

Data Science 1 - SS 2020

Dr. Karsten Tolle

ZOOM handling ...

- Please mute your micro! ... unless you want to speak!
- ... use also the chat function if possible.
- ... voting by pressing yes/no under „Teilnehmer“! ... at least we try it with so many😊

Who is in the air ... ?

- Bachelor Computer Science
- Bachelor other
- Master Computer Science
- Master Computer Science Spezialisierung „Data Science“
- Wirtschaftsinformatik Master
- Master other

Situation: Präsidium ...


- Lehrende sind gebeten, für das Sommersemester 2020 soweit möglich auf Klausuren als Prüfungsform zu verzichten und großzügig auf **Äquivalenzleistungen** auszuweichen. Dies gilt explizit auch dann, wenn die studiengangspezifische Ordnung lediglich eine Klausur als Prüfungsleistung vorgesehen hat.

All the following information is based on the currently applied and accepted rules. It is not impossible that these rules change, which could result in changes of the rules for this course as well!

... official registration ...

- I am waiting for details from the Prüfungsamt.
- ... currently an online registration similar to a written exam via QIS/LFS is envisioned.


How to pass DS1 ...

1. Mandatory (without grading): do selected courses at  <https://cognitiveclass.ai/> (former Big Data University – IBM)
2. Mandatory (with grading): do a project in a small team (two persons – in exceptional cases up to three): ...

Deadlines:

1. ... till 28.07.2020 8pm
2. ... till 26.06.2020 8pm

How to pass DS1 ...

1. Mandatory (without grading): do selected courses at  <https://cognitiveclass.ai/> (former Big Data University – IBM)
(Deadline 28.07.2020 8pm)

Task 1: Mandatory (without grading):

You need to perform at least 5 points via courses from <https://cognitiveclass.ai/> (former Big Data University – IBM).

The point per course are according their rated level:

Beginner: 1 point;

Intermediate: 2 points;

Advanced: 3 points.

You will need to upload your certificates in the OLAT course (all in one shot). **The certificates must be dated after 01. April 2020!**

Note: It would be great, if you would write a small report within the OLAT forum in order to give feedback to the other students how helpful the course was!

Set of PDFs plus a .txt-File with the list of URLs as a ZIP-File, would be best!

List of courses that are accepted for DS1. In case you want to take different courses, this should be discussed upfront (in this case send a request via OLAT or directly to tolle@dbis.cs.uni-frankfurt.de).



[Introduction to Data Science](#) Cognitive Class / Fireside Analytics Inc. DS0101EN Beginner



[Statistics 101](#) Cognitive Class ST0101EN Beginner



[Python for Data Science](#) PY0101EN Beginner



[Machine Learning with Python](#) Cognitive Class ML0101ENv3 Intermediate



[Data Visualization with Python](#) DV0101EN Intermediate



[Data Analysis with Python](#) Cognitive Class DA0101EN Intermediate



[R 101](#) Cognitive Class RP0101EN Beginner



[Machine Learning with R](#) Cognitive Class ML0151EN Intermediate



[Data Visualization with R](#) Cognitive Class DV0151EN Intermediate



[Deep Learning Fundamentals](#) DeepLearning.TV ML0115EN Intermediate



[Deep Learning with TensorFlow](#) Cognitive Class ML0120ENv2 Advanced



[Data Science Methodology](#) Cognitive Class DS0103EN Beginner



[Data Science Tools](#) DS0105EN Beginner



[Machine learning with Apache SystemML](#) Cognitive Class ML0111EN Intermediate



[Predictive Modeling Fundamentals I](#) Cognitive Class PA0101EN Intermediate



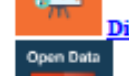
[Data Science for Scala](#) Lightbend SC0105EN Intermediate



[Machine Learning – Dimensionality Reduction](#) Cognitive Class ML0109EN Beginner



