasets. In summary, we proved the applicability of federated learning for prognostic model acquisition from multi-centric questionnaire-based healthcare data.

Keywords:

Federated Learning
Questionnaire data
Osteoarthritis
prediction model

Link to contribution: [Privacy-preserving federated prognostic model learning from multi-centric questionnaire data to predict pain changes in osteoarthritis](#)

Predicting multiple conformations of flexible proteins

Author: Karen Manalastas-Cantos¹

Co-authors: Kish Adoni²; Matthias Pfeifer³; Birgit Märten³; Kay Grünewald³; Konstantinos Thalassinos²; Maya Topf³

¹ *Center for Data and Computing in Natural Science, Universität Hamburg*

² *Institute of Structural and Molecular Biology, Division of Biosciences, University College London*

³ *Leibniz-Institut für Virologie (LIV), Centre for Structural Systems Biology (CSSB), Hamburg, Germany*

Corresponding Author: karen.manalastas@uni-hamburg.de

We developed a pipeline for predicting multiple conformations of flexible proteins. A range of conformations is first generated using the deep learning model AlphaFold2 (AF2), which are then filtered using distance constraints and solvent-accessibility data from crosslinking mass spectrometry (XL-MS), using two scoring functions we developed: the crosslink and monolink probability scores (XLP, MP). The scoring functions were first benchmarked on 200 proteins (each with 300 structural decoys) using simulated XL-MS data, before being tested on an experimental test dataset. We showed that AF2 alone can only identify two out of six conformations in the test dataset, while a combination of AF2 and XLP/MP was able to identify four of six conformations, highlighting the complementarity between AF2 and XL-MS.

Keywords:

protein structure prediction
computational modeling

Link to contribution: [Predicting multiple conformations of flexible proteins](#)

Shaking out Intrinsic Structural Variation in Femtosecond Protein Crystallography using Deep Learning

Authors: Florian Griese; José Licón-Saláiz; Samuel Flewett; Thomas White

Corresponding Authors: florian.griese@tuhh.de, samuel.flewett@desy.de, jose.licon-salaiz@uni-hamburg.de, twhite@desy.de

Current femtosecond crystallography data processing routines such as Cryst-FEL by T. White generally assume that identical molecular structure for each of the tens of thousands of protein crystals used. This assumption is however known to be unphysical, and with the recent development of deep learning technology, we are exploring whether we can build a net capable of shaking out the subtle structural variation information hidden behind a cloud of noise. We report progress on the development of a variational autoencoder, with an architecture inspired both from work in particle physics and Cryo EM.

Keywords:

Crystallography

Deep Learning

Variational Autoencoder

Link to contribution: [Shaking out Intrinsic Structural Variation in Femtosecond Protein Crystallography using Deep Learning](#)

Machine Learning for the Automated Selection and Reconstruction of Multi-Modal Nanotomography Data of Bone-Implant Interfaces

Author: Benedict Schacht

Co-authors: Berit Zeller-Plumhoff ; Christian Wilms ; Imke Greving ; Simone Frintrop

Corresponding Authors: berit.zeller-plumhoff@hereon.de, benedict.schacht@hereon.de, christian.wilms@uni-hamburg.de, simone.frintrop@uni-hamburg.de, imke.greving@hereon.de

Synchrotron radiation-based imaging and structural characterization provides unique opportunities to study hierarchical materials systems, such as bone. Correlative 3D imaging enables linking cellular connectivity and bone mineralization via transmission X-ray microscopy, X-ray diffraction and X-ray fluorescence. Generated data is terabyte-scale. We have acquired multi-modal datasets of bone from the bone-to-implant interface at the Hereon beamlines at PETRA III, DESY to study the impact of different implant types on bone mineralization. We employ machine learning methodologies from computer vision, such as segmentation and object detection, in combination with clustering to enhance data evaluation and future data acquisition.

Keywords:

Machine Learning

Link to contribution: [Machine Learning for the Automated Selection and Reconstruction of Multi-Modal Nanotomography Data of Bone-Implant Interfaces](#)

Integrated Climate Data Center

Authors: Stefan Kern; Remon Sadikni

Corresponding Authors: stefan.kern@uni-hamburg.de, remon.sadikni@uni-hamburg.de

The Integrated Climate Data Center (ICDC) is an established service component of the Center for Earth System Science and Sustainability (CEN) of the University of Hamburg (<http://www.cen.uni-hamburg.de/icdc>). Our main mandate is to provide easy-to-use observational (in situ, remote sensing) data products describing the Earth's climate system, to consult scientists across all career levels in data product usage, to support them in data publication, and to assist them in data management and adhering to FAIR principles. In addition to these tasks we are involved in visualizing data products, creating our own data sets, and data product quality assessment such as data set evaluation studies.

Keywords:

climate data, data publication, quality assessment, scientific user support

Link to contribution: [Integrated Climate Data Center](#)

Open Access Publishing for Hamburg Research. Quality-Assured Scholarly Open Access Publication Services

Authors: Isabella Meinecke¹; Oliver Krüger¹; Tobias Buck¹

¹ *Staats- und Universitätsbibliothek Hamburg*

Corresponding Authors: oliver.krueger@sub.uni-hamburg.de, isabella.meinecke@sub.uni-hamburg.de, tobias.buck@sub.uni-hamburg.de

The digital publication of research results is an important step in the life cycle of academic research. Today, a publication needs to be as open as possible in order to unleash its full potential. In the context of the Open Access (OA) transformation institutional publishing services are emerging that enable quality-assured OA publications according to professional standards.

We will present the publication services of the of the Hamburg State and University Library with the open access publication service Hamburg University Press as its core offer. We will focus on the benefits for researchers to publish their results according to scholar-led publishing: The technical infrastructure remains in science-related institutions, thus guaranteeing cost-effective and sustainable publication with increased visibility in the digital world.

Keywords:

open access, publication service, open science, quality assurance, standards, policies

Link to contribution: [Open Access Publishing for Hamburg Research. Quality-Assured Scholarly Open Access Publication Services](#)

Scholarly Publishing in the European Research Area (ERA) - Supporting the Visibility of Hamburg Research Results

Authors: Isabella Anna Greiner¹; Isabella Meinecke¹; Oliver Krüger¹

¹ *Staats- und Universitätsbibliothek Hamburg*

Corresponding Authors: isabella.meinecke@sub.uni-hamburg.de,
oliver.krueger@sub.uni-hamburg.de, isabella.greiner@sub.uni-hamburg.de

Open Science is on the political and funding agenda on the national (DFG, BMBF, KMK) as well as on the international level. The European Commission funds the projects CRAFT-OA, DIAMAS, and PALOMERA in order to establish diamond open access publishing services. Networked institutional infrastructure for journals and books is being built on high quality standards and will be integrated in platforms like the EOSC or OpenAIRE, thus supporting the formal quality and isibility of scholar-led publications.

The Hamburg State and University Library (SUB) participates in theses processes. Its publication services are excellently networked on the European level. With this poster, we will explain the above mentioned projects and show how publications of Hamburg researchers can benefit from these developments.

Keywords:

open access, open science, networking, scholarly communication, publishing, EU projects

Link to contribution: [Scholarly Publishing in the European Research Area \(ERA\) - Supporting the Visibility of Hamburg Research Results](#)

Legal Hypergraphs

Authors: Corinna Coupette; Dirk Hartung¹; Daniel Katz

¹ *Bucerius Law School gGmbH*

Corresponding Author: dirk.hartung@law-school.de

Network analysis has been widely adopted to investigate law as a complex system. However, the utility of dynamic higher-order networks in the legal domain has remained largely unexplored. Setting out to change this, we introduce temporal hypergraphs as a powerful tool for studying legal network data. Temporal hypergraphs generalize static graphs by (i) allowing any number of nodes to participate in an edge, and (ii) permitting nodes or edges to be added, modified, or deleted. We develop models and methods to explore legal hypergraphs that evolve over time and present case studies on legal citation and collaboration networks that change over more than 70 years. Our work demonstrates the potential of dynamic higher-order networks for studying complex legal systems.

Keywords:

legal networks
temporal networks
higher-order networks
hypergraphs
complex systems

Link to contribution: [Legal Hypergraphs](#)

A Personalized Web Application to Predict Changes in Knee Pain among Patients with Osteoarthritis Participating at the GLA:D® program

Author: Mahdie Rafiei¹

Co-authors: Supratim Das ; Jan Baumbach ²; Linda Baumbach ³

¹ *Institute for Computational Systems Biology, University of Hamburg, Hamburg, Germany*

² *Director of the Institute for Computational Systems Biology University of Hamburg*

³ *Department of Health Economics and Health Services Research, University Medical Center Hamburg-Eppendorf, Germany*

Corresponding Authors: mahdie.rafiee@uni-hamburg.de, supratim.das@uni-hamburg.de, jan.baumbach@uni-hamburg.de, l.baumbach@uke.de

GLA:Dai (<https://apps.cosy.bio/gladaai>) is a web application for predicting individualized changes in knee pain for patients considering enrollment with the GLA:D® program (Good Life with osteoArthritis in Denmark). The statistical model is based on self-reported patient information to predict the individual change in pain intensity (VAS scale 0 to 100) from before to after GLA:D® enrollment. We select important variables from 51 candidates using the “Gini impurity” and clinical reasoning. We trained a random forest regression model and compared it against predicting the mean change in pain. Our individualized model yields a 7% improvement, contributes to personalized healthcare by enabling patients to estimate their change in pain intensity, and allows for more data-driven treatment decisions.

Keywords:

osteoarthritis,
prediction,
pain Intensity,
web-based tool,
exercise therapy

Link to contribution: [A Personalized Web Application to Predict Changes in Knee Pain among Patients with Osteoarthritis Participating at the GLA:D® program](#)

Privacy-Preserving Federated Differential Protein Expression Analysis with FedProt

Author: Yuliya Burankova¹

Co-authors: Olga Zolotareva²; Miriam Abele³; Fabian Gruhn⁴; Christine von Törne⁴; Teresa Barth⁵; Lisa Schweizer⁶; Pieter Giesbertz⁷; Mohammad Bakhtiari²; Anne Hartebrodt⁸; Tobias Frisch⁸; Chen Meng³; Julian Matschinske²; Julian Späth²; Richard Röttger⁸; Veit Schwämmle⁸; Stefanie M. Hauck⁴; Stefan Lichtenthaler⁷; Axel Imhof⁵; Matthias Mann⁶; Christina Ludwig³; Bernhard Küster³; Jan Baumbach²

¹ *Technical University of Munich, University of Hamburg*

² *University of Hamburg*

³ *Technical University of Munich*

⁴ *Helmholtz Center Munich*

⁵ *Ludwig Maximilian University of Munich*

⁶ *Max Planck Institute of Biochemistry*

⁷ *German Center for Neurodegenerative Diseases*

⁸ *University of Southern Denmark*

Corresponding Authors: yuliya.burankova@uni-hamburg.de, olga.zolotareva@uni-hamburg.de

Integrating distributed patient-derived proteomics data poses privacy concerns, risking genotype reconstruction attacks. To enable privacy-preserving analysis of distributed data, we developed FedProt, the first tool for federated differential protein expression analysis. Based on DEqMS and utilizing the hybrid methodology of federated learning and additive secret sharing from Flimma, FedProt allows collaborative model training without violating data privacy. Tested on a DIA dataset of 99 *Escherichia coli* samples from five research centers, FedProt results matched the centralized DEqMS results and did it more precisely than typical meta-analyses. FedProt manages proteomic complexity, enhancing statistical power without sacrificing accuracy. Its user-friendly implementation will be accessible as a FeatureCloud App (<https://featurecloud.ai>), making privacy-aware differential protein expression analysis available to a broad community.

Keywords:

Differential Protein Expression Analysis
Proteomics Data Privacy
Federated Learning

Link to contribution: [Privacy-Preserving Federated Differential Protein Expression Analysis with FedProt](#)

Detection of alternative splicing: deep sequencing or deep learning

Author: Lena Maria Hackl

Co-authors: Jan Baumbach ¹; Olga Tsoy ¹

¹ *Universität Hamburg*

Corresponding Authors: olga.tsoy@uni-hamburg.de, lena.hackl@uni-hamburg.de, jan.baumbach@uni-hamburg.de

Alternative splicing enables the expression of a variety of isoforms coding for functionally diverse proteins from a single gene. RNA sequencing (RNAseq) has become the state-of-the-art tool for profiling the transcriptome, but still reliable detection of alternative splicing events in RNAseq from virus-infected cells with low number of reads is challenging. Few computational tools, such as Junction coverage compatibility (JCC) and Deep Splice, try to address this. We compare those tools' performance with sequence-based prediction tool SpliceAI on subsampled 50M read data from four dilated cardiomyopathy patients. The tools were found to show high precision but poor recall. Using these insights, we developed SpliceStat, leveraging both sequence and RNAseq data for improved accuracy in alternative splicing prediction.

