



Faculty IV – Electrical Engineering and Computer Science

The Data Science and Engineering Master's Track

Juan Soto and Prof. Dr. Volker Markl

Chair of Database Systems and Information Management (DIMA)

Key Objectives

1. Data Analytics Master's Track established in **Fall 2013** under the purview of Faculty IV – Electrical Engineering and Computer Science
2. Renamed **Data Science & Engineering (DSE) Master's Track** in **April 2019**
3. Provides students with a **structured roadmap**
4. Offers **guidance** in course selection
5. Enables students to **specialize** in DSE
6. Better **prepare** students to pursue careers in DSE

The screenshot shows the homepage of the Data Analytics Lab at TU Berlin. At the top right is a navigation bar with links for Contact, Impressum, Sitemap, Deutsch, Index A-Z, Mobil, Datenschutz, TUB-Login, and a search bar. Below the navigation is a banner for the Data Analytics Lab. On the left, there's a sidebar with links for News, About us (which is highlighted in red), Motivation, Mission, Subject Area Experts, and Data Science and Engineering Track. The main content area features a heading "About the Data Analytics Laboratory" with a subtext about the lab's history and mission. To the right is a photograph of a server room.

search for ...

Contact Impressum Sitemap Deutsch Index A-Z Mobil Datenschutz

TUB-Login with Password

Electrical Engineering and Computer Science
Data Analytics Lab

Quick Access:
Go to:

Auxiliary Functions
 maximize aux. functions

Data Science and Engineering Track Contact
Juan Soto
+49 30 314 23551
Room Office: E-N 724
 e-mail query

About the Data Analytics Laboratory

About the Data Analytics Laboratory

As a leading academic institution, the Technische Universität Berlin (TUB) is spearheading numerous initiatives to deliver impactful/groundbreaking research in the big data arena. One of these initiatives was the establishment of the [Data Analytics Laboratory](#) in 2011, to serve as a focal point for innovative research and curriculum development.

Our aim is to coordinate our efforts, bring about synergies, and leverage our collective expertise to address the growing big data challenges.

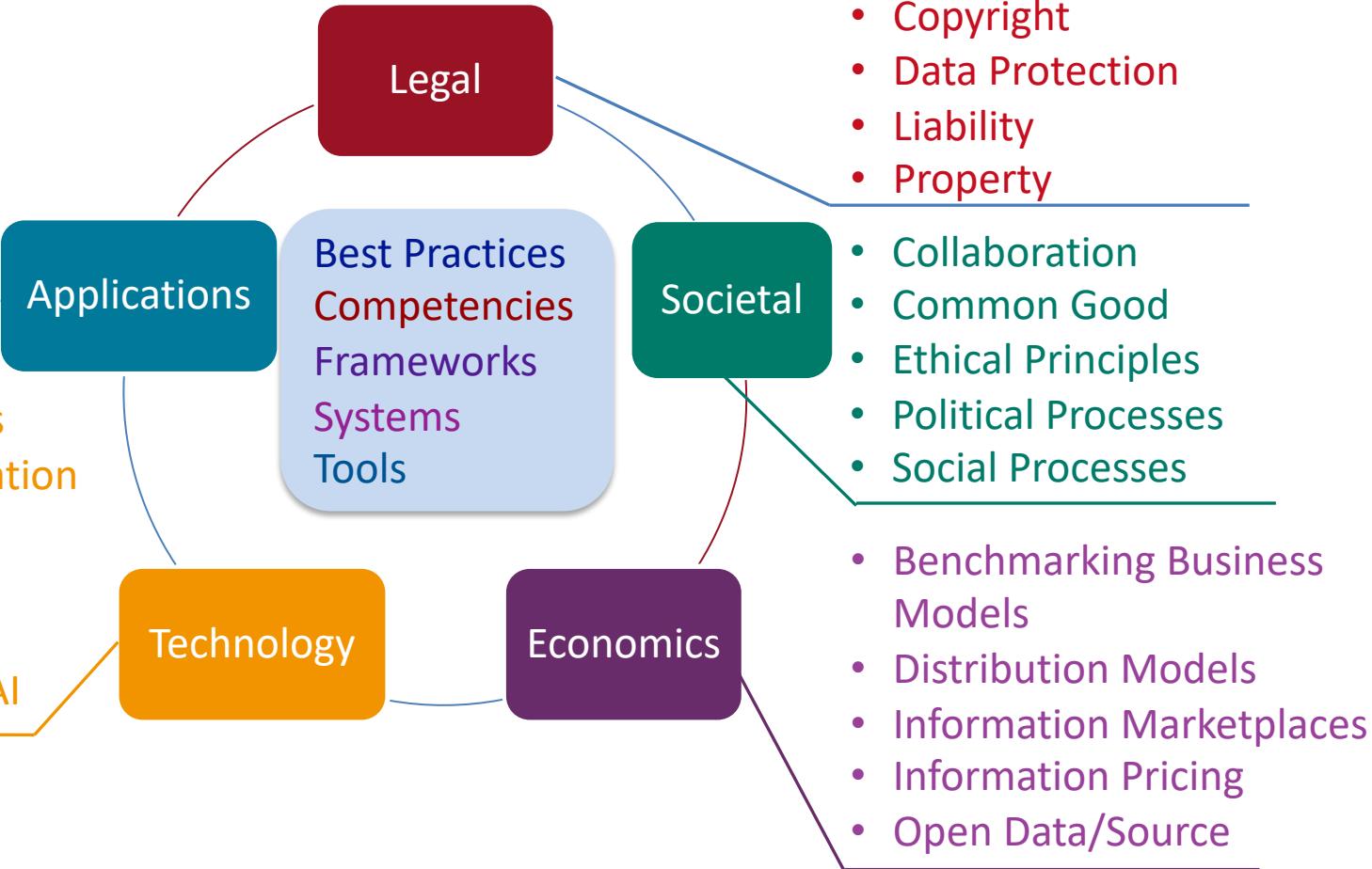
The Data Analytics Lab will build upon the extensive expertise of TUB investigators in **scalable and high-velocity data processing, information integration, statistics, visualization, machine learning, signal processing, compression, security, and networking**.