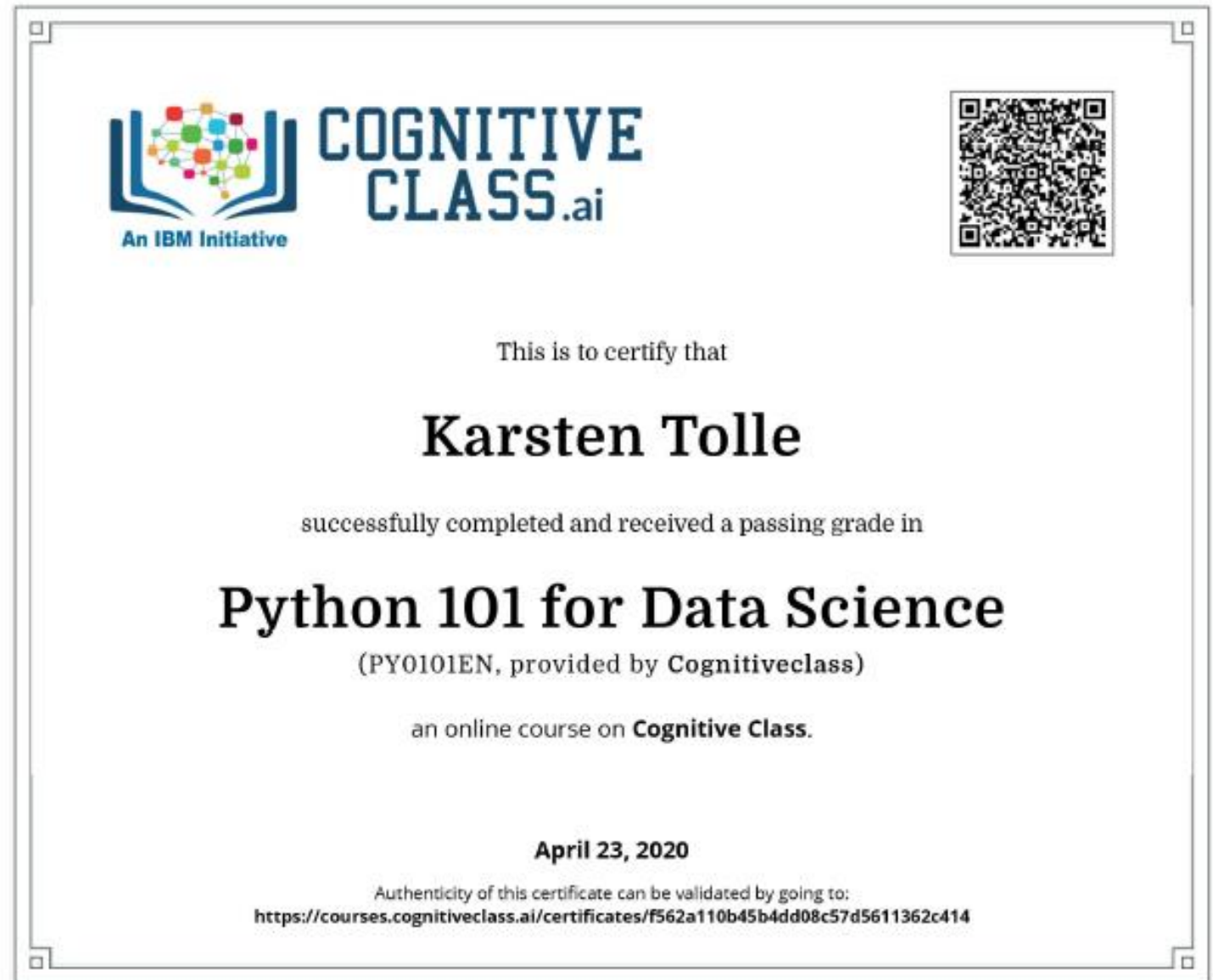
[Digital Analytics & Regression](#) Cognitive Class / Fireside Analytics Inc. ML0103EN Beginner