Keywords:

alternative splicing
RNA sequencing
virus-infected cells
computational systems biology
bioinformatics

Link to contribution: [Detection of alternative splicing: deep sequencing or deep learning](#)

Identification of differentially expressed biclusters for unsupervised patient stratification

Authors: Michael Hartung¹; Andreas Maier²; Fernando Miguel Delgado Chaves¹; Yuliya Burankova³; Olga Isaeva⁴; Daniel He⁵; Katharina Kaufmann¹; Fábio Malta de Sá Patróni⁶; Alexey Savchik⁷; Zoe Chervontseva¹; Niklas Probul¹; Alexandra Abisheva⁸; Evgenia Zotova⁸; Olga Tsoy¹; David Blumenthal⁹; Martin Ester¹⁰; Olga Zolotareva¹; Jan Baumbach¹

¹ *Universität Hamburg*

² *Institute for Computational Systems Biology*

³ *Technical University of Munich, University of Hamburg, Germany*

⁴ *The Netherlands Cancer Institute*

⁵ *University of British Columbia*

⁶ *University of Campinas*

⁷ *ACMetric*

⁸ *Altius Institute for Biomedical Sciences*

⁹ *Universität Erlangen*

¹⁰ *Simon Fraser University; Vancouver Prostate Centre*

Corresponding Authors: michael.hartung@uni-hamburg.de, andreas.maier-1@uni-hamburg.de, olga.tsoy@uni-hamburg.de, david.b.blumenthal@fau.de, yuliya.burankova@uni-hamburg.de, fernando.miguel.delgado-chaves@uni-hamburg.de, olga.zolotareva@uni-hamburg.de, jan.baumbach@uni-hamburg.de

Unexplored molecular heterogeneity of human diseases causes treatment inefficacy and hinders the investigation of causative disease mechanisms. Since the number and frequencies of disease subtypes are usually unknown, unsupervised methods are applied to omics data to identify patients subgroups with similar molecular profiles. Here, we present UnPaSt, a novel biclustering algorithm for unsupervised patient stratification and demonstrate its superior performance compared to traditionally used clustering, factorization, and biclustering methods in benchmarks with simulated and real data. Moreover, besides accurate identification of well-known PAM50 subtypes of breast cancer, UnPaSt detected the rare neuroendocrine subtype, which was overlooked in previous analyses due to its low frequency.

Keywords:

patient stratification
clustering
biclustering

disease heterogeneity
omics

Link to contribution: [Identification of differentially expressed biclusters for unsupervised patient stratification](#)

Decentralized Batch Effect Correction in Multi-Center Biomedical Studies: Introducing FedComBat, a Federated Learning Approach

Author: Mohammad Mahdi Kazemi Majdabadi

Co-authors: Jan Baumbach ; Olga Zolotareva ; Yuliya Burankova

Corresponding Authors: jan.baumbach@uni-hamburg.de, olga.zolotareva@uni-hamburg.de, yuliya.burankova@uni-hamburg.de, mohammad.kazemi.majdabadi@uni-hamburg.de

Batch effect correction is a pivotal challenge in biological data analysis, affecting fields such as genomics, transcriptomics, and proteomics. Traditional correction methods like `limma::removeBatchEffect` and `ComBat` require centralization of data from various institutions, which can be problematic due to privacy and data governance concerns. Our innovative approach, named FedComBat, circumvents this issue by utilizing federated learning techniques. Unlike conventional methods, FedComBat allows for decentralized batch effect correction, enabling data to remain at their respective institutes while still achieving comparable analysis results. This approach not only maintains data privacy but also paves the way for efficient multi-center biomedical studies.

Keywords:

Batch Effect Correction, Federated Learning, FedComBat, Data Privacy, Decentralized Analysis

Link to contribution: [Decentralized Batch Effect Correction in Multi-Center Biomedical Studies: Introducing FedComBat, a Federated Learning Approach](#)

The CDCS - A Collaborative Research Center for Data Science in Hamburg

Authors: Annett Ungethüm¹; Antonin Sulc¹; José Licón-Saláiz¹; Karen Manalastas-Cantos²; Khaliq Newaz³; Lennart Rustige¹; Lennart Rustige; Patrick L.S. Connor; Samuel Flewett

¹ *Universität Hamburg*

² *Center for Data and Computing in Natural Science, Universität Hamburg*

³ *University of Hamburg*

Corresponding Authors: annett.ungethuem@uni-hamburg.de, lennart.rustige@uni-hamburg.de, jose.licon-salaiz@uni-hamburg.de, karen.manalastas@uni-hamburg.de, khaliq.newaz@uni-hamburg.de, antonin.sulc@uni-hamburg.de

The Center for Data and Computing in Natural Sciences (CDCS) is a joint facility of the University of Hamburg, the German Electron Synchrotron DESY and the Hamburg University of Technology. An increasing amount of scientific research projects rely on modern information technology. To accommodate for this development, the CDCS incorporates four Cross-Disciplinary Labs (CDLs) which are supported by a Computational Core Unit (CCU). The CDLs focus on the following topics:

CDL1: Computational Astro- and Particle Physics

CDL2: Computational Poton Science

CDL3: Computational Systems Biology

CDL4: Computational Control of Accelerators

We are going to give an overview of the CDCS and the current projects we are working on, as well as the first results of this collaborative setup.

Keywords:

CDCS

Data Science

Research Center

Collaborative Research

Link to contribution: [The CDCS - A Collaborative Research Center for Data Science in Hamburg](#)

ERIS: Hamburg Information System on the Representation of Greek and Roman Violence

Authors: Justine Diemke¹; Werner Rieß¹

¹ *Universität Hamburg*

Corresponding Authors: werner.riess@uni-hamburg.de, justine.diemke@uni-hamburg.de

Das MyCoRe-basierte Informationssystem Eris, welches seit 2012 unter der Leitung von Prof. Dr. Werner Rieß von Mitarbeiter*innen aus dem Arbeitsbereich Alte Geschichte an der Universität Hamburg aufgebaut wird, ermöglicht zum ersten Mal eine inhaltliche und multidimensionale Erschließung aller historiographischen und biographischen Texte zur antiken Gewalt.

Keywords:

ancient history
text mining
named entity recognition
data visualization

Link to contribution: [ERIS: Hamburg Information System on the Representation of Greek and Roman Violence](#)

FAIR-Based Strategy for Data Management and Handling in Large Climate Modeling Projects: Experiences from PalMod-II

Author: Swati Gehlot¹

Co-authors: Andrea Lammert ²; Hannes Thiemann ²; Karsten Peters-von Gehlen ¹

¹ DKRZ GmbH

² DKRZ

Corresponding Authors: peters@dkrz.de, gehlot@dkrz.de

Here we present the research data management (RDM) approach developed and employed for project PalMod-II, making large scale climate model data available for reuse by the global paleo-climate science community in-line with the FAIR (Findable, Accessible, Inter-operable, Reusable) data principles. The compilation and maintenance of a project-wide data management plan (DMP) was prepared and maintained for keeping the project on track and serving as a central focal point for all data-related aspects. These include the specification of data responsible scientists, allocation of storage and compute resources on a high-performance computing system, documentation of simulation output requirements, definition of (meta)data standardisation, and publication workflows in-line with the FAIR data principles.

Keywords:

RDM, DMP, FAIR, Climate model data, data standardization

Link to contribution: [FAIR-Based Strategy for Data Management and Handling in Large Climate Modeling Projects: Experiences from PalMod-II](#)

World Data Center for Climate (WDCC) - Improvement of FAIRness of an established repository

Author: Eileen Hertwig¹

Co-authors: Andrea Lammert ¹; Heinke Höck ¹; Andrej Fast ¹; Hannes Thiemann ¹

¹ DKRZ

Corresponding Author: hertwig@dkrz.de

The World Data Center for Climate (WDCC), hosted by DKRZ, provides access to and offers long-term archiving for datasets relevant for Earth System research in a highly standardized manner. Its services are aimed at researchers who produce data and those who re-use published data for new research.

To meet user's needs it is essential to ensure high quality of data, i.e. guaranteeing that datasets are really Findable, Accessible, Interoperable, and Reusable (FAIR). FAIRness at WDCC has been systematically assessed and developments aim at further improvements, including publishing data as citable entities (assigning DOIs to dataset collections and now PIDs to individual datasets), implementing the schema.org standard (machine-actionable metadata), and continually enhancing the quality of metadata (e.g. including funder information).

Keywords:

WDCC

FAIR

Repository

Research data

Link to contribution: [World Data Center for Climate \(WDCC\) - Improvement of FAIRness of an established repository](#)

DRaCOoN: Differential Regulation and CO-expression Networks

Author: Fernando Miguel Delgado Chaves¹

Co-authors: Ferdinand Spurny¹; Jan Baumbach¹; Olga Zolotareva¹

¹ *Universität Hamburg*

Corresponding Authors: olga.zolotareva@uni-hamburg.de, jan.baumbach@uni-hamburg.de, fernando.miguel.delgado-chaves@uni-hamburg.de

To address challenges in identifying crucial regulatory elements, we introduce DRaCOoN, a data-driven method for differential co-expression and regulatory networks between unique conditions. DRaCOoN uses established metrics to better handle large datasets and offers algorithmic and benchmarking strategies for accuracy and relevance. The method employs permutation tests and a background model for significance estimation. Numba optimization enhances runtimes and scalability.

We tested DRaCOoN using simulated datasets and node-based perturbations. DRaCOoN's entropy-based metrics outperform others in multiple scenarios, particularly with datasets containing many perturbed genes. A GSEA-inspired approach revealed optimal performance when combining entropy-based association metrics with the shift differential metric.

In summary, DRaCOoN effectively identifies key factors in complex biological processes and shows promise for studying intricate conditions or diseases.

Keywords:

Differential Regulation Analysis,
Co-expression Networks,
disease mechanism elucidation,
Differential networking,

Link to contribution: [DRaCOoN: Differential Regulation and CO-expression Networks](#)

Automated symptom evaluation using sensor data from wearables in patients with Parkinson's disease utilizing explainable AI

Authors: Alexander Wiederhold¹; Qi Rui Zhu²; Sören Spiegel¹; Frank Ückert¹; Christopher Lukas Gundler¹

¹ *Universität Hamburg*

² *Universität Osnabrück*

Corresponding Authors: s.spiegel@uke.de, christopher.gundler@uni-hamburg.de, zqirui@uni-osnabrueck.de, frank.ueckert@uni-hamburg.de, alexander.wiederhold@studium.uni-hamburg.de

Parkinson's disease is a neurodegenerative disorder that affects primarily dopaminergic neurons often showing a characteristic tremor or akinesia. While the severity estimation remains a stationary UPDRS scoring, we aim for an additional sensor-driven approach to develop an automated symptom evaluation for a second opinion. Our team collected over 200 million samples of accelerometer data and over 5000 corresponding UPDRS labels at the neurology department of the University Hospital Hamburg-Eppendorf. We implemented an explainable AI algorithm as well as a high-performance baseline model. First results show a high accuracy even for explainable classifiers. Consequently, we determined that a simple algorithm might be superior to a complex model in clinical routine to support clinicians' decision process.

Keywords:

Parkinson
Neurology
Sensor-driven-approach
AI
Accelerometer-data

Link to contribution: [Automated symptom evaluation using sensor data from wearables in patients with Parkinson's disease utilizing explainable AI](#)

On the power and limits of structure-based exon-exon interaction prediction

Author: Jeanine Liebold¹

Co-authors: Aylin Del Moral Morales²; Karen Manalastas-Cantos³; Olga Tsoy¹; Stefan Kurtz¹; Jan Baumbach¹; Khaliq Newaz⁴

¹ *Universität Hamburg*

² *Universidad Autónoma Metropolitana-Cuajimalpa (UAM-C), Mexico*

³ *Centre for Structural Systems Biology*

⁴ *University of Hamburg*

Corresponding Authors: aylindmm@gmail.com, jeanine.liebold@uni-hamburg.de, jan.baumbach@uni-hamburg.de, stefan.kurtz@uni-hamburg.de, olga.tsoy@uni-hamburg.de, khaliq.newaz@uni-hamburg.de

To perform important cellular functions, Proteins physically interact with each other through protein-protein interactions (PPIs). However, PPIs are not constant but change with varying cellular conditions, e.g., those induced by infections or other diseases. These changes in PPI can be caused by alternative splicing (AS), a biological process, by which a single gene gives rise to multiple proteins. Our goal is to predict PPIs that occur in the event of AS. We explore whether we can exploit existing computational PPI interface prediction methods, based on geometric and convolutional neural networks, to create a network of exon-exon interactions (EEIs). A main task is to evaluate how effectively such an EEI network can be used to predict PPI rewiring caused by AS.

Keywords:

Protein-protein interaction

Alternative splicing

Exon-exon interaction

Link to contribution: [On the power and limits of structure-based exon-exon interaction prediction](#)

sPLINK: a hybrid federated tool as a robust alternative to meta-analysis in genome-wide association studies

Authors: Julian Matschinske¹; Mohammad Bakhtiari¹

¹ *PhD Student*

Corresponding Authors: mohammad.bakhtiari@uni-hamburg.de,
julian.matschinske@uni-hamburg.de

Meta-analysis has been established as an effective approach to combining summary statistics of several genome-wide association studies (GWAS). However, the accuracy of meta-analysis can be attenuated in the presence of cross-study heterogeneity. We present sPLINK, a hybrid federated and user-friendly tool, which performs privacy-aware GWAS on distributed datasets while preserving the accuracy of the results. sPLINK is robust against heterogeneous distributions of data across cohorts while meta-analysis considerably loses accuracy in such scenarios. sPLINK achieves practical runtime and acceptable network usage for chi-square and linear/logistic regression tests. sPLINK is available at <https://featurecloud.ai/app/splink>

Keywords:

GWAS, privacy-aware

Link to contribution: [sPLINK: a hybrid federated tool as a robust alternative to meta-analysis in genome-wide association studies](#)

Research data management at the German Institute for Global and Area Studies (GIGA)

Author: Catharina Hibler

Corresponding Author: catharina.hibler@giga-hamburg.de

Um einen kontinuierlichen Service im Forschungsdatenmanagement (FDM) am German Institute for Global and Area Studie (GIGA) zu gewährleisten, gibt es dort seit 2014 eine zuständige Person für diesen Themenbereich. 2017 führte das GIGA eine institutseigene Forschungsdatenleitlinie ein. Seitdem wurden und werden Kooperationen zu Infrastrukturpartnern aufgebaut und ein andauernder Kulturwandel Richtung „Daten teilen“ vorangetrieben. Die Komplexität des Themas am GIGA wird an folgenden Merkmalen sichtbar: Verschiedene sozialwissenschaftliche Disziplinen, Methoden und Datentypen,

erfordern unterschiedliche spezialisierte Infrastrukturlösungen. Ethische Aspekte, wie beispielsweise die Sicherheit von Interviewpartnern oder das Thema Datensicherheit spielen eine große Rolle. Die Forschung findet in verschiedenen Sprach- und Rechtsräumen statt. Das Poster möchte einen Einblick in das FDM am GIGA anhand von Beispielen geben.