<http://analytics.tu-berlin.de>

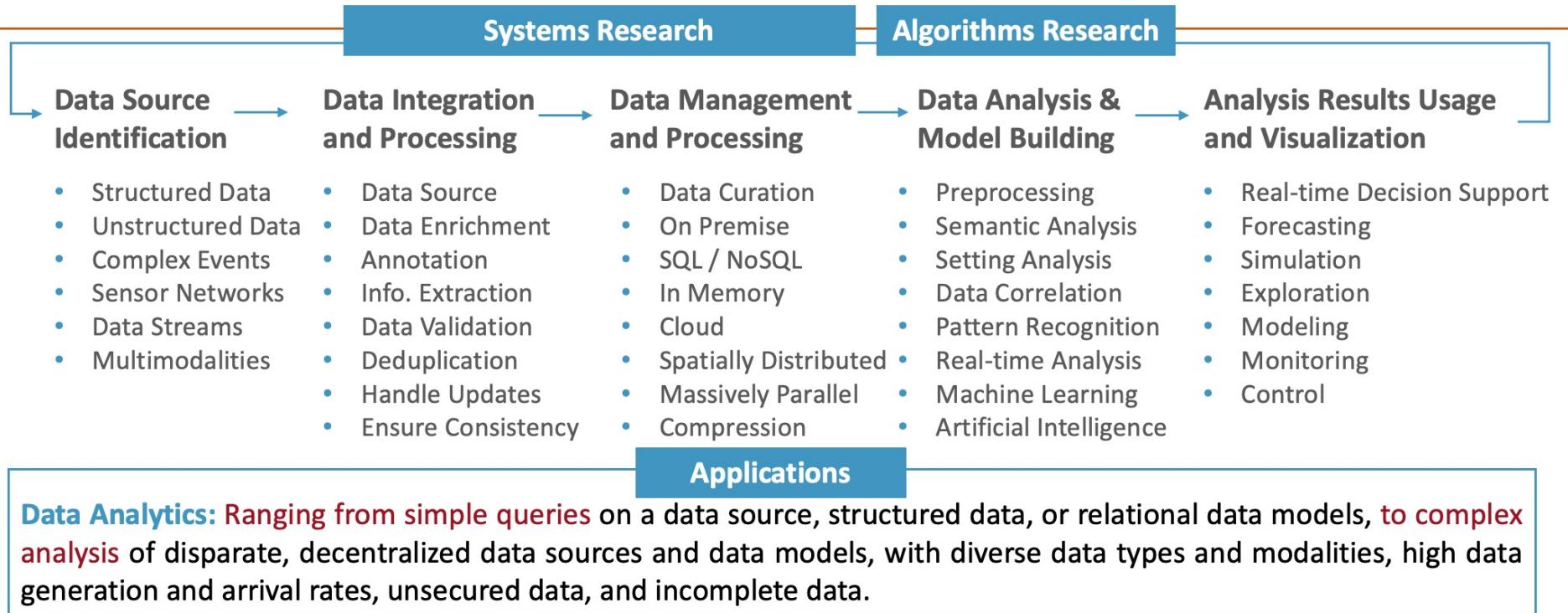
Facets of Big Data/Data Science

- Economy
- Engineering
- Humanities
- Medicine
- Politics
- Sciences