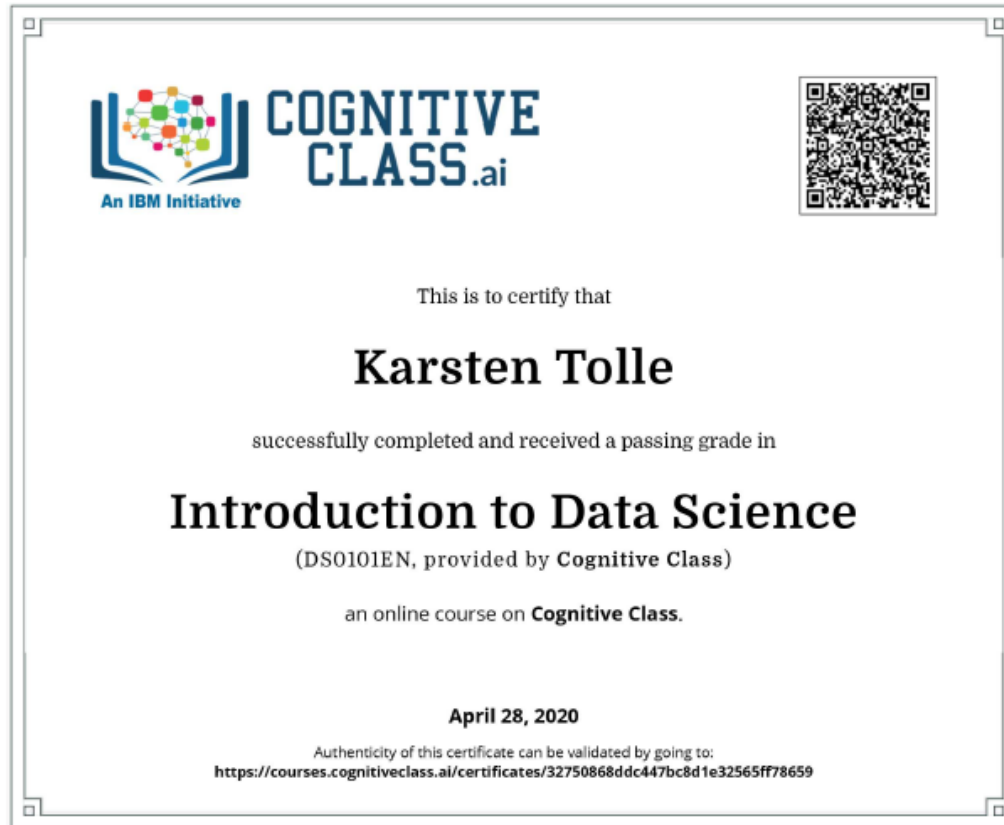
[Data Science with Open Data](#) Fireside Analytics Inc. / National Energy Board of Canada DS0110EN Beginner

Certificate ...

<https://courses.cognitiveclass.ai/certificates/f562a110b45b4dd08c57d5611362c414>



Certificate and badge ...



cclaim
by Credly

This badge was issued to [Karsten Tolle](#) on 28 April 2020.

Data Science Foundations - Level 1

Issued by [IBM](#)

This badge earner has an understanding of the possibilities and opportunities analytics and big data bring to new applications in any industry.

Skills

- Data Science
- Data Scientist Workbench
- Dswb
- Python
- R
- R Program
- Statistics

[Additional Details](#)

Insights

TOP JOB TITLES

- Senior Lead Healthcare Architect Watson Health Service
- Associate Partner Cognitive Analytics

How to pass DS1 ...

1. Mandatory (without grading): do selected courses at <https://cognitiveclass.ai/> (former Big Data University – IBM)
2. **Mandatory (with grading): do a project in a small team (two persons – in exceptional cases up to three):**
 1. Do a plan what you want to do ... define the goal ... then:
 2. Take existing open data from at least two different sources.
 3. Merge and clean the data.
 4. Test and verify your data quality.
 5. (optional) Further preprocessing.
 6. Apply two different algorithms of the same kind ...
 7. Evaluate and verify the results of them!
 8. Come up with a conclusion!



Regression
Classification
Clustering
Decision Trees
Data Generation
Image Processing
Speech Processing
Natural Language Processing
Recommender Systems
Adversarial Networks
Reinforcement Learning

... as output for 2. you need to
(Deadline: 26.06.2020 8pm):

1. **Project report** – max. 5 pages (up to 1500 characters per page including blanks) ... details on an expected structure will follow!
2. Provide the **merged and cleaned datasets** you generated together with the **code** you used (reproducible), e.g. Github with JupiterNotebook
3. Short **presentation** of you project – details will follow, depends also on the number of final participants.

Sprache: Deutsch oder Englisch!

What happens during the semester ...

Datum	Inhalt
28.04.2020	Orga und so
05.05.2020	Was ist DS?, Master mit Schwerpunkt "Data Science" , Team-Building
12.05.2020	Vorstