

# HeKKSaGOn ファンドプログラム成果報告書

2017年12月8日

文責：佐藤彰洋

ドイツ・ゲッティンゲン大学にて開催された **Gottingen Data Science Summer School** (2017年7月10日～21日)はゲッティンゲン大学 **Ramin Yahyapour** 教授が実行委員長をつとめ、HeKKSaGOn のデータサイエンス分科会より派遣された日独の大学教員、研究者が行う講義と演習から構成される。以下に期間中のプログラムを示す。2週間にわたるサマースクールに世界中から学生および若手研究者が集まり、約40名の受講があった。京都大学からも修士課程の学生1名がこのサマースクールに参加した。

## Program - Göttingen Data Science Summer School

Website: <https://www.uni-goettingen.de/de/555493.htm> Twitter: #DS3GOE Status: 1.7.2017

Day	16.7.	17.7.	18.7.	19.7.	20.7.	21.7.	22.7.			
Time	Introduction 8:30-9:30 Registration & Welcome Visit 9:30-10:00 Welcome Visit President Casper-Helms and Organizers	Data life cycle and Data Management S01 Schwedtmann: "Data Lifecycle"	Excursion and exchange	Data science methods I S07 Kneib/Grül: "Intro to linear models: Spines, Regularization (LASSO, Boosting) I"	Data science methods II S13 Terada: "Semi-supervised learning for functional data and its applications"	Infrastructure and analytics S19 Keynote: "Data Analytics Infrastructures and Use Cases", NN	Data science practical examples I S25 Sabo: "Indicators of sustainable tourism for development and their applications"	Data science practical examples II S31 Born: "Big data integration and inference in biomedicine"	Ethics, security, privacy, legal aspects S27 Jørgensen/Dierkes: "Open Science"	Panel and Presentations S41 Student panel I: "How does my research profit from Data Science?"
8:30 - 10:00	Coffee break	Coffee break	Coffee break	Coffee break	Coffee break	Coffee break	Coffee break			
10:00 - 10:30	Kocherle: On-line Mathieu-P. Schaprona "TBO"	S03/033 Base: Creation and publishing of data: "The story of Datacube"	Excursions	S08 Gruesz: "Sparse modeling: a revolution on various measurements"	S14 Erasm: "Social information systems"	S20 Yehyapour: "Overview of Infrastructure & Tools, Incentives to Infrastructure for data analysis, and various openly available tools"	S28 Oberweis: "Biological data analysis"	S30 Hensch: "Anonymization of sensitive data"	S45 Student panel II: "How does my research profit from Data Science?" & Q&A I: "Students ask lecturers"	
10:30 - 11:30	Dierkes/Gruel: "Introduction the summer school"		S09 Kamran: "Markov chain Monte Carlo in high-dimension"	S15 Erasm: "Storytelling with web mapping tools"	S21 Yehyapour: "Build your own virtual research environment for reproducible research - Part I: Introduction"	S27 Bestsorn: "High dimensional data in Biomedicine"	S33 Grabowski: "Data Science in Software Engineering I"	S39 Moghaddam: "Security, data privacy?"	S46: Q&A II "Students ask lecturers"	
11:30 - 12:30										
12:30 - 14:00	Intro of participants I	S04 Schwedtmann: Hands-On "Persistent identifier"	Lunch break	S10 Kneib/Grül: "Intro to linear models: Spines, Regularization (LASSO, Boosting) II"	S16 Sporeder/Lenz: "Introduction to Sentiment Analysis & Opinion Mining"	S23 Kallmann: Hands On "Build your own virtual research environment for reproducible research - Part II: Data Organization, Project Organization, and Data Engineering"	S28 Baum/Bauer: Hands On "Data management with Jupyter/SMART"	S34 Schmitt/Mache: "Research data in the Arts and Humanities"	S40 Umbach: "An IT perspective on managing data from genomic high throughput technologies (GenoPerspectiv/IAS)"	S47 Panel discussion: "Data Science: Dedicated profession or everyday business?"
14:00 - 14:30	Coffee break									
14:30 - 15:30	Intro of participants II & topics	S05 Kurzwase: "Challenges for data management in the context of research projects"		S11 Kneib/Wiemann: "Linear Models: Tutorials & Exercises I"	S17 Sporeder/Lenz: "Introduction: Sentiment Analysis & Opinion Mining"	S23 Kallmann: Hands On "Build your own virtual research environment for reproducible research - Part III: Visualization"	S28 Baum/Bauer: Hands On "Data management with Jupyter/SMART"	S35 Blömer: "Hands on TSI & Textgrid"	S41 Schwede: "Ethical Aspects of Data Science Applications"	Wrap-up 16:30 Panel
15:30 - 16:00	Coffee break									
16:00 - 17:30	Intro of participants III & topics	S06 Kurzwase, Weiler: "Challenges for data management in the context of research projects - Bringing in perspectives from library and compute center"	Research Bazaar at Altes-Hesseler-Saal, entrance Building of SUB, Papendiek str. 1 at Adam-von-Trobs-Saal	S12 Kneib/Wiemann: "Linear Models: Tutorials & Exercises II"	S18 Nishida: "Knowledge graph and its application in text mining"	S23 Kallmann: Hands On "Build your own virtual research environment for reproducible research - Part IV: Publication and Sharing"	S30 Ulrich Star: "FAIR perspectives on high quality data management pipeline"	S36 Koenigsmann: "High Energy Physics Big Data and the ATLAS experiment with a hands-on tutorial"	S42 Schwede: "Ethical Aspects of Data Science Applications - hands on"	
17:30 - 18:30										Departure

## Gottingen Data Science Summer School のプログラム

京都大学から講師2名がこのサマースクールに派遣され、佐藤彰洋特定准教授（情報学研究科）は2017年7月18日8:30-10:00までの講義に出講し、「Indicators of sustainable tourism for development and their applications」という講義を担当した。

この講義では、国連2017年開発のための持続可能な観光国際年で取り扱われる持続可能な観光を実現するためのKey Performance Indicators(KPIs)開発を行うためのデータの利活用方法について取り扱った。持続可能性の概念を環境、社会、経済の3分野における各活動の長期にわたるバランスと捉えることにより多目的最適化問題として定式化されることを述べ、国連持続可能な開発アジェンダ(Sustainable Development Goals; SDGs)が

多目的最適化問題におけるパレートフロンティア上に複数の多目的最適化解集合を許容し得ることを述べた。

更に、国連世界観光機関が発行する旅行・観光サテライト勘定（Tourism Satellite Account; TSA）について説明すると共に、データ収集時における同一性と一致性を担保するための標準の重要性を解説した。更に、KPI として利用できる時間と空間の分解能に基づく指標の分類を説明した。

特に地域レベルにおける持続可能性を定量的に評価できる指標を全世界レベルで実現するために、我が国で40年以上使用されている地域メッシュコード(JIS X0410)と地域メッシュ統計を全世界拡張した世界メッシュコードと世界メッシュ統計を説明した。また、R 言語を用いた世界メッシュ統計作成方法の実習および、佐藤彰洋准教授が科学技術振興機構さきがけ研究「グローバルシステムの持続可能性評価基盤に関する研究」で独自開発を行っている、世界メッシュ統計を用いた世界メッシュ統計定量化可視化基盤(MESHSTATS)の実演解説を行った。

国連世界観光機関が提唱する3フェーズから構成される12ステップの持続可能な観光のための指標開発方法と、この開発方法論のフェーズ1に位置付けられる異質的参加者からなる参加型アプローチによる指標開発の実習を受講生とともにを行った。

サマースクール開催中の2017年7月13日～14日には、HEKKSaGOn データサイエンス分科会のF2F 会合が開催された。佐藤彰洋特定准教授はこの会議に参加し、今後の交流について日独の大学教員、研究者と共に意見交換を行った。