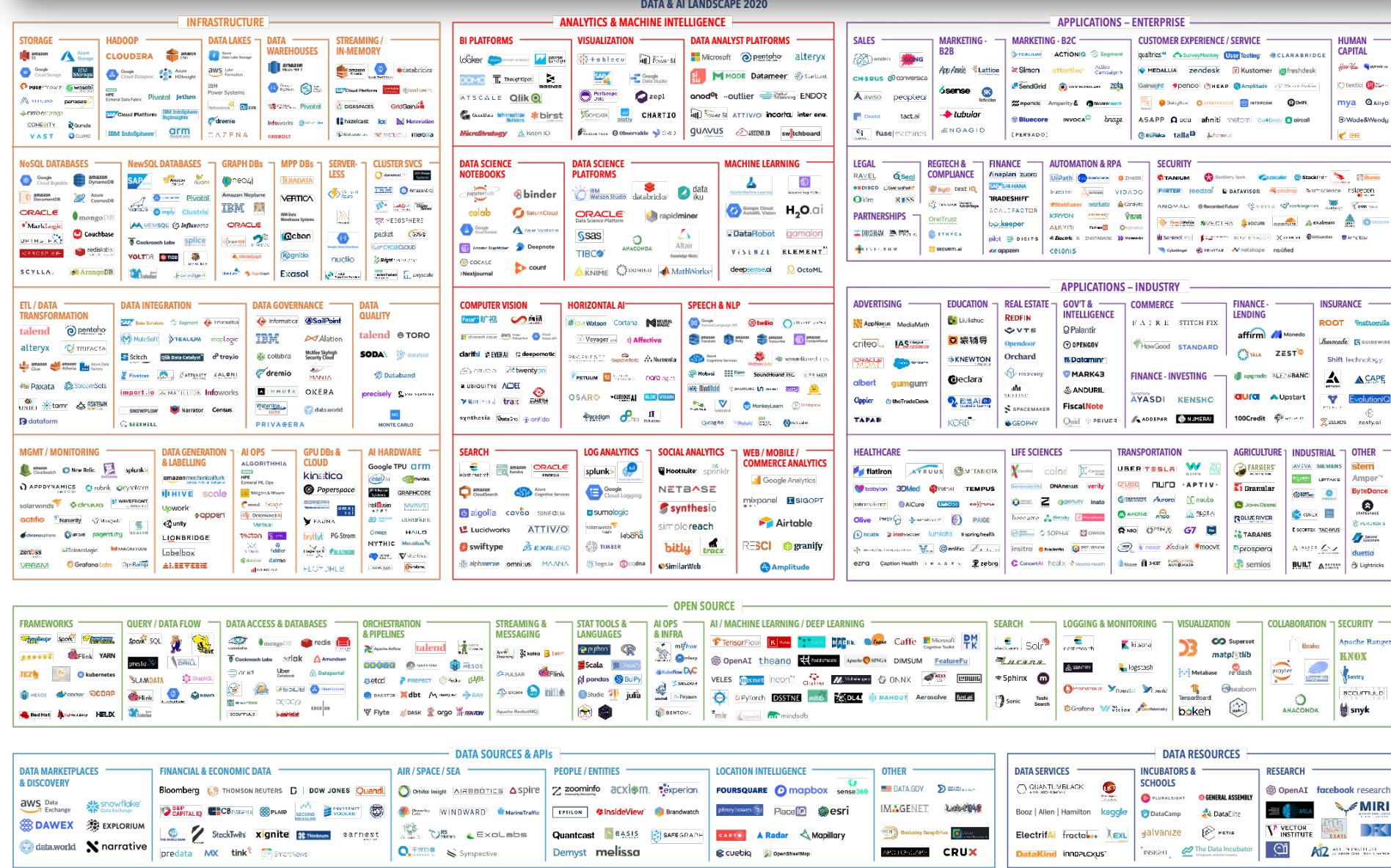
- Data Management
- Distributed Systems
- Interactive Visualization
- Language
- Security
- Signal Processing
- Statistics, ML, and AI



Big Data: A New Standard for Data Analytics



Constellation of big data, data science, and AI tools and technologies



A Call to Action

- **Educate Data Scientists**

- T-shaped and even Π (pi-shaped) students that possess *breadth* and *depth*
 - information literacy
 - data analytics curriculum

- **R&D Big Data Analytics Technologies**

- **data management** (uncertainty, query processing under near real-time constraints, information extraction)
 - **programming models**
 - **machine learning and statistical methods**
 - **systems architectures**
 - **information visualization**

- **Innovate to Maintain Competitiveness**

- demonstrate flagship use-cases to raise awareness
 - promote startups in the area of data analytics
 - transfer technologies to German enterprises, in particular SMEs
 - determine legal frameworks and business models

... to ensure German technological leadership in big data.



Latest Guidance Document



The Technische Universität Berlin Faculty IV Electrical Engineering and Computer Science The Data Science and Engineering (DS&E) Master's Track: A Guidance Document (Version 4.1)

Juan Soto and Prof. Dr. Volker Markl

Database Systems and Information Management (DIMA) Group

Last Updated: October 14, 2020

Synopsis. In Fall 2013, TU Berlin's Faculty IV Electrical Engineering & Computer Science (EECS) approved a new track, which enables students pursuing a M.Sc. in Computer Science, Information Systems Management or Computer Engineering, to specialize in data science and engineering. To meet the track requirements, students must complete courses in three core competencies: (1) *scalable data analytics*, (2) *scalable data management*, and (3) *a domain-specific application area*. This guidance document offers students general advice: in the selection of courses, the procedure to follow when identifying a thesis topic, and prospective career possibilities. **In April 2019, the track was renamed, the *Data Science & Engineering Master's Track*.** From SS 2019 on, students who complete both their respective M.Sc. degree and track requirements, will receive – besides their M.Sc. degree – a *Data Science & Engineering Master's Track Certificate* issued by Faculty IV. **Questions or comments concerning this document should be directed to lehre@dima.tu-berlin.de.**



Data Science & Engineering Track Description

2. Detailed Descriptions of the Data Science and Engineering Master's Track Rules

Please study the following subsections very carefully, most of your questions should be answered.

2.1 Qualification and Main Competence Areas. The Data Science and Engineering Master's Track qualifies students to pursue careers as a *Data Scientist*, *Data Analyst*, or *Data Engineer*. They will learn about data analysis methods, their application to real-world problems in varying domains, learn more about the internals of database systems, and develop programming skills with a focus on massively-parallel data processing systems.

2.2 Requirements. Students following the track should be enrolled in one of the following TU Berlin Master's Programs: *Computer Science* ('Informatik'), *Information Systems Management* ('Wirtschaftsinformatik') or *Computer Engineering* ('Technische Informatik'). *Their acceptance to the Data Science and Engineering Track is by default.*

2.3 Prerequisites: Students interested in joining the track should possess: (a) very strong English language skills, (b) programming skills in functional (e.g., Scala) and object oriented (e.g., Java) programming languages, (c) fundamental skills in database management systems, and (d) knowledge in mathematical foundations (e.g., linear algebra, probability, statistics).

2.4 Credit Points and Track Structure. To earn a M.Sc. degree, students must achieve 120 ECTS credit points. Of these, 90 ECTS credit points must fulfill the requirements described further below, to qualify for the track certificate.