Keywords:

research data management

Link to contribution: [Research data management at the German Institute for Global and Area Studies \(GIGA\)](#)

Hyperbolic Learning on Omics Data

Authors: Ilya Petrov; Kate Trifonova; Semyon Karaban; Ariana Migutskaya; Jackie Khodzhaeva; Mikhail Gelfand; Zoe Chervontseva

Corresponding Author: zoia.chervontseva@uni-hamburg.de

Many objects in biology are connected by hierarchical relationships. To clarify how snippets of data are associated, we apply embeddings, that is, mapping of multidimensional objects into space so that similar objects are positioned at close points. Analyzing such data with tools operating in Euclidean spaces is problematic as the tools may not account for the underlying data hierarchy. We applied several dimensionality reduction methods based on hyperbolic geometry to diverse datasets including gene expression, gene interaction, microbiome composition, and gene phyletic patterns. We compare hyperbolic embeddings to the more conventional ones (PCA, UMAP) and discuss the difference.

Keywords:

hyperbolic learning, dimensionality reduction

[Link to contribution: Hyperbolic Learning on Omics Data](#)

Exploring supervised machine learning algorithms for predicting pitting corrosion damage of AA1050 exposed to ethanol-containing fuels

Author: Lukas Jarren

Co-authors: Christian Feiler ; Daniel Höche ; Eugen Gazenbiller ; Matthias Oechsner ; Mikhail Zheludkevich ; Rüdiger Reitz ; Visheet Arya

Corresponding Authors: lukas.jarren@hereon.de, ruediger.reitz@tu-darmstadt.de, daniel.hoeche@hereon.de, matthias.oechsner@tu-darmstadt.de, mikhail.zheludkevich@hereon.de, visheet.arya@tu-darmstadt.de, eugen.gazenbiller@hereon.de, christian.feiler@hereon.de

Aluminum alloys are crucial in car manufacturing, but the rise of bioethanol fuels has raised concerns about alcoholate corrosion. Gazenbiller et al. have already explored temperature-induced alcoholate pitting corrosion in AA1050 aluminum exposed to anhydrous ethanol.¹ To extend this study, a specially constructed reactor is used, which allows for in situ tracking of chemical corrosion damage by providing video and image documentation of the AA1050 corrosion exposed to different ethanol-gasoline blends. We leverage this data to estimate corrosion damage by utilizing artificial neural networks, ensemble methods (random forest, XGBoost) and support vector machines. These models serve as black-box estimators within specific usage limits and using both weight loss and the corroded surface area from images as corrosion damage quantification.

Keywords:

machine learning
corrosion

Link to contribution: [Exploring supervised machine learning algorithms for predicting pitting corrosion damage of AA1050 exposed to ethanol-containing fuels](#)

Selection of Effective Corrosion Inhibitors for Aluminium Alloys based on Data-Driven Techniques

Author: Lisa Sahlmann

Co-authors: Can Özkan ; Christian Feiler ; Daniel Höche ; J.M.C. Mol ; Mikhail Zheludkevich ; Natalia Konchakova ; Peyman Taheri ; Sviatlana Lamaka

Corresponding Authors: daniel.hoeche@hereon.de, mikhail.zheludkevich@hereon.de, lisa.sahlmann@hereon.de, p.taheri@tudelft.nl, j.m.c.mol@tudelft.nl, gercek-canozkan@gmail.com, sviatlana.lamaka@hereon.de, natalia.konchakova@hereon.de, christian.feiler@hereon.de

The replacement of highly protective but toxic hexavalent-chromium-based corrosion inhibitors with novel and safer inhibitors for aluminum (Al) alloys is urgently required. Small organic molecules have emerged as safe and potent alternatives showing promising corrosion inhibition for Al alloys ¹. Experimental techniques alone can only screen a tiny fraction of the vast chemical space of available compounds despite tremendous progress in the screening of potential inhibitors by efficient high throughput techniques. Quantitative structure-property relationship models offer the opportunity to preselect a list of promising candidates for experimental testing without in-depth knowledge of the underlying chemical mechanisms. In this work, we used a set of small organic molecules to develop a data driven model to predict their corrosion response for AA2024-T3.

Acknowledgements: The VIPCOAT project: H2020-NMBP-TO-IND-2020, Grand Agreement 952903.

¹ D.A. Winkler, M. Breedon, P. White, A.E. Hughes, E.D. Sapper, I. Cole, Using high throughput experimental data and in silico models to discover alternatives to toxic chromate corrosion inhibitors, *Corros. Sci.* 106 (2016) 229-235.

Keywords:

Corrosion Science
Quantitative structure-property relationships
Machine Learning

Link to contribution: [Selection of Effective Corrosion Inhibitors for Aluminium Alloys based on Data-Driven Techniques](#)

A Computational Approach to Magnesium Corrosion Engineering

Author: Tim Würger¹

Co-authors: Bahram Vaghefinazari ¹; Christian Feiler ¹; David Winkler ²; Di Mei ¹; Mikhail Zheludkevich ¹; Robert Meißner ³; Sviatlana Lamaka ¹

¹ *Helmholtz-Zentrum Hereon*

² *La Trobe University*

³ *Technische Universität Hamburg*

Corresponding Author: tim.wuerger@hereon.de

Small organic molecules constitute useful materials to modify the service environment of light metals, such as magnesium (Mg). As the lightest engineering metal, Mg is promising for advanced technologies that will tackle climate change through improved battery technologies and advanced transport applications. However, due to its high chemical reactivity, target applications require tailoring of the degradation properties. The vast chemical space of potentially effective compounds can be explored by quantitative structure-property relationship (QSPR) models. We use molecular similarities in a kernel ridge regression model to predict the experimental performance of a large number of chemicals. The model robustness is confirmed by blind validation. Finally, a workflow is presented that facilitates the automated discovery of chemicals with desired dissolution modulating properties.

Keywords:

Magnesium

Corrosion

Machine Learning

Design of Experiments

Link to contribution: [A Computational Approach to Magnesium Corrosion Engineering](#)

Leveraging Semantic Information for Efficient Self-Supervised Emotion Recognition with Audio-Textual Distilled Models

Authors: Danilo de Oliveira¹; Navin Raj Prabhu²; Timo Gerkmann

¹ *Signal Processing (SP), Universität Hamburg*

² *Signal Processing*

Corresponding Authors: timo.gerkmann@uni-hamburg.de, danilo.oliveira@uni-hamburg.de, navin.raj.prabhu@uni-hamburg.de

In large part due to their implicit semantic modeling, self-supervised learning (SSL) methods have significantly increased valence recognition performance in speech emotion recognition (SER) systems. Yet, their large size may often hinder implementation in applications such as virtual assistants and digital customer service. In this work, we analyze the relevance for SER of each of HuBERT's layers, showing that shallower/deeper layers are more important for arousal/valence recognition, respectively. This motivates the use of additional textual information for accurate valence recognition, as the distilled model lacks the depth of its teacher. Thus, we propose a framework that, while having only ~20% of the trainable parameters of a large SSL model, achieves on par SER performance on the MSP-Podcast dataset.

Keywords:

speech emotion recognition
self-supervised learning
knowledge distillation
paralinguistics
semantics

Link to contribution: [Leveraging Semantic Information for Efficient Self-Supervised Emotion Recognition with Audio-Textual Distilled Models](#)

Feature Reduction Improves Drug Response Prediction

Author: Farzaneh Firoozbakht¹

Co-authors: Behnam YOUSEFI²; Olga Tsoy³; Jan Baumbach⁴; Benno Schwikowski²

¹ *Hamburg University*

² *Institut Pasteur*

³ *Hamburg Universität*

⁴ *Universität Hamburg*

Corresponding Authors: jan.baumbach@uni-hamburg.de, olga.tsoy@uni-hamburg.de, farzaneh.firoozbakht@uni-hamburg.de, yousefi.bme@gmail.com, benno.schwikowski@pasteur.fr

Machine learning approaches play an important role in precision medicine by predicting drug responses based on the molecular profiles of patients. However, the high dimensionality of molecular profiles requires feature reduction. To assess the effectiveness of different feature reduction strategies for drug response prediction (DRP) on cancer, we evaluated four feature reduction methods on cell line and tumor gene expression data. These methods include Landmark genes, drug pathway genes, pathway activities, and Transcription Factor (TF) activities. Our analysis showed that TF activities provide the best predictive performance. While our study was necessarily limited to a discrete spectrum of feature reduction methods, we believe that it can provide a new and meaningful reference point for machine learning research in DRP.

Keywords:

Drug response prediction, Feature reduction, Transcription factor activity

Link to contribution: [Feature Reduction Improves Drug Response Prediction](#)

Studying protein 3D structures using network science

Author: Khalique Newaz¹

Co-author: Tijana Milenkovic²

¹ *University of Hamburg*

² *University of Notre Dame*

Corresponding Authors: tmilenko@nd.edu, khalique.newaz@uni-hamburg.de

Understanding protein 3-dimensional (3D) structures is important because functions of proteins depend on them. We proposed (static or dynamic) network approaches to model protein 3D structures as protein structure networks (PSNs). Static PSNs model the whole 3D structure of a protein as a single-layer PSN. Because the folding of a protein entails some protein parts folding before others, we additionally proposed to model a protein as a dynamic (i.e., multi-layer) PSN. Using 44, 000 proteins, we evaluated our PSN models in the task of protein structural classification (PSC); a supervised problem of assigning proteins into pre-defined structural classes. We showed that our models outperformed state-of-the-art PSC approaches, showing our models' promising application in understanding 3D structures of proteins.

Keywords:

Protein 3D structures

Protein structure networks

Protein structural classification

Link to contribution: [Studying protein 3D structures using network science](#)

NeDRex - an integrative and interactive network medicine platform for drug repurposing

Authors: James Skelton¹; Sepideh Sadegh²

Co-authors: Andreas Maier²; Jan Baumbach²

¹ *School of Computing, Newcastle University, Newcastle upon Tyne, Uk*

² *Institute for Computational Systems Biology, University of Hamburg, Hamburg, Germany*

Corresponding Authors: jan.baumbach@uni-hamburg.de, andreas.maier-1@uni-hamburg.de, james.skelton3@googlemail.com, ssadegh@health.sdu.dk

Drug discovery's prevailing paradigms grapple with efficacy challenges. Drug repurposing, with its agility and cost-efficiency, emerges as a potent alternative. Complex diseases are unraveled through disease modules, highlighting potential drug targets discernible in silico via network algorithms. However, essential data remains dispersed across various databases, underscoring a demand for unified tools. We present NeDRex, an integrated platform for network-based drug repurposing at [NeDRex.net](https://nedrex.net). NeDRex seamlessly integrates diverse data, underpinning tailored biomedical inquiries. Enhanced by NeDRexApp for Cytoscape and [NeDRex-Web](#) (NeDRex Web Interface), and emphasizing the expert-in-the-loop paradigm, it facilitates iterative feedback, allowing researchers to meld domain-specific knowledge with computational insights. Built on the same paradigm [CoVex](#) a COVID-19 drug repurposing platform was realized.

Keywords:

network medicine
drug repurposing
knowledge graph
disease module identification

Link to contribution: [NeDRex - an integrative and interactive network medicine platform for drug repurposing](#)

The AIME registry for artificial intelligence in biomedical research

Author: Julian Matschinske¹

Co-authors: David Blumenthal²; Jan Baumbach¹; Nina Wenke³; Olga Zolotareva

¹ *Universität Hamburg*

² *Universität Erlangen*

³ *Institute for Computational Systems Biology*

Corresponding Authors: julian.matschinske@uni-hamburg.de,
david.b.blumenthal@fau.de, nina.wenke@uni-hamburg.de, olga.zolotareva@uni-hamburg.de, jan.baumbach@uni-hamburg.de

Artificial intelligence's vast potential in biomedicine demands rigorous research with comprehensive method and data details. Despite journal peer review and checklists, variations in validation, metrics, data/code accessibility persist in publications, hampering comparability and reproducibility. The AIME registry, a community-driven platform, enhances biomedical AI model accessibility, reproducibility, and usability. Numerous examples illustrate the critical need for comprehensive reporting in biomedical research to avoid misleading claims. An extensive yet adaptable questionnaire (at <https://aime-registry.org>) can significantly enhance research documentation and reproducibility in this field.

Keywords:

accessibility, reproducibility, usability of biomedical AI models

Link to contribution: [The AIME registry for artificial intelligence in biomedical research](#)

Inferring Tissue-Specific Alternative Splicing from Blood Samples

Authors: Jan Baumbach¹; Olga Tsoy¹; Silke Van Elferen

¹ *Universität Hamburg*

Corresponding Author: silke.van.elferen@uni-hamburg.de

This study explores an innovative, less invasive approach to diagnosing complex diseases like dilated cardiomyopathy (DCM) by predicting target tissue expression profiles using gene expression and alternative splicing profiles from blood samples. Machine learning approaches, specifically dimensionality reduction and linear regression, are utilized. Preliminary results indicate that while this approach is insufficient for robust tissue splicing factor expression predictions, it holds significant promise. The study suggests that incorporating additional data types (e.g., proteomics, RNA-Seq on exosomes) or alternative methods (e.g., deep learning techniques) may yield stronger predictions. Despite limitations, this study emphasizes the potential of this approach and encourages further research.

Keywords:

Systems biology, tissue expression prediction, linear regression, dimensionality reduction

Link to contribution: [Inferring Tissue-Specific Alternative Splicing from Blood Samples](#)

Visualization and Sign Representation in the Digital Dictionary DGS (DW-DGS)

Authors: Anke Müller¹; Felicitas Otte¹; Gabriele Langer¹; Sabrina Wähl¹; Thomas Hanke¹

¹ IDGS / Universität Hamburg

Corresponding Author: gabriele.langer@uni-hamburg.de

Creating a dictionary of a visual language for diverse target groups presents unique challenges and calls for new technological solutions. We show ways of visualizing information, representing German Sign Language, and navigating the use of different languages in the *Digitales Wörterbuch DGS (DW-DGS)*. As signs can best be represented in video form, we created a new format called a micon for the representation of DGS in the form of a moving thumbnail. The visualization of further information on a sign's use includes maps showing regional variation directly generated from DGS corpus language data. As written language is a necessity for navigation, we discuss the use of written German as both a metalanguage and an object language in the dictionary.

Keywords:

German Sign Language
Dictionary
Sign representation
Distribution maps
metalanguage

Link to contribution: [Visualization and Sign Representation in the Digital Dictionary DGS \(DW-DGS\)](#)

MY DGS –ANNIS: ANNIS and the Public DGS Corpus

Authors: Amy Isard¹; Reiner Konrad¹

¹ *IDGS, Universität Hamburg*

Corresponding Author: amy.isard@uni-hamburg.de

The Public DGS Corpus is a collection of annotated German Sign Language data which is available through several different interfaces. In 2022 we published our third data portal (MY DGS –ANNIS), using the ANNIS browser-based corpus software. ANNIS is a corpus query tool for visualization and querying of multi-layer corpus data which has its own query language, AQL, and is accessed from a web browser without requiring a login. It allows more complex queries and visualizations than those provided by our existing research portal. We used the Pepper platform to convert the corpus data into the Salt data format used by ANNIS.

Keywords:

sign language
corpus query and visualization tools
digital humanities

Link to contribution: [MY DGS –ANNIS: ANNIS and the Public DGS Corpus](#)

The DW-DGS: A corpus-based dictionary of German Sign Language is being compiled

Authors: Anke Müller¹; Gabriele Langer¹; Sabrina Wähl¹; Thomas Hanke¹

¹ IDGS / Universität Hamburg

Corresponding Author: gabriele.langer@uni-hamburg.de

The Digitales Wörterbuch DGS (DW-DGS) is the first corpus-based dictionary of German Sign Language (DGS). As such, it supplies dictionary users with new types of information based on the analysis of real natural language data. This, for example, includes information on collocations and examples taken directly from the corpus. We show how the various target groups of the dictionary may benefit from the different access points and information types in the dictionary. As a corpus-based dictionary, the DW-DGS offers valuable insights into the presentation of the results of language data analysis to a wide audience.

Keywords:

German Sign Language
Corpus-based SL Dictionary
Maps
Collocations
User Groups

Link to contribution: [The DW-DGS: A corpus-based dictionary of German Sign Language is being compiled](#)

Diffusion Models for Audio-Visual Speech Enhancement

Author: Julius Richter

Co-author: Timo Gerkmann

Corresponding Authors: julius.richter@uni-hamburg.de, timo.gerkmann@uni-hamburg.de

This poster showcases a selection of our work on diffusion models for speech enhancement. While diffusion models have proven successful in natural image generation, we adopt them for speech enhancement by introducing a task-adapted diffusion process in the complex short-time Fourier domain. Our results show competitive performance compared to strong predictive methods, while generalization is better when evaluated in a mismatched training scenario. However, for very challenging input, the model tends to produce speech-like sounds without semantic meaning. To address this problem, we condition the diffusion model on visual input with the speaker's lips, resulting in improved speech quality and intelligibility. This improvement is reflected in a reduced word error rate of a downstream automatic speech recognition model.

Keywords:

Diffusion models
Speech Enhancement
Audio-Visual
Generative Models

Link to contribution: [Diffusion Models for Audio-Visual Speech Enhancement](#)

Multimodal Modelling of Cultural Artefacts in Digital Space

Authors: Frank Steinicke¹; Julia Nantke¹; Qianqi Huang¹

¹ *Universität Hamburg*

Corresponding Authors: julia.nantke@uni-hamburg.de, frank.steinicke@uni-hamburg.de

The poster presents the cross-disciplinary methods from artificial intelligence and humanities that we plan to apply in the project MuMokA, which deals with the multimodal modeling of cultural artefacts. We apply digital methods from the area of digital restoration, natural language processing, computer vision, and affective computing to extract modalities from the unfinished multimodal work *Ortslinien* by Walter Kempowski, and to annotate and interlink between them automatically. We also explore generative AI methods to develop concepts of autocompletion of the unfinished parts of the work. Additionally, we work on concepts to present multimodal cultural artefacts in a digital, immersive environment in a sustainable way, which takes into account the FAIR and CARE principles for storage and representation of digital data.

Keywords:

multimodality
cultural artefacts
artificial intelligence
natural language processing
computer vision

Link to contribution: [Multimodal Modelling of Cultural Artefacts in Digital Space](#)

In-the-wild Speech Emotion Conversion Using Disentangled Self-Supervised Representations and Neural Vocoder-based Resynthesis

Author: Navin Raj Prabhu¹

Co-authors: Nale Lehmann-Willenbrock ²; Timo Gerkmann

¹ *Signal Processing*

² *Industrial and organizational psychology*

Corresponding Authors: nale.lehmann-willenbrock@uni-hamburg.de, timo.gerkmann@uni-hamburg.de, navin.raj.prabhu@uni-hamburg.de

Speech emotion conversion aims to convert the expressed emotion of a spoken utterance to a target emotion while preserving the lexical information and the speaker's identity. In the context of human-machine interaction systems (e.g., social robots), to improve the naturalness of machine communication, the generation of emotionally expressive speech is required. In this work, we introduce a methodology that uses self-supervised networks to disentangle the lexical, speaker, and emotional content of the utterance, and subsequently uses a HiFiGAN vocoder to resynthesize the disentangled representations to a speech signal of the targeted emotion. Results confirm that the proposed approach is aptly conditioned on the emotional content of input speech and is capable of synthesising natural-sounding speech for a target emotion.

Keywords:

Speech Emotion Conversion
HiFiGAN
Self-supervised representations
Speech Synthesis
Human-machine interaction

Link to contribution: [In-the-wild Speech Emotion Conversion Using Disentangled Self-Supervised Representations and Neural Vocoder-based Resynthesis](#)

Comparative Annotation to Explore and Explain Text Similarities (CompAnno)

Authors: Janis Pagel¹; Julia Nantke²; Marie Flüh²; Nils Reiter¹

¹ *Universität Köln*

² *Universität Hamburg*

Corresponding Authors: nils.reiter@uni-koeln.de, janis.pagel@uni-koeln.de, marie.flueh@uni-hamburg.de, julia.nantke@uni-hamburg.de

Das DFG-Projekt *CompAnno* entwickelt einen vergleichenden Annotationsworkflow zur computergestützten Detektion und Klassifizierung von literarischen Textähnlichkeiten am Beispiel von Figureneigenschaften als einer Kategorie, die sowohl für die Gestaltung literarischer Texte im Allgemeinen als auch für die Interpretation intertextueller Beziehungen zentral ist. Im Gegensatz zu etablierten Annotationsansätzen beruht die vergleichende Annotation auf der gleichzeitigen Betrachtung mehrerer Textausschnitte. Figureneigenschaften werden zunächst manuell ermittelt und kategorisiert. Darauf aufbauend wird ein Ranking der Ähnlichkeiten erstellt, das zum Trainieren eines maschinellen Lernsystems verwendet wird. Wir entwickeln einen Workflow für eine computergestützte Untersuchung von Textähnlichkeit, die über die Erkennung von text-reuse hinausgeht und nicht auf ein festes Korpus bezogen ist. Gleichzeitig zeigen wir mit der vergleichenden Annotation einen neuen Weg für die Arbeit mit interpretativen Kategorien auf.

Keywords:

Annotation

NLP

Maschinelles Lernen

Intertextualität

literarische Figuren

Link to contribution: [Comparative Annotation to Explore and Explain Text Similarities \(CompAnno\)](#)

123 - DisKo –Das Diversitäts Korpus

Authors: Mareike Schumacher; Marie Flüh

Corresponding Authors: mareike.schumacher@sprachlit.uni-regensburg.de,
marie.flueh@uni-hamburg.de

DisKo steht für Diversitäts-Korpus und ist ein literaturwissenschaftliches Projekt mit Digital-Humanities-Komponente. Mit Methoden des maschinellen Lernens trainieren wir einen Machine Learning Classifiers, der nicht nur weibliche, männliche und neutrale Genderrollen in literarischen Texten aus unterschiedlichen Epochen erkennt, sondern auch weniger binäre Genderzuschreibungen. Für den überwachten Trainingsprozess bauen wir in Kooperation mit der Deutschen Nationalbibliothek Frankfurt ein Trainingskorpus auf, das Texte beinhaltet, in denen nicht-binäre Genderzuschreibungen vorkommen. Für die Korpuskonstituierung nutzen wir *Citizen Humanities*. Die manuelle Annotation der Trainingstexte bildet die Grundlage für das Machine Learning. Den projekteigenen Workflow von der Textbeschaffung bis zum Training des Classifiers und dessen Anwendung möchten wir auf der “Digital Total” vorstellen.

Keywords:

Computational Literary Studies
Gender Studies
Citizen Humanities
Machine Learning

Link to contribution: [123 - DisKo –Das Diversitäts Korpus](#)

Extending the Public DGS Corpus in Size and Depth

Authors: Elena Jahn¹; Marc Schulder¹; Reiner Konrad¹; Thomas Hanke¹

¹ *UHH, IDGS*

Corresponding Authors: reiner.konrad@uni-hamburg.de, thomas.hanke@uni-hamburg.de, marc.schulder@uni-hamburg.de, elena.jahn@uni-hamburg.de

In 2018 the DGS-Korpus project published the first full release of the Public DGS Corpus with new data formats, corpus annotation conventions, and OpenPose pose information all transcripts. The community and research portal websites of the corpus also received upgrades, including persistent identifiers, archival copies of previous releases and improvements to their usability on mobile devices. In 2020 a third portal with the corpus research tool ANNIS was launched. Release 4 of the data and further improvements of the three portals was planned for 2022, but postponed to end of 2023.

The poster provides an overview of the content of the Public DGS Corpus and its accessibility and searchability in the three portals.

Keywords:

German Sign Language (DGS), Linguistic Resource, Corpus, Resource Extension

Link to contribution: [Extending the Public DGS Corpus in Size and Depth](#)

Semi-Automatic Subtitle Alignment for Sign Languages

Author: Maren Brumm¹

Co-author: Sam Bigeard ¹

¹ *IDGS Universität Hamburg*

Corresponding Authors: sam.bigeard@uni-hamburg.de, maren.brumm@uni-hamburg.de

To provide subtitles for videos in sign language the translation has to be aligned with the signing.

Doing this manually is time consuming but fully automatic alignment is often not accurate enough. We therefore propose an interactive tool that uses human feedback to optimize the alignment process. For the automatic alignment we use signs detected by a sign spotter and map them to words in the translation. These mappings are used as landmarks for the alignment algorithm. Users can change the alignment directly but also alter, add or delete landmarks to guide the automatic process, so that a correct alignment is reached after a few feedback loops.

Keywords:

Sign Languages
Machine Learning
Natural Language Processing
Interface Design
Semi-Automation

[Link to contribution: Semi-Automatic Subtitle Alignment for Sign Languages](#)

The Multilingual Sign Languages Wordnet

Author: Sam Bigeard¹

Co-authors: Marc Schulder ¹; Maria Kopf ¹; Thomas Hanke ¹

¹ *IDGS, Universität Hamburg*

Corresponding Authors: sam.bigeard@uni-hamburg.de, marc.schulder@uni-hamburg.de

Wordnets are a popular type of lexical resource, used in a variety of computational and linguistic applications. They are appreciated for their sense-based representation of lexical items and their typed relationship network. Wordnets for many different languages exist, but while research has been published toward creating Wordnets for sign languages, our work is the first resulting in a publicly available dataset. It covers 5 different European sign languages including German Sign language, and contains more than 9000 signs. This poster presents the method used to generate this data, a combination of automatic processes and manual validation by signers. The dataset can be browsed online and downloaded at sign-net.meine-dgs.de

Keywords:

natural language processing

resource creation

dataset

wordnet

sign languages

Link to contribution: [The Multilingual Sign Languages Wordnet](#)

The Sign Language Dataset Compendium: Creating an Overview of Digital Sign Language Resources

Authors: Maria Kopf¹; Marc Schulder¹; Thomas Hanke¹

¹ IDGS, Universität Hamburg

Corresponding Author: marc.schulder@uni-hamburg.de

Identifying suitable datasets is a common challenge for data scientists working in domains with scarce data. For research on sign languages, this usually involves extensive literature review or word-of-mouth. Information on individual datasets may be distributed across different publications, data repositories and (potentially defunct) project websites. We introduce the Sign Language Dataset Compendium, an extensive overview of linguistic resources for sign languages. It covers corpora, lexical resources, and commonly used data collection tasks. Special attention is paid to covering many different languages from around the globe. All information is provided in a standardised format to make entries comparable, but kept flexible enough to allow for differences in content. The compendium is a growing resource that is updated regularly.

Keywords:

Sign languages
Corpora
Lexical resources
Survey
Metadata

Link to contribution: [The Sign Language Dataset Compendium: Creating an Overview of Digital Sign Language Resources](#)

KI für die Internationale Raumstation

Authors: Alexander Diedrich¹; Henrik Steude¹; Lukas Moddemann¹; Mark Tappe¹; Oliver Niggemann¹; Stephan Myschik¹

¹ *HSU/UniBw H*

Corresponding Authors: lukas.moddemann@hsu-hh.de, mark.tappe@hsu-hh.de, henrik.steude@hsu-hh.de, alexander.diedrich@hsu-hh.de, oliver.niggemann@hsu-hh.de, stephan.myschik@hsu-hh.de

In der technologisch komplexen Umgebung der Internationalen Raumstation (ISS) ist die schnelle und akkurate Identifikation von Fehlerursachen sowie deren Behebung, insbesondere bei Störungen im Lebenserhaltungssystem, kritisch. Aktuell stellen die Analyse von Daten aus mehr als 20.000 Sensoren und das Verständnis der komplexen Wirkzusammenhänge in der Station eine bedeutende Herausforderung für die Experten in der Bodenstation dar. In diesem Poster stellen wir vor, wie wir durch die Verknüpfung von symbolischer KI und modernen Deep-Learning-Methoden ein vollautomatisiertes Diagnosesystem entwickelt haben, welches sich aktuell in der Validierungsphase befindet. So können automatisch Abweichungen erkannt, Diagnosen berechnet und Rekonfigurationsanweisungen im Prozess durchgeführt werden. Eine im Projekt entwickelte MLOps-Plattform ermöglicht dabei eine effiziente Ressourcennutzung und unterstützt schnelle Iterationen durch den gesamten ML-Entwicklungszyklus.

Keywords:

International Spacestation
Seep Learning
Diagnosis
Time Series

Link to contribution: [KI für die Internationale Raumstation](#)

Akidu - Automatische, durch KI-integrierte Disposition für Universalterminals

Authors: Lothar Hotz¹; Matthias Kerzel¹; Pascal Rost¹; Rainer Herzog¹

¹ HITeC e.V.

Corresponding Authors: pascal.rost@hitec-hamburg.de, matthias.kerzel@hitec-hamburg.de, rainer.herzog@hitec-hamburg.de, lothar.hotz@hitec-hamburg.de

One aim of the Akidu project is to automatically determine the position of goods at the Unikai port terminal in Hamburg. Goods such as unit loads, cranes or other types of goods with typically abnormal dimensions might have been moved without updating their position in the terminal system in order to create free space for upcoming goods. If goods are equipped with markers such as Datamatrix tags or Bluetooth beacons, a drone, flying over the area, could detect and locate them using e.g. image processing methods and then update their position stored in the terminal system, speeding up the search process for specific goods when they are needed.

Keywords:

image processing
localization
marker identification
cargo identification
tracking

Link to contribution: [Akidu - Automatische, durch KI-integrierte Disposition für Universalterminals](#)

DaFne - Platform Data Fusion Generator

Authors: Ulrike Steffens¹; Jörg Rainer Noennig²; Pamela Kunert¹; Ayse Glass²

¹ *HAW Hamburg*

² *HafenCity Universität Hamburg*

Corresponding Authors: ulrike.steffens@haw-hamburg.de, pamela.kunert@haw-hamburg.de, joerg.noennig@hcu-hamburg.de, ayse.glass@hcu-hamburg.de

Der Generierung synthetischer Daten kommt in der Entwicklung von KI-Methoden dort eine große Bedeutung zu, wo Trainingsdaten für diese Methoden nicht ausreichend oder in passender Form verfügbar sind. Die praktische Nutzbarkeit bisher existierender Methoden zur Datengenerierung ist oft eingeschränkt, da sie auf spezifische Einsatzgebiete zugeschnitten sind, ihre Anwendung erhebliche Expertise erfordert oder ihre Anpassung aufwändig ist. Wir stellen hier das Projekt DaFne vor, das die Nutzbarkeit von Methoden zur Datengenerierung für KI-Forscher und -Entwickler durch Zurverfügungstellung einer innovativen flexiblen und erweiterbaren Plattform zur Datengenerierung verbessern will. Hierzu werden auf der Plattform parametrisierbare Algorithmen zur Datengenerierung ebenso wie Evaluationsmethoden für die generierten Daten bereitgestellt. DaFne ist offen für Erweiterungen durch Dritte und wird aktuell in Use-Cases aus dem Gebiet Smart-Cities erprobt.

Keywords:

Datengenerierung
Datenevaluierung
Softwareplattform
Smart Cities

Link to contribution: [DaFne - Platform Data Fusion Generator](#)

AI for Happiness: Pedestrian Path Generation through Agent-Based Simulations with Deep Reinforcement Learning

Authors: Ayse Glass¹; Burak Bek¹; Eylül Kibele Mengeş¹; Iryna Okhrin²; Jörg Rainer Noennig¹; Mariela Rossana Sanchez³; Pramod Baddam³; rene jaekel²; roman glass⁴

¹ *HafenCity University Hamburg*

² *Center for Information Services and High Performance Computing (ZIH) Technische Universität Dresden, Dresden; ScaDS.AI - Center for Scalable Data Analytics and Artificial Intelligence Dresden/Leipzig*

³ *Center for Information Services and High Performance Computing (ZIH) Technische Universität Dresden, Dresden*

⁴ *University Grenoble Alpes, Grenoble, France*

Corresponding Authors: iryna.okhrin@tu-dresden.de, pramod.baddam@tu-dresden.de, joerg.noennig@hcu-hamburg.de, burakbek7@gmail.com, rene.jaekel@tu-dresden.de, roman.glass@gmx.net, mariela_rossana.sanchez_figuerola@tu-dresden.de, eylulkibelemenges@gmail.com, ayse.glass@hcu-hamburg.de

AI integration in Smart Cities, primarily through agent-based simulations, holds transformative potential for understanding and enhancing citizen behavior. Striking a balance between complexity and computational feasibility is essential. Our research question is, how can we make agents behave more realistically? We assumed that happiness is a motivating factor for the mobility. Insights from a survey of 130 citizens inform our weightings. We used reinforcement learning (RL) as a method and Q-learning as an algorithm to generate a baseline, further enhanced with neural networks for adaptability. This study contributes to data-driven urban design by offering efficient intelligent agent solutions. The research lays foundations for smart agents in urban design, which can be used to generate synthetic data.

Keywords:

artificial intelligence, synthetic data, smart cities, deep reinforcement learning, design

Link to contribution: [AI for Happiness: Pedestrian Path Generation through Agent-Based Simulations with Deep Reinforcement Learning](#)

Generative Machine Learning for Particle Physics

Authors: Gregor Kasieczka¹; Thorsten Buss¹

¹ *Institut für Experimentalphysik, Universität Hamburg*

Corresponding Authors: thorsten.buss@uni-hamburg.de, gregor.kasieczka@uni-hamburg.de

To analyze events in particle colliders, a comparable amount of events has to be simulated as events are recorded. The Monte Carlo simulation of the detector needs most of the computing resources. However, the required resources will exceed the available ones soon.

To tackle this problem, we are developing several generative machine learning models in our research group at the Institute for Experimental Physics. We use them as a fast and resource-efficient alternative to Monte Carlo simulation while preserving high accuracy. We develop models for the generation of particle jets and of calorimeter showers. To this end, we employ a number of different architectures such as generative adversarial networks, diffusion-based models as well as discrete and continuous normalizing flows.

Keywords:

particle physics
generative ai

[Link to contribution: Generative Machine Learning for Particle Physics](#)

Anomaly detection for discovery in particle physics

Author: Manuel Sommerhalder¹

Co-authors: Gregor Kasieczka ¹; Louis Moureaux ¹; Tobias Quadfasel ¹

¹ *Universität Hamburg, Institut für Experimentalphysik*

Corresponding Authors: gregor.kasieczka@uni-hamburg.de, tobias.quadfasel@uni-hamburg.de, manuel.sommerhalder@uni-hamburg.de, louis.moureaux@uni-hamburg.de

In their quest to understand the fundamental law governing both the cosmos and interactions between particles, physicists are searching for small deviations from established theories in massive amounts of experimental data. Searches targeting specific signatures have yielded no positive results since the discovery of the Higgs boson, raising the question of whether the right signatures have been searched for. We present ML-based methods to address this problem by performing searches with minimal theory guidance.

Keywords:

particle physics
anomaly detection
weak supervision
normalizing flow

Link to contribution: [Anomaly detection for discovery in particle physics](#)

Fast machine learning in particle physics

Authors: Gregor Kasieczka¹; Karim El Morabit¹; Sven Bollweg²

¹ *Universität Hamburg*

² *Fachbereich Physik, Institut für Experimentalphysik*

Corresponding Authors: sven.bollweg@studium.uni-hamburg.de, karim.el-morabit@desy.de, gregor.kasieczka@uni-hamburg.de

Particle physics studies the nature of elementary particles. There are strong hints that the current theories are not complete yet and that so far undiscovered particles exist. They could be produced in proton collisions at very high energy. To allow significant production of new particles in reasonable time, collisions are performed at 40MHz. Only a small fraction of the data is stored, requiring the use of an online filtering algorithm called trigger. Instead of trigger criteria based on theory predictions, this work focuses on a machine learning based anomaly detection trigger. To fulfill the time constraints and to be compatible with the existing system, the trained model needs to be ported to Field Programmable Gate Arrays.

Keywords:

FPGA

Particle Physics

anomaly detection

machine learning

Link to contribution: [Fast machine learning in particle physics](#)

Inflation Narratives from a Machine Learning Perspective

Authors: Cedric Möller¹; Junbo Huang¹; Max Weinig¹; Ricardo Usbeck²; Ulrich Fritsche¹

¹ *Universität Hamburg*

² *Leuphana Universität Lüneburg*

Corresponding Authors: ulrich.fritsche@uni-hamburg.de, max.weinig@uni-hamburg.de, cedric.moeller@uni-hamburg.de, ricardo.usbeck@uni-hamburg.de, junbo.huang@uni-hamburg.de

Inflation narratives explain inflation changes and affect expectations. Manually identifying them is cumbersome, prompting the need for scalable algorithms. Narratives comprise events, causal relations, and arguments, represented as graphs with event and argument nodes. Causal relations indicate cause-and-effect relationships between events using directed edges. Our main objective is to extract narratives from text to enhance a knowledge graph for analysis like social network analysis or edge prediction. We address two sub-problems: event extraction, involving event type and argument identification, and event deduplication. Second, we extract causal relations as expressed by authors, not necessarily true causal links between events in the text.

Keywords:

Inflation Narratives

Event Extraction

Causal Relations Extraction

Knowledge Graphs

Link to contribution: [Inflation Narratives from a Machine Learning Perspective](#)

Inflation Narratives from an Economics Perspective

Authors: Cedric Möller¹; Junbo Huang¹; Max Weinig¹; Ricardo Usbeck¹; Ulrich Fritsche¹

¹ *Universität Hamburg*

Corresponding Authors: cedric.moeller@uni-hamburg.de, max.weinig@uni-hamburg.de, ricardo.usbeck@uni-hamburg.de, ulrich.fritsche@uni-hamburg.de, junbo.huang@uni-hamburg.de

Previous research in the research group has focused on two main areas: First, building on the work of Andre et al. (2023), Eliaz & Spiegler (2020) and others, a survey has been conducted and narratives have been represented in the form of “directed acyclic graphs” (DAGs, Pearl, 2009). Second, based on the DAGs of Andre et al. (2023), a study of the distribution of narratives in a large text corpus (Dow Jones Newswires) was conducted using a semisupervised topic model (keyATM). The results of the study by Andre et al (2023) were used for the structural identification of the text model.

Keywords:

directed acyclic graphs
large text corpora
topic model

[Link to contribution: Inflation Narratives from an Economics Perspective](#)

SmartShip: KI für Echtzeit-Anomalieerkennung und automatischer Objekterkennung in maritimen Such- und Rettungsaktionen

Authors: Alexander Windmann¹; Lukas Augustin

Co-authors: Jonas Schöttler ; Juan Zhang ; Oliver Niggemann ; Pavle Ivanovic ; Philipp Neumann

¹ *Helmut-Schmidt-Universität*

Corresponding Author: alexander.windmann@hsu-hh.de

In dieser Studie entwickeln wir ein KI-gestütztes System für maritime Such- und Rettungsaktionen. Das System führt in Echtzeit eine Analyse sensorischer Daten und Verhaltensanomalien von Schiffen durch und integriert automatische Objekterkennung zur Identifikation von Personen in Not. Ziel ist eine präzise Vorhersage für Wartungsarbeiten, eine genauere Zustandsüberwachung und effizientere Rettungsmissionen. Unsere Architektur umfasst Software- und Hardwarekomponenten und integriert verschiedene Datenquellen wie Sensor-, Kamera- und AIS-Daten. Durch die Simulation von End-to-End-Datenflüssen validieren wir die Effektivität unseres Ansatzes, der mittels Anomalieerkennung und maschinellem Lernen frühzeitige Motorausfälle erkennen kann.

Keywords:

KI, Search and Rescue, Objekterkennung, Prognostics and Health Management

Link to contribution: [SmartShip: KI für Echtzeit-Anomalieerkennung und automatischer Objekterkennung in maritimen Such- und Rettungsaktionen](#)

Smart Assembly im Labor für intelligente Leichtbauproduktion

Authors: Phillip Johann Overlöper¹; Sven Karras²

Co-authors: Constantin Hildebrandt²; Henrik Eschen²; Jonas Ehrhardt¹; Oliver Niggemann¹

¹ HSU/UniBwH

² CTC GmbH (An Airbus Company)

Corresponding Authors: oliver.niggemann@hsu-hh.de, sven.karras@airbus.com, constantin.hildebrandt@airbus.com, jonas.ehrhardt@hsu-hh.de, henrik.eschen@airbus.com, phillip.overloeper@hsu-hh.de

Im “Labor für intelligente Leichtbauproduktion” (LaiLa) werden in Kooperation zwischen der HSU und der CTC GmbH digitalisierte Produktionssysteme für Bauteile und Großstrukturen aus Faserverbundmaterial in der Luftfahrt vernetzt und durch künstliche Intelligenz (KI) und Maschinelles Lernen (ML) optimiert. Im Use-Case “Smart Assembly” innerhalb von LaiLa kommen intelligente Tools und kollaborative Automatisierungssysteme in Montageprozessen zum Einsatz. Dabei erfassen zahlreiche Sensoren aus elektrischen Handwerkzeugen und Kameras Prozessdaten während der Montage, welche mittels KI-Algorithmen analysiert werden. Dies ermöglicht das automatische Erlernen und Dokumentieren manueller Arbeitsschritte allein durch digitale Beobachtung. Auf Basis dieser erlernten Prozessschritte sollen KIs neue Montageabläufe autonom und effizient planen. Dies fördert die Integration des Menschen in digitale Produktionsprozesse (humanzentriert) und steigert die Montageresilienz sowie die Digitalisierung in sogenannten Brownfield-Umgebungen.