Description of the Track Structure

Credit Points	Competence	Course	Notes
24 ECTS	Data Analytics (DA)	Machine Learning 1 or Machine Intelligence I	<i>mandatory</i> course
		DA Elective 1	see Appendix A, Table 1
		DA Elective 2	
		DA Elective 3	
18 ECTS	Scalable Data Management (SDM)	Database Technology	<i>mandatory</i> course
		SDM Elective 1	see Appendix A, Table 2
		SDM Elective 2	
6 ECTS	Domain Specific Application (DSA)	DSA Elective	see Appendix A, Table 3
9 ECTS	Project	Project Elective	see Appendix A, Table 4
3 ECTS	Seminar	Seminar Elective	see Appendix A, Table 5
30 ECTS	Thesis	Master's Thesis	The thesis must be a <i>data science oriented</i> topic, supervised by a TU Berlin Data Analytics Lab Professor .
Total: 90 ECTS			



Enrollment, Mentorship, Updates

2.5 Enrolling in the Track. To enroll in the track, students must join the “*Data Science & Engineering Track*” course located at <https://isis.tu-berlin.de/course/view.php?id=16781>. Students are advised to complete the Excel spreadsheet located here: <https://isis.tu-berlin.de/mod/folder/view.php?id=694766> and forward it on to Juan Soto ([juan dot soto at tu-berlin dot de](mailto:juan.soto@tu-berlin.de)) for review.

2.6 Mentoring Program. Track participants are invited to contact a [member of the Data Analytics Lab](#) to identify a mentor and request guidance.

2.7 Changes to the Track. Track requirements may change annually. Therefore, students are required to regularly monitor announcements posted on the ISIS *Data Science and Engineering Track* forum.

Today, there are over **175 students** who have expressed interest in completing the **data science and engineering track**.

To date, twelve students have completed the track.

Several others intended to complete the track, however, they **did not fully meet** all of the track requirements.

Appendix A. Representative List of Elective Master's Courses Across Competency Areas

Special Instructions (Read Carefully):

1. Below we list a *representative* list of elective courses that should meet track requirements across varying competencies. If a student wishes to enroll in a course that is not explicitly listed in one of the tables listed below, then you are urged to reach out to *Juan Soto* via email or in person, to obtain assurance that the course meets track requirements, **prior to enrolling in the course**.
2. **TU Berlin's course catalog is fairly vast. Thus, in this document, we are unable to maintain an accurate record.** For example, regarding when a course will be offered (i.e., WiSe or SS), the specific target language spoken in class (i.e., EN or DE), or whether new courses will be coming online, among other things. Therefore, students are responsible to obtain the latest information. Students are urged to review the latest course offerings as contained in the Technische Universität Berlin *Course Catalog*: <https://moseskonto.tu-berlin.de/moses/modultransfersystem/bolognamodule/suchen.html>.
3. Unfortunately, **course schedules (i.e., day and time) are subject to change.** There have been instances where some courses are offered at the exact day and time. In these cases, students should seek to resolve scheduling conflicts by appropriately selecting their courses.
4. **Project / Seminar courses can only be applied to the Project / Seminar requirement, respectively.**
5. **For a current list of courses students are advised to visit the following groups and their respective webpages.** Bear in mind that we cannot list all group at TU Berlin. *The compilation below is representative and incomplete!*

Group	URL
Agent Technologies in Business Applications and Telecommunication	https://www.aot.tu-berlin.de/
Algorithmics and Computational Complexity	https://www.akt.tu-berlin.de/menue/teaching/
Artificial Intelligence	https://www.ki.tu-berlin.de/menue/teaching
Database Systems and Information Management	https://www.dima.tu-berlin.de/menue/teaching/
Distributed and Operating Systems	https://www.dos.tu-berlin.de/menue/teaching/
Econometrics and Business Statistics	https://www.statistik.tu-berlin.de/menue/studium_und_lehre/aktuelles_lehrangebot/
Embedded Systems Architecture	https://www.aes.tu-berlin.de/menue/courses/
Machine Learning	http://wiki.ml.tu-berlin.de/wiki/
Models and Theory of Distributed Systems	https://www.mtv.tu-berlin.de/menue/lehre/parameter/en/
Neural Information Processing	https://www.ni.tu-berlin.de/menue/teaching_activities/
Open Distributed Systems	https://www.ods.tu-berlin.de/menue/teaching/parameter/en/
Quality and Usability Lab	https://www.qu.tu-berlin.de/menue/studium_und_lehre/parameter/en/
Remote Sensing Image Analysis	https://www.rsim.tu-berlin.de/menue/teaching/parameter/de/
Service Centric Networking	https://www.snet.tu-berlin.de/menue/teaching_and_exams/

Table 1. A Representative List of Eligible *Data Analytics* Courses.

Course Title	ECTS	Professor
Machine Learning 2	9	Klaus-Robert Müller
Machine Learning Lab	9	Klaus-Robert Müller
Machine Intelligence II	6	Klaus Obermayer
Monte Carlo Methods in Machine Learning and AI	6	Manfred Opper
Probabilistic and Bayesian Modelling in ML and AI	6	Manfred Opper
Digital Communities	6	Axel Küpper
Econometric Analysis of Longitudinal and Panel Data	6	Axel Werwatz
Microeometrics	6	Axel Werwatz
Multivariate Analysis/Business Statistics	6	Axel Werwatz
Time Series Analysis	6	Axel Werwatz
Treatment Effect Analysis	6	Axel Werwatz
Ökonometrie (Econometrics)	6	Axel Werwatz
Numerische Mathematik für Ingenieure II	10	Jörg Liesen
Stochastische Modelle (Stochastic Models)	10	Michael Scheutzow
Digitale Signalverarbeitung (Digital Signal Processing)	12	Reinhold Orglmeister

Table 2. A Representative List of Eligible *Scalable Data Management* Courses.

Course Title	ECTS	Professor
AIM-2 Management of Data Streams	6	Volker Markl
AIM-3 Scalable Data Science: Systems & Methods (SDSSM)	6	Volker Markl
IDB-PRA: Implementation of a Database Engine (Database Technology Lab Course)	6	Volker Markl
CIT 9 - Cloud Computing	6	Odej Kao

Table 3. A representative list of eligible *domain specific application* courses.

Course Title	ECTS	Professor
Energiewirtschaft - Elektrizitätswirtschaft	6	Christian Hirschhausen
Energiewirtschaft - Technologie und Innovation	6	Christian Hirschhausen
Energy Economics	6	Georg Erdmann
Experimental and Behavioral Economics	6	Dorothea Kübler
Gesundheitsökonomie II	6	Marco Runkel
Integriertes Informationsmanagement	6	Rüdiger Zarnekow
IT-Service-Management	6	Rüdiger Zarnekow
Intelligente Sicherheit in Netzwerken (IT Sec. in Networks)	9	Sahin Albayrak
Patentrecht/Patentmanagement (Patent Rights / Mgmt.)	6	Jürgen Ensthaler
Speech Signal Processing and Speech Technology	6	Sebastian Möller
The Economics of Climate Change	6	Ottmar Edenhofer

Other potential courses include:

- Computational Biology
- Image Processing for Remote Sensing
- Intelligent Cybersecurity Applications,
- Electric Energy Networks
- Medical Image Processing
- Planning and Security of Smart Grids
- Vehicular Networking and Cooperative Driving
- ...

Table 4. A representative list of eligible *project* courses.

Course Title	ECTS	Professor
IMPRO3 - Big Data Analytics Project (BDAPRO)	9	Volker Markl
Verteilte Systeme (Distributed Systems Project)	9	Odej Kao
Project Machine Learning	9	Klaus-Robert Müller
Project Neural Information Processing	9	Klaus Obermayer
Project: Statistical Methods in AI and ML	9	Manfred Opper
Projekt Nachrichtenübertragung (Signal Processing Project)	6	Thomas Sikora

Table 5. A representative list of eligible *seminar* courses.

Course Title	ECTS	Professor
Anwendungen Kognitiver Algorithmen (Applied Cognitive Algorithms)	3	Klaus-Robert Müller
BDASEM - Big Data Analytics Seminar	3	Volker Markl
CIT 8 - Aktuelle Themen aus dem Bereich der verteilten Systeme (Hot Topics in Distributed Systems)	3	Odej Kao
Hot Topics in Operating Systems & Distributed Systems	3	Hans-Ulrich Heiß
IMSEM - Seminar Hot Topics in Info. Management	3	Volker Markl
Introduction to Computational Genomics	3	Manfred Opper
Seminar: Operating Complex IT Systems	3	Odej Kao
Recent Advances in Computer Architecture	3	Bernardus Juurlink
Recent Advances in Multicore Systems	3	Bernardus Juurlink
Synchronous and Asynchronous Interactions in Distributed Systems	3	Uwe Nestmann



Appendix C. Questions and Answers

Q1. What is a track?