Keywords:

Smart Assembly
Airbus
Artificial Intelligence
Machine Learning
Production Processes

Link to contribution: [Smart Assembly im Labor für intelligente Leichtbauproduktion](#)

hpc.bw –Competence Platform for Software Efficiency and Supercomputing (dtec.bw)

Authors: Alexander Kolling¹; Hauke Preuß¹; Marie Rathmann²; Philipp Neumann¹; Piet Jarmatz¹

¹ *Helmut Schmidt University Hamburg / University of the Federal Armed Forces Hamburg*

² *Deutsch*

Corresponding Authors: marie.rathmann@hsu-hh.de, jarmatz@hsu-hh.de, philipp.neumann@hsu-hh.de, preuss@hsu-hh.de, alexander.kolling@hsu-hh.de

The project „hpc.bw –Competence Platform for Software Efficiency and Supercomputing“ (dtec.bw) aims to strengthen innovative cross-location research in the field of high performance computing (HPC) and to promote the transfer of relevant expertise to a wide range of disciplines. The established HPC Competence Platform (HPCCP) and container-based HPC computing center (CBRZ) are leveraged by performance engineering projects, and trainings are offered to access HPC competencies at different competence levels (e.g. workshops, seminar series, networking activities).

The poster further shows application areas such as AI/ML, numerical simulations, or optimization problems for the CBRZ, HPCCP and a building block system for creating appealing and useful offers for the diverse and interdisciplinary user group.

Keywords:

high performance computing
container-based HPC computing center
Competence Platform

Link to contribution: [hpc.bw –Competence Platform for Software Efficiency and Supercomputing \(dtec.bw\)](#)

EvalSpek-ML

Author: Daniel Boschmann¹

Co-authors: Oliver Niggemann¹; Willi Großmann¹

¹ *HSU/UniBW Hamburg*

Corresponding Authors: daniel.boschmann@hsu-hh.de, oliver.niggemann@hsu-hh.de, grossmaw@hsu-hh.de

Compositional analysis identifies the elemental makeup of materials, with applications including tracing the geological origins of archaeological artifacts. Spectra, encompassing for example 16, 384 distinct energy channels, provide a detailed elemental composition when analyzed by experts. This process is complex and time consuming, therefore EvalSpek-ML aims to use machine learning algorithms to streamline the analysis of those spectra. The project begins by setting a performance benchmark using 1CNNs, Random Forests, and FeedForward-NNs. It then advances to cutting-edge NNs, such as Encoder-Decoder Networks, to enhance the efficacy of foundational ML algorithms. The dataset used to train the algorithms includes expert-evaluated spectra along with artificially generated data. Further, this approach will be generalized to map other types of spectra.

Keywords:

Machine Learning

Deep Learning

Spectra

Compositional Analysis

Project

Link to contribution: [EvalSpek-ML](#)

Data-based tool try out in sheet metal forming

Author: Michael Hohmann¹

Co-authors: Adili Yiming ²; Lars Penter ²; Oliver Niggemann ¹; Steffen Ihlenfeldt ²

¹ *HSU/UniBw Hamburg*

² *TU Dresden*

Corresponding Author: michael.hohmann@hsu-hh.de

In sheet metal forming, the development cost and complexity of the tool try out is a major challenge. Data-based tool try out focuses on two main areas: the interpretation of die spotting images as indicators of tool quality, and the generation of tool designs based on the parts to be formed. In addition to the part geometry, various physical conditions are taken into account. Limited data availability increases the complexity of the task. As a solution, generative models such as CVAE or cGAN will be included, which can be pre-trained with simulations. For the analysis of the die spotting images, computer vision techniques like CNNs are applied to evaluate the relationship between the images and the excess material.

Keywords:

generative model

computer vision

3d-design

smart manufacturing

Link to contribution: [Data-based tool try out in sheet metal forming](#)

Morphology-based Molecular Classification of Spinal Cord Ependymomas Using Deep Neural Networks

Author: Yannis Schumann¹

Co-authors: Julia Neumann²; Leonille Schweizer³; Maja Krech⁴; Matthias Dottermusch²; Philipp Neumann¹; Tasja Lempertz²; Ulrich Schüller⁵

¹ *Helmut-Schmidt-Universität/Universität der Bundeswehr Hamburg - Chair for High Performance Computing*

² *University Medical Center Hamburg-Eppendorf (UKE) - Center for Molecular Neurobiology Hamburg (ZMNH)*

³ *University Hospital Frankfurt, Goethe University - Institute for Neurology*

⁴ *Charité Berlin - Institute for Neuropathology*

⁵ *University Medical Center Hamburg-Eppendorf (UKE) - Neuropathology Department*

Corresponding Authors: philipp.neumann@hsu-hh.de, ju.neumann@uke.de, schumany@hsu-hh.de

Introduction

Ependymomas are tumors of the central nervous system and DNA methylation profiling is used to distinguish them into 10 molecular types. However, traditional assessment by neuropathologists of tumor tissue stained with hematoxylin and eosin (H&E) does often not match to the correct molecular group. Focusing on the molecular ependymoma types myxopapillary (MPE) and spinal (SP-EPN), we resolve the relation of morphological and DNA-methylation patterns by employing deep-learning techniques to predict the molecular type from H&E-stained whole-slide images.

Computational Methodology and Results

Siamese representation learning and attention-based multiple instance learning is used to classify ependymomas using high-resolution images with hundreds of gigapixels. We demonstrate accurate prediction (98% test accuracy) and discuss relevant challenges for the application in diagnostics.

Keywords:

Ependymoma
Neural Networks

Image Classification
Self-Supervised Learning
Multiple Instance Learning

Link to contribution: [Morphology-based Molecular Classification of Spinal Cord Ependy-
momas Using Deep Neural Networks](#)

Understanding and Improving Large Language Models

Author: Michaela Regneri

Co-authors: Alhassan Abdelhalim ; Merle Mosch ; Sören Laue ¹

¹ *Universität Hamburg*

Corresponding Authors: merle.mosch@studium.uni-hamburg.de,
michaela.regneri@uni-hamburg.de, soeren.laue@uni-hamburg.de,
alhassan.ahmed.said.abdelhalim@studium.uni-hamburg.de

The Machine Learning Group at UHH researches the core mechanisms behind large language models (LLMs). With the rise of ChatGPT, LLMs moved to the focus of public interest. Many research projects analyse the application of LLMs, while comparably little work sheds light on the algorithmic and mathematical foundations behind training, reasoning and abstraction capabilities. We will advance LLM predictability, and provide theoretically sound explanations of their reasoning mechanisms.

The results will make LLMs more transparent, and thus directly help to predict and avoid bias, to make training much more data-efficient, and to compare the reasoning in LLMs with the language understanding in humans.

The project aligns with our commitment to providing more efficient, explainable and transparent algorithms for artificial intelligence.

Keywords:

Large Language Models

Transformers

Explainable AI

Link to contribution: [Understanding and Improving Large Language Models](#)

Machine Learning Group at Universität Hamburg

Author: Sören Laue¹

Co-authors: Michaela Regneri ; Tomislav Prusina

¹ *Universität Hamburg*

Corresponding Authors: soeren.laue@uni-hamburg.de, tomislav.prusina@uni-hamburg.de, michaela.regneri@uni-hamburg.de

The Machine Learning Group at Universität Hamburg is at the forefront of research in the field of efficient machine learning. Our specialized focus lies in developing cutting-edge optimization methods that enhance the training of machine learning models. Leveraging innovative techniques and algorithms, our group strives to make machine learning more efficient, scalable, and sustainable. We are dedicated to addressing the computational challenges that arise in modern machine learning, allowing for faster convergence and reduced resource consumption. Our research aims to empower both academia and industry, fostering advancements that pave the way for more accessible and resource-efficient machine learning applications. We will present some results and tools that we have developed within this area.

Keywords:

machine learning
efficient algorithms
resource-efficient learning
efficient optimization methods

Link to contribution: [Machine Learning Group at Universität Hamburg](#)

Multilingual Racial Hate Speech Detection Using Transfer Learning

Authors: Abinew Ali Ayele¹; Chris Biemann¹; Seid Muhie Yimam²; Skadi Dinter¹

¹ *Language Technilogy Group, University of Hamburg, Hamburg, Germany*

² *House of Computing and Data Sceince*

Corresponding Authors: chris.biemann@uni-hamburg.de, skadi.dinter@posteo.de, abinew.ali.ayeale@uni-hamburg.de, seid.muhie.yimam@uni-hamburg.de

The rise of social media eases the spread of hateful content, especially racist content with severe consequences. In this paper, we analyze the tweets written in French targeting the death of George Floyd in May 2020 as the event accelerated debates on racism globally. Using the Yandex Toloka platform, we annotate the tweets into categories as hate, offensive, or normal. Tweets that are offensive or hateful are further annotated as racial or non-racial. We build French hate speech detection models based on the multilingual BERT and CamemBERT and apply transfer learning by fine-tuning the HateXplain model. We compare different approaches to resolve annotation ties and find that the detection model based on CamemBERT yields the best results in our experiments.

Keywords:

Racial hate speech, offensive speech, transfer learning, Toloka

Link to contribution: [Multilingual Racial Hate Speech Detection Using Transfer Learning](#)

German Large Language Legal Model (GLLLM)

Authors: Chris Biemann¹; Fereidoun Rashidi¹; Martin Semmann¹; Seid Muhie Yimam¹

¹ *House of Computing and Data Science*

Corresponding Authors: seid.muhie.yimam@uni-hamburg.de, fereidoun.rashidi@uni-hamburg.de, chris.biemann@uni-hamburg.de, martin.semmann@uni-hamburg.de

Our project's primary focus is to systematically collect legal documents in German from European Union administrative tiers. This effort enables seamless access to German legislation for researchers, lawyers, and citizens alike. Using intuitive Python scripts, we've automated data collection from publicly accessible websites, streamlining the process for users. Furthermore, we employ fine-tuning techniques on models like BERT and GPT to enhance their ability to comprehend and extract information from German legal texts. Our initiative is driven by the conspicuous absence of a comprehensive German legal document database. By creating this valuable and versatile resource, we aim to bridge this gap, catering to a diverse range of users, including the general public, researchers, and legal professionals.

Keywords:

Legal documents
German legislation
LLLM
Fine-tuning for legal documents

Link to contribution: [German Large Language Legal Model \(GLLLM\)](#)

ARDIAS: AI-Enhanced Research Discovery and Advisory System

Authors: Aditya Jogalekar; Chris Biemann¹; Debayan Banerjee²; Martin Semmann¹; Seid Muhie Yimam¹

¹ *House of Computing and Data Science*

² *Universität Hamburg*

Corresponding Authors: chris.biemann@uni-hamburg.de, seid.muhie.yimam@uni-hamburg.de, debayan.banerjee@uni-hamburg.de, martin.semman@uni-hamburg.de, adiitya08@gmail.com

ARDIAS is a web-based application that aims to provide researchers with a suite of discovery, collaboration, and recommendation tools. ARDIAS allows searching for authors and articles by name and gaining insights into the research topics of a particular researcher. It also presents a conversational interface where users can chat about a particular author or paper. Lastly, it implements a “Community” feature where a group of like-minded researchers can come together and share ideas or organize events.

Keywords:

research
scholarly
AI
NLP
chatgpt

Link to contribution: [ARDIAS: AI-Enhanced Research Discovery and Advisory System](#)

House of Computing and Data Science

Authors: Chris Biemann¹; Katrin Schönig-Stierand²; Martin Semmann³; Seid Muhie Yimam⁴

¹ *House of Computing & Data Science*

² *House of Computing and Data Science (HCDS)*

³ *House of Computing and Data Science, Universität Hamburg*

⁴ *House of Computing and Data Science*

Corresponding Authors: martin.semmann@uni-hamburg.de, seid.muhie.yimam@uni-hamburg.de, chris.biemann@uni-hamburg.de, katrin.schoening-stierand@uni-hamburg.de

As a central institution of the University of Hamburg, the House of Computing and Data Science (HCDS) supports interdisciplinary research and application of innovative digital methods in close cooperation with its partners from science and industry in the Hamburg metropolitan region. It coordinates and supports the implementation of the digital strategy in research at the University of Hamburg. Located in the Science City Bahrenfeld and at various other locations in Hamburg, it fuels the easy adoption, usage and research of digital methods in its Methodology Competence Center and offers various disciplines and projects a forum for the exchange of information and collaboration at the interface between methodological sciences and applied sciences in the Cross-Disciplinary Labs.

The HCDS is coordinating the network “KIEZ of Computing and Data Science”, a platform for interconnectivity in all areas of data science and computation within the metropolitan region of Hamburg.

Keywords:

Digitization

Digital Transformation

Artificial Intelligence

Service Center

Link to contribution: [House of Computing and Data Science](#)

Transformation of ERIS Information System to Advanced Search

Authors: Haruka Inoba; Seid Muhie Yimam¹; Martin Semmann²; Chris Biemann³

Co-authors: Werner Rieß⁴; Justine Diemke⁴; Ahmet Yilmaz⁴

¹ *House of Computing and Data Science*

² *House of Computing and Data Science*

³ *House of Computing & Data Science*

⁴ *Universität Hamburg*

Corresponding Authors: chris.biemann@uni-hamburg.de, haruka.inoba@studium.uni-hamburg.de, seid.muhie.yimam@uni-hamburg.de, martin.semmann@uni-hamburg.de

Our project focuses on developing the advanced search for ERIS information system which has been operated since 2012 by the Ancient History department at the University of Hamburg.

Since the current ERIS information system shows the database in text format, we aim to support historical researchers in finding hidden information in ancient biographical data by visualizing the data in charts.

Applying the Elastic Search, we enable the user to retrieve data and visualize them in a variety of charts just in a few seconds. Furthermore, the user can dynamically change the visualizations, filtering all data by the free text and keywords.

Keywords:

Visualization, Elastic Search, ERIS Information System,

Link to contribution: [Transformation of ERIS Information System to Advanced Search](#)

D-WISE - Digitale Wissenssoziologische Diskursanalyse

Authors: Liv Ohlsen¹; Tim Fischer¹

Co-authors: Chris Biemann²; Florian Schneider¹; Fynn Petersen-Frey¹; Gertraud Koch¹; Isabel Eiser¹

¹ *Universität Hamburg*

² *House of Computing & Data Science*

Corresponding Authors: florian.schneider-1@uni-hamburg.de, chris.biemann@uni-hamburg.de, liv.winberg.ohlsen@studium.uni-hamburg.de, fynn.petersen-frey@uni-hamburg.de, tim.fischer@uni-hamburg.de, gertraud.koch@uni-hamburg.de, isabel.eiser@uni-hamburg.de

The D-WISE Tool Suite (DWTS) is a web-based working environment for digital qualitative discourse analysis in the Digital Humanities (DH). It addresses the limitations of current DH tools induced by the ever-increasing amount of heterogeneous, unstructured, and multi-modal data in which the discourses of contemporary societies are encoded. To provide meaningful insights from such data, the system combines state-of-the-art machine learning technologies.

The tool is conceived and developed within the D-WISE project in close co-creation by an interdisciplinary team of cultural anthropologists and computer scientists to ensure the tool's usability for modern DH research.

Central features include: a) import of multi-modal data b) ML-based preprocessing pipelines c) lexical and semantic search d) manual annotations e) documentation of the research process.

Keywords:

D-WISE
Digital Humanities
Natural Language Processing
Discourse Analysis

Link to contribution: [D-WISE - Digitale Wissenssoziologische Diskursanalyse](#)

FUNDus meets HCDS - Novel avenues to explore scientific collections

Authors: Florian Schneider¹; Iris Vogel¹; Martin Semmann²

¹ *Universität Hamburg*

² *House of Computing and Data Science*

Corresponding Authors: martin.semmann@uni-hamburg.de, iris.vogel@uni-hamburg.de

Die Sichtbarmachung von derzeit insgesamt 31745 Objekten aus 59 Sammlungen mit 32404 Digitalisaten wird durch die enge Zusammenarbeit des Zentrums für nachhaltiges Forschungsdatenmanagement (ZFDm) mit der Zentralstelle für wissenschaftliche Sammlungen und dem LIB (ehemals Centrum für Naturkunde) ermöglicht.

In Kooperation mit dem House of Computing and Data Science (HCDS) arbeiten wir daran, die Funktionalität des Portals um innovative digitale Methoden zur Exploration und Qualitätssicherung zu erweitern.

Keywords:

scientific collection
data science
artificial intelligence
user-centricity

Link to contribution: [FUNDus meets HCDS - Novel avenues to explore scientific collections](#)

RESCUE-MATE - Rethinking Innovation in Natural Disasters

Authors: Heiko Bornholdt¹; Julia Bräker¹; Martin Semmann²; Mathias Fischer¹

¹ *Universität Hamburg*

² *House of Computing and Data Science*

Corresponding Authors: martin.semmann@uni-hamburg.de, heiko.bornholdt@uni-hamburg.de, mathias.fischer@uni-hamburg.de, julia.braeker@uni-hamburg.de

A consequence of drastic climate change is increased extreme weather events. One of those are storm surges. Hamburg, Germany, is well-known for its dramatic storm surges in 1962, with over 280 deaths. Since then, the measures taken to prevent disasters and cope with the weather have matured. Nevertheless, the application of state-of-the-art information systems falls short. The project RescueMate seeks to implement a platform that enables the aggregation of publicly available data, sensor data and governmental sources. This platform allows the distributed use of artificial intelligence to support the needs of a plethora of actors that aim to improve safety throughout the disaster. Such instances are law enforcement, firefighters, ministries, or civil society like the red cross and others.

Keywords:

natural disaster
climate change
artificial intelligence
IoT

Link to contribution: [RESCUE-MATE - Rethinking Innovation in Natural Disasters](#)

Mikropolis - Orientierung in der digitalen Transformation

Authors: Arno Rolf¹; Janick Edinger¹; Larissa Gebken¹; Lothar Hotz¹; Martin Semmann²

¹ *Universität Hamburg*

² *House of Computing and Data Science*

Corresponding Authors: martin.semmann@uni-hamburg.de, janick.edinger@uni-hamburg.de, arno.rolf@uni-hamburg.de, lothar.hotz@uni-hamburg.de

Mikropolis versteht sich als Digital Literacy-Projekt mit der Zielsetzung, Studierenden Orientierung und Urteilsfähigkeit über den Prozess der digitalen Transformation zu vermitteln und dieses Lernziel mit der nachhaltigen Gestaltung der Digitalisierung (sustainable development Goals/ESG-Kriterien) zu verbinden.

Zielsetzung ist, der Transfer der Plattform MikroPolis.org sowie die inhaltliche, didaktische und IT-gestützte Weiterentwicklung für ein Studium Generale-Studienangebot für eine breite Studierendenschaft aus unterschiedlichen Fächern und Fakultäten zu schaffen. Die transdisziplinäre Ausrichtung bezieht ethische, soziale, ökologische, informatische und ökonomische Inhalte ein.

Keywords:

digital transformation
digital literacy
data literacy
sustainability

[Link to contribution: Mikropolis - Orientierung in der digitalen Transformation](#)

InteGreatDrones: New Perspectives on Ports Through Drones and AI

Authors: André Kelm¹; Helen Schloh¹; Janick Edinger¹; Liviana Franke¹; Simone Frintrop¹

¹ *Universität Hamburg*

Corresponding Authors: helen.schloh@uni-hamburg.de, andre.kelm@uni-hamburg.de, janick.edinger@uni-hamburg.de, simone.frintrop@uni-hamburg.de, liviana.franke@studium.uni-hamburg.de

The rise in production of cost-efficient unmanned aerial vehicles (“UAVs” or “drones”) opens up new possibilities for process optimizations in industrial settings. The InteGreatDrones project focuses on the use of drone swarms for data collection in dynamic inland terminals, addressing challenges of changing cargo types and stakeholder interactions. The project’s goal is to employ autonomous drones to capture terminal activities, ensuring privacy through local data processing and selective sharing. With the help of computer vision algorithms, a digital twin of the port is created. A central middleware connects drones and existing systems, supporting terminal-specific applications like route optimization and predictive maintenance. The project seeks to provide real-time data for informed operational and strategic decisions.

Keywords:

Computer Vision
Digital Twins
Distributed Systems
Industrial IoT
Drones

Link to contribution: [InteGreatDrones: New Perspectives on Ports Through Drones and AI](#)

Group dynamics in the Metaverse: A conceptual model and first empirical insights

Author: Marvin Grabowski¹

Co-authors: Frank Steinicke ²; Anita Blanchard ³; Sebastian Rings ²; Nale Lehmann-Willenbrock ⁴

¹ *PhD Candidate, Industrial & Organizational Psychology*

² *Universität Hamburg*

³ *University of North Carolina at Charlotte*

⁴ *Industrial and organizational psychology*

Corresponding Authors: frank.steinicke@uni-hamburg.de, sebastian.rings@uni-hamburg.de, nale.lehmann-willenbrock@uni-hamburg.de, marvin.grabowski@uni-hamburg.de, anita.blanchard@charlotte.edu

The metaverse offers new technological affordances to conduct 3D immersive meetings with head-mounted displays that can enrich virtual teamwork. We present a conceptual model of effective group interactions in the metaverse, along with a pilot study that begins to explore relevant design factors, attendee experiences, and behavioral group dynamics. Our findings show that participants intensely engaged in immersive meetings, despite having no prior virtual reality experience. Moreover, we identified positive linkages between the observed interaction flow, emergent group entitativity, and overall satisfaction with the immersive meeting. We discuss implications for future investigations of group dynamics in the metaverse.

Keywords:

metaverse

meetings

group dynamics

entitativity

interaction flow

Link to contribution: [Group dynamics in the Metaverse: A conceptual model and first empirical insights](#)

Advanced mathematical solutions for circulating cardiovascular biomolecules

Authors: Benedikt Schrage; Christoph Al-Taie; Julius Obergassel; Olga Tsoy; Olga Zolotareva; Silke Van Elferen

Corresponding Author: b.schrage@uke.de

Cardiovascular biomolecules contain quantifiable information about pathophysiologic processes and are thus frequently used in modern medicine. However, the magnitude of targets, their biological collinearity and their time-dependent variation constitute challenges for conventional statistical models, so that new mathematical and computational solutions are needed. Within our Cross-Disciplinary Lab we aim to create an interdisciplinary sandpit, where researchers from experimental and computational biology develop and test ideas together to push the boundaries of biomolecule research and to advance our understanding of cardiovascular diseases. We work with the unique dataset that includes biomolecule data from 3 clinical trials involving >5.000 patients and predict incident/worsening heart failure in atrial fibrillation patients, ultimately identifying those with the highest need and largest benefit from earlier therapy.

Keywords:

CDL
biomolecules
atrial fibrillation
heart failure

Link to contribution: [Advanced mathematical solutions for circulating cardiovascular biomolecules](#)

FLIGHT: Federated LearnIng Guided digital Health

Authors: Jens Lohmann¹; Olga Zolotareva¹; Guido Sauter²; Stefan Bonn³; Jan Baumbach¹

¹ *Institute for Computational Systems Biology*

² *Institute of Pathology*

³ *Institute of Medical Systems Biology*

Around 60,000 men are annually diagnosed with prostate cancer in Germany, which makes it the second-most frequent cancer. We previously developed the state-of-the-art analysis method eCaReNet (explainable Cancer Relapse prediction Network) for survival prediction of prostate cancer patients based on tissue microarray (TMA) data. To build a more robust and accurate model, data from multiple study sites can be combined and used for training, but this poses serious privacy risks to patient-derived data. To enable model training on distributed TMA data while minimizing privacy risks, we are developing FLIGHT, a privacy-aware version of eCaReNet, protecting patient-derived data with the use of Federated Learning (FL) and Secure Multi-Party Computation (SMPC) techniques.

Keywords:

federated learning
prostate cancer
artificial intelligence

Link to contribution: [FLIGHT: Federated LearnIng Guided digital Health](#)

Sustainable data management @ Tamilex

Authors: Charles Li¹; Eva Wilden¹

¹ *Universität Hamburg*

Corresponding Author: charles.li@uni-hamburg.de

The Tamilex project, funded by a long-term grant from the Akademie der Wissenschaften in Hamburg, aims to create the first historical dictionary of Classical Tamil. It comprises an online corpus of annotated texts, linked to manuscript images, which provides the material evidence for lexical entries. Each of these layers requires careful consideration in order to ensure that the data is preserved and managed sustainably for generations to come, and that it remains accessible for future scholars to use and re-use in their research projects.

Traditionally, Classical Tamil texts were transmitted by copying onto palm leaves. Data management, then, consisted of making new copies before the old ones were destroyed by the heat, humidity, and insects of south India. But each time a text was copied, it would be slightly different, whether due to scribal errors or intentional revisions or corrections. Now, we aim to critically edit these texts so that their evolution becomes apparent, while preserving the rich textual variation that has been accumulated over almost two millennia. Earlier editors may have ignored these variations, selecting or correcting one preferred version of the text as definitive, and thus inadvertently making the same mistakes as the manuscript scribes of the past. When dictionaries have been compiled on the basis of these editions, this has led to appearance of lexical phantoms — words that do not exist except as faulty conjectures on the part of an editor. These conjectures have now become part of the history of the text and its language as well. Today, our challenge is to collect and analyze all of this complex data in a way that is transparent, reusable, and traceable back to its material sources.

Keywords:

Digital Humanities
Cultural Heritage
Manuscripts
Lexicography
Philology

Link to contribution: [Sustainable data management @ Tamilex](#)

Research in Geoinformatics and Geovisualization at HCU

Author: Jochen Schiewe¹

¹ *HAfenCity Universität Hamburg*

Corresponding Author: jochen.schiewe@hcu-hamburg.de

Research at the Lab for Geoinformatics and Geovisualization (g2lab) is about

- visualizing spatio-temporal data to enable the understanding of complex problems and decision making
- applying task-oriented approaches to generate effective and efficient visualizations
- modeling and communicating uncertainties to make decisions more certain

This poster shows selected research and PhD projects of g2lab.