A1. In general, a track is a suggested sequence of courses that profile a specific specialization. Students who successfully complete the track will be awarded a certificate from Faculty IV. A certificate indicates that a student has followed a structured academic program with the intent to pursue specialization in data science.

Q2. Who can follow a track?

A2. By default, students enrolled in the Computer Science (“*Informatik*”), Information Systems Management (“*Wirtschaftsinformatik*”) or Computer Engineering (“*Technische Informatik*”) Master’s programs are eligible to pursue the track. **Unfortunately, due to resource constraints, we are unable to consider other study programs at this time beyond the three mentioned above.**

Q3. Will my study period be extended, if I follow the track?

A3. No, neither the amount of ECTS credit points, nor the number of semesters will increase. Moreover, a longer study period will not lead to a disqualification from the track.

Q4. How to go about selecting a thesis topic?

A4. Students should speak with Senior Researchers, Postdocs, or PhD students, in the participating research groups, i.e. “Chairs,” to identify an open thesis topic of mutual interest. For a list of representative data science oriented publications have a look at [3, 4], and for Master’s Thesis topics see [5]. For a glimpse into ongoing research activities in big data/data science see [6]. For open problems and a vision of the future of computer science see [7, 8, 9], respectively.

Q5. What are my prospective career possibilities?

A5. Students who complete the data analytics track are prepared to pursue careers as *Data Analysts*, *Data Engineers*, or *Data Scientists*. For information about big data projects in industry within Germany have a look at [10]. In some cases, students enter a PhD program with the aim to further specialize in a research topic, such as *deep learning* or *streaming systems*. Examples of recent (DIMA specific) PhD thesis topics, include [11, 12]. For more information about job opportunities and earning potential across Europe have a look at [13].

Q6. If I still have questions or doubts, not answered yet?

A6. This document is assumed to be comprehensive. It should address the most relevant questions. In case of any doubt (e.g., you are enrolled in a different study programme) or concern, please contact us at lehre@dima.tu-berlin.de. Also, please look for announcements (e.g., the bi-annual “*Data Science and Engineering Track Intro Presentation*”) posted on the *Data Science and Engineering Track* forum in ISIS.

Q7. How do I obtain my certificate?

A7. You will need to present evidence (e.g., academic transcript) that you have met the track requirements. Once this has been verified, DIMA staff will prepare your certificate.

Seeking Advice and Plan Approval



Use this Template When Seeking Advice About the Track

by Juan Soto - Montag, 4 Juni 2018, 6:47

To all track participants,

Attached is an Excel spreadsheet. This *template* should be used by students to document their course selections.

Please complete and forward the template to me via email for review.

Best,

Juan

Data Analytics Masters Track - Template.xlsx

[PERMALINK](#) | [EDIT](#) | [DELETE](#)

[DISCUSS THIS TOPIC](#) (0 REPLIES SO FAR)

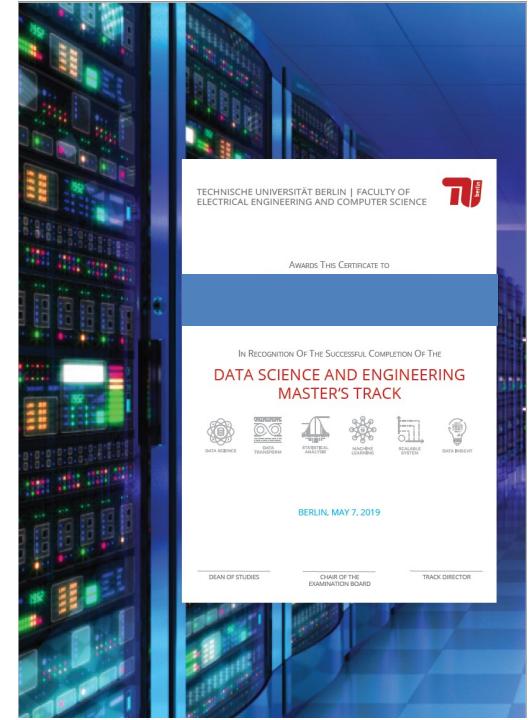
Please complete the Excel sheet named **Proposed Plan** below. You should specify the courses you plan to complete, that satisfy the requirements of the certificate. Then forward the Excel spreadsheet on to **Juan Soto (juan.soto@tu.berlin.de)** for review. In those cases, when a course does not appear on the track guidance document, please include a link to the course syllabus for review and approval.

Once all of your courses are complete, you should specify the **Actual Plan** sheet below and then **forward the Excel spreadsheet, jointly with your academic transcript** for a final review. Afterwards, our support staff will generate a certificate and we will then forward it on to obtain the appropriate signatures. You will subsequently be contacted once the certificate is ready for pick up.

	Course	ECTS
Data Analytics	Machine Learning I Machine Intelligence I	6
	Data Analytics Elective 1	6
	Data Analytics Elective 2	6
	Data Analytics Elective 3	6
Scalable Data Management	Database Technology	6
	Scalable Data Management Elective 1	6
	Scalable Data Management Elective 2	6
Project	Data Science Oriented Project	9
Seminar	Data Science Oriented Seminar	3
Domain Specific Application	Domain Specific Application Elective	6
Master's Thesis	Data Science Oriented Thesis	30
	Total	90

TU Berlin Recipients of a Data Science and Engineering Master's Track Certificate

1. M. Nowacka, *Data Engineer*, Netlight
2. S. Alaniz, *PhD Student*, University of Tübingen
3. B. Pietrowicz, *Engineer*, Biotronik
4. A. Bartnik, *Consultant*, Netlight
5. J. Meiners, *Research Associate*, TU Berlin
6. J. Dikow, *CTO*, Mapegy
7. P. Lehmann, *Data Scientist*, Modomoto
8. Z. Lux, *PhD Student*, TU Berlin
9. O. Chumak, *Data Engineer*, CARIAD
10. L. Füchsel, *Software Engineer*, ONSEI GmbH
11. I. Winata, *Data Scientist*, Center of Excellence for Data Analytics at Deutsche Bank
12. M. Hienen, *Business Intelligence Specialist*, Aurubis AG





References

- [1] *“Breaking the chains: On declarative data analysis and data independence in the big data era,”* Volker Markl, PVLDB, 7(13):1730–1733, 2014. URL: www.vldb.org/pvldb/vol7/p1730-markl.pdf.
- [2] *Towards a Thriving Data Economy: Open Data, Big Data, and Ecosystems* (Presentation), Volker Markl, European Competitiveness Council, March 2015. URL: goo.gl/eDRSS3.
- [3] FG DIMA Data Science Publications: <http://www.dima.tu-berlin.de/menue/publications/publications/>.
- [4] FG ML/IDA Machine Learning Publications: <http://doc.ml.tu-berlin.de/publications/>.
- [5] *Completed Master’s Theses:* Many Data Science Oriented, DIMA Group. URL: http://www.dima.tu-berlin.de/menue/theses/completed_mastersdiploma_theses/.
- [6] BIFOLD (Berlin Institute for the Foundations of Learning and Data), <https://bifold.berlin/>.
- [7] *Future Directions in Computer Science Research* (Presentation: TU Berlin, Big Data Workshop), John Hopcroft, Cornell University, September 2013. URL: <http://www.eecs.tu-berlin.de/index.php?id=139969>.
- [8] *50 Years of Data Science* (Version 1.00), David Donoho, Stanford University, September 2015. URL: <http://courses.csail.mit.edu/18.337/2015/docs/50YearsDataScience.pdf>.
- [9] **Frontiers in Massive Data Analysis**, National Academies Press, 2013. URL: <http://nap.edu/18374>.
- [10] Germany – Excellence in Big Data, Bitkom, 2016. URL: goo.gl/wUSZWv.
- [11] *Scaling Data Mining in Massively Parallel Dataflow Systems* (PhD Thesis), S. Schelter, November 2015.
- [12] *Visualization-Driven Data Aggregation* (PhD Thesis), U. Jugel, TU Berlin, April 2017.
- [13] *The European Data Science Salary Survey: Tools, Trends, What Pays (and What Doesn’t) for Data Professionals in Europe*, John King & Roger Magoulas, O’Reilly Press, 2017.