Keywords:

Geoinformatics
Geovisualization
Cartography
Deep Learning
Generlaization

Link to contribution: [Research in Geoinformatics and Geovisualization at HCU](#)

Index

- (Partial) Autocorrelation, 39
- 3d-design, 149
- Accelerometer-data, 110
- accessibility, reproducibility, usability of
biomedical AI models, 122
- Artificial Intelligence, Super Resolution,
Inpainting, Climate Modeling,
Observations, 77
- Aging, 10
- AI, 44, 110, 156
- Airbus, 146
- Alternative splicing, 111
- alternative splicing, 101
- ancient history, 106
- Annotation, 130
- annotation, 19
- anomaly detection, 141, 142
- antibiotic resistance, 29
- App Development, 57
- Artificial Intelligence, 146
- artificial intelligence, 128
- Artificial Intelligence, 38, 157
- artificial intelligence, 41, 160, 161,
166
- artificial intelligence, synthetic data,
smart cities, deep
reinforcement learning, design,
139
- atrial fibrillation, 165
- Audio-Visual, 127
- Augmented Reality, 13
- Autoimmune diseases, , 31
- avatar animation, 66
- Batch Effect Correction, Federated
Learning, FedComBat, Data
Privacy, Decentralized
Analysis, 104
- beam control, 88
- biclustering, 102
- Big Data, 20, 81
- Bioinformatics, 25
- bioinformatics, 26, 30, 101
- biomedicine, machine learning, AI,
Omics, Image processing, EHR,
74
- biomolecular data, 30
- biomolecules, 165
- cargo identification, 137
- Cartography, 168
- Causal AI, 44
- Causal Inference, 42
- Causal ML, 44
- Causal Relations Extraction, 143
- Causality, 42
- CDCS, 105
- CDL, 165
- chatgpt, 156
- chemical fragment spaces, 35
- Chemical Spaces, 33
- Cheminformatics, 33
- chemoinformatics, 35
- Chemometrics, 25
- chemometrics, 26
- Citizen Humanities, 131
- climate change, 161
- Climate Data Operators, 65
- climate data, data publication, quality
assessment, scientific user
support, 95
- climate protest, 79
- Clinical Lab Data, 39
- Cloud-Based Data Evaluation, 45

- clustering, 102
- Co-creation, 57
- Co-expression Networks, , 109
- collaboration, 71
- collaborative learning, 67
- Collaborative Research, 82, 105
- Collocations, 126
- Competence Platform, 147
- Complex diseases, 31
- complex systems, 98
- Compositional Analysis, 148
- Computational Literary Studies, 131
- Computational methods, 80
- computational methods, 63
- computational modeling, 92
- computational systems biology, 101
- ComputationalMolecularDesign, 32
- computer science, 71
- Computer Vision, 163
- computer vision, 128, 149
- container-based HPC computing center, 147
- content analysis, 19, 79
- Corpora, 135
- Corpus Linguistics, 85
- corpus query and visualization tools, 125
- Corpus-based SL Dictionary, 126
- Corrosion, 117
- corrosion, 115
- Corrosion Science, 116
- COVID-19, 43
- CPS, 69
- CRIS, 15
- Critical infrastructures, 48
- Crystallography, 93
- cultural artefacts, 128
- Cultural Heritage, 167
- cultural heritage, 70
- Cyber-physical systems, , 48

- D-WISE, 159
- Data Analysis, 73
- data analysis, 55
- Data Literacy, 42, 78
- data literacy, 162
- data mining, 21, 51

- Data processing, 65
- data reduction, 55
- Data Science, 85, 105
- data science, 160
- Data Science, , 34
- Data Science, Machine Learning, Information Geometry, Embodied Intelligence, Causal Inference, 86
- Data Stream Processing, 81
- data sustainability, 64
- data visualization, 106
- data, climate, sustainability, digital tools, visualisation, 54
- DataScience, 32
- dataset, 134
- Datenevaluierung, 138
- Datengenerierung, 138
- Datenprojekte, 78
- decentralisation, 51
- Deep Learning, 93, 148, 168
- Deep learning, , 31
- Design of Experiments, 117
- design science research, 41
- Diagnosis, 136
- Dictionary, 124
- Differential Gene Expression, 90
- Differential networking, , 109
- Differential Protein Expression Analysis, 100
- Differential Regulation Analysis, , 109
- Diffraction, 37
- Diffusion models, 127
- Digital Data Handling, 18
- Digital Humanities, 84, 159, 167
- digital humanities, 125
- digital humanities, digital edition, manuscript, 49
- Digital Literacy, 78, 84
- digital literacy, 162
- Digital Research, 84
- Digital Scholarship Services, 84
- digital traces, 21
- Digital Transformation, 157
- digital transformation, 162
- digital twin, 67, 88
- Digital Twins, 163
- Digitale Spurendaten, 20

- Digitalisierung in der Hausarztpraxis, 40
- Digitalization, Financial Performance, Ecological Performance, Organizational Support, Socio-Technical Systems Theory, 28
- digitisation, 70
- Digitization, 157
- directed acyclic graphs, 144
- Discourse Analysis, 159
- Disease, 10
- disease heterogeneity, 103
- disease mechanism elucidation, , 109
- disease module identification, 121
- Distributed Systems, 163
- Distribution maps, 124
- Drones, 163
- Drug Design, 33
- Drug discovery, 10
- drug discovery, 35
- drug repurposing, 121
- Drug response prediction, Feature reduction, Transcription factor activity, 119
- Drug targets, 10
- DrugDesign, 32
- Eco-Friendly Technology, 45
- efficient algorithms, 153
- efficient optimization methods, 153
- Electronic Lab Notebook (ELN), 52
- embodied intelligence, 87
- entitativity , 164
- Ependymoma, 150
- Epigenetics, 10
- Evaluation , 89
- Event Extraction, 143
- Evolutionary Algorithm, 33
- exercise therapy, 99
- Exon-exon interaction, 111
- Explainable AI, 152
- FAIR, 60, 108
- fair data, 14
- FAIR principles, 59
- fdb, 56
- Federated Learning, 81, 90, 100
- federated learning, 166
- Federated Learning , 91
- few-shot, 19
- filter bubbles, 63
- Fine-tuning for legal documents, 155
- FIS, 15
- Forschungsethik, 20
- Forschungsinfrastruktur, 20
- FPGA, 142
- free-electron laser, 88
- Gender Studies, 131
- Generative AI, 24
- generative AI, 41
- generative ai, 140
- generative model, 149
- Generative Models, 127
- Generlaization, 168
- genetic variants, 29
- Genetics, , 31
- Geoinformatics, 168
- Geovisualization, 168
- German legislation, 155
- German Sign Language, 124, 126
- German Sign Language (DGS), Linguistic Resource, Corpus, Resource Extension, 132
- GitHub, 82
- Graduate School, , 34
- GRIB, 65
- group dynamics, 164
- group profile, 30
- GWAS, privacy-aware, 112
- Handwriting Style Analysis, 17
- Harbour, 13
- heart failure, 165
- HiFiGAN, 129
- High performance computing, 65
- high performance computing, 147
- high-performance computing, 75, 76
- High-throughput processing, 37
- higher-order networks, 98
- Historical Document Analysis, 17
- Historical Documents Analysis, 16, 23, 24
- HPC, 56, 83

- human-AI collaboration, 41
- Human-machine interaction, 129
- Human-Robot-Interaction, 12
- hyperbolic learning, dimensionality reduction, 114
- hypergraphs, 98
- ICON, 56
- Image Classification, 151
- image processing, 137
- Image Reconstruction, Image Analysis, AI in Imaging, 61
- Industrial IoT, 163
- infectious disease, 11
- Inflation Narratives, 143
- Informatics, , 34
- information theory, 87
- Inpainting, 24
- interaction flow, 164
- interdisciplinary research, 34
- Interface Design, 133
- International Spacestation, 136
- Internet Measurement, 73
- Intertextualität, 130
- Inverse problems, image reconstruction, machine learning, neural networks, robust training., 62
- IoT, 161
- Journalism, 80
- journalism, 79
- Jupyter, 83
- KI, Search and Rescue, Objekterkennung, Prognostics and Health Management, 145
- knowledge distillation, 118
- knowledge graph, 11, 121
- Knowledge Graphs, 143
- language corpora, , 64
- Large Language Models, 152
- large language models, 19
- large text corpora, 144
- Law, 89
- Legal Data Science, 89
- Legal documents, 155
- legal networks, 98
- Lehrinnovationen, 78
- Lexical resources, 135
- Lexicography, 167
- Library, 84
- literarische Figuren, 130
- LLLM, 155
- localization, 137
- lung, 30
- Machine Learning, 25, 44, 81, 94, 116, 117, 131, 133, 146, 148
- Machine learning, 43
- machine learning, 26, 75, 76, 87, 88, 115, 142, 153
- Machine learning, , 31
- Magnesium, 117
- Manuscripts, 167
- Maps, 126
- marker identification, 137
- Maschinelles Lernen, 130
- mastodon, 51
- Mathematics, , 34
- media effects, 63
- medical imaging, 76
- meetings, 164
- mesh-based methods, 75
- Metadata, 135
- metalanguage, 124
- Metallic Biomaterials, 52
- metaverse, 164
- microbial genome-wide association study , 29
- minimal self, 87
- Mixed Methods, 85
- modeling, 69
- molecular dynamics, cfd, hpc, multiscale, 46
- Molecular Optimization, 33
- monitoring, 69
- motion capture, 66
- motion editing, 66
- mpox, 11
- multi-agent systems, 67
- multimodality, 128
- Multimorbidität, 40
- Multiphysics Simulation, 36
- Multiple Instance Learning, 151
- Multiple Sclerosis, 53
- named entity recognition, 106
- natural desaster, 161

- Natural Language Processing, 133, 159
- natural language processing, 19, 128, 134
- Natural Language Processing , 89
- Natural Sciences, , 34
- NetCDF, 65
- network analysis, 79
- network medicine, 121
- network sampling, 21
- Neural Networks, 150
- Neurodegeneration, 53
- neuroimaging, 60
- Neurology, 110
- News recommender systems, 80
- NFDI, digital infrastructure, digital humanities, 50
- NLP, 130, 156
- normalizing flow, 141
- numerical methods, 76
- Object Detection, 23
- ocean modeling, 75
- OCR, 70
- OLR, 70
- Omics, 53
- omics, 103
- Ontology, 52
- open access, open science, networking, scholarly communication, publishing, EU projects, 97
- open access, publication service, open science, quality assurance, standards, policies, 96
- open data, 60
- open science, 14
- Open Source, 42
- OPTical Mirror Design, 68
- Optimization, 68
- Osteoarthritis, 91
- osteoarthritis, , 99
- pain Intensity, , 99
- Palimpsests, 24
- paralinguistics, 118
- Parkinson, 110
- Participatory Design, 57
- Particle Physics, 142
- particle physics, 140, 141
- patent landscaping, 9
- patent profiling, 9
- patent trend annotation, 9
- patient stratification, 102
- Patientenzentrierung, 40
- Pattern Analysis, 16
- PEMT, 9
- Philology, 167
- physics, 59
- physics-based modeling, 76
- Post-quantum resistance, , 48
- prediction, 29, 69
- prediction model, 91
- prediction, , 99
- Privacy-Preserving learning, 90
- Process Monitoring, Machine Learning, Drilling, Aircraft Assembly, 58
- Production Processes, 146
- Project, 148
- prostate cancer, 166
- Protein 3D structures, 120
- Protein structural classification, 120
- Protein structure networks, 120
- protein structure prediction, 92
- Protein-protein interaction, 111
- ProteinStructures, 32
- Proteomics Data Privacy, 100
- Pulse Shaping, 68
- PUNCH4NFDI, 59
- quality assurance, 60
- quality assurance, , 64
- Quantitative structure-property relationships, 116
- Queer Linguistics, 85
- Questionnaire data, 91
- Racial hate speech, offensive speech, transfer learning, Toloka, 154
- radio communication, deep learning, asr, maritime, 47
- random forest, 26
- RDM, DMP, FAIR, Climate model data, data standardization, 107
- Reference Limit Estimator, 39
- reinforcement learning, 87
- repositories, 14

- Repository, 108
- reproducibility, 60
- research, 156
- Research Center, 105
- Research data, 108
- research data, 14, 22, 70
- Research Data Management, 52, 59
- research data management, 14, 30, 113
- research information, 22
- Research Resource, 85
- research support, 22
- resource creation, 134
- resource-efficient learning, 153
- right-wing counterpublics, 21
- RNA sequencing, 101

- Scalable Systems, 73
- Scanners, 73
- scholarly, 156
- scientific collection, 160
- Secure communication, , 48
- Security, 73
- Seep Learning, 136
- Self-Supervised Learning, 151
- self-supervised learning, 118
- Self-supervised representations, 129
- Semantic data management, 56
- semantics, 118
- Semi-Automation, 133
- Sensor Integration, 45
- Sensor-driven-approach, 110
- Service Center, 157
- services, 22
- sign language, 66, 125
- Sign Languages, 133
- Sign languages, 135
- sign languages, 134
- Sign representation, 124
- Smart Assembly, 146
- Smart Cities, 138
- smart manufacturing, 149
- Smart Window, 13
- SMPC, 90
- social media, 79
- social networks, 21
- Software Tools, 16
- Softwareplattform, 138

- Soziale Medien, 20
- spatio-temporal data, 67
- Spectra, 148
- Speech Emotion Conversion, 129
- speech emotion recognition, 118
- Speech Enhancement , 127
- Speech Synthesis, 129
- super-resolution, 75
- Survey, 135
- sustainability, 162
- sustainabilty, 22
- Sustainable Research Data Management, 18
- Synchrotron-radiation, 38
- Systems biology, tissue expression prediction, linear regression, dimensionality reduction, 123

- Teaching, 42
- TEI XML, 82
- Teleoperation, 12
- temporal networks, 98
- testing, 69
- text corpus analysis, multilingualism, educational science, 72
- text mining, 106
- Time Series, 136
- topic model, 144
- tracking, 137
- Transformers, 152
- transverse emittance, 88
- Treatment targets, 53
- Twin Transformation, Socio-Technical Transitions Theory, Smart and Sustainable Cities, 27
- Twitter, 43

- Ultrashort Laser, 68
- Unmanned Aerial Systems, 45
- User Groups, 126
- user-centricity, 160

- Variational Autoencoder, 93
- Versorgungsforschung, 40
- Virtual Reality, 12
- virus-infected cells, 101
- Visual-Pattern Classification, 17
- Visual-Pattern Detection, 23
- Visualization, Elastic Search, ERIS Information System, , 158

Voting behavior, 80

WDCC, 108

weak supervision, 141

web-based tool, , 99

WebServer, 32

wordnet, 134

Worker Wellbeing, 57

workflow management, , 64

Writer Identification, 17

Written Artefacts, 18, 25

XSLT, 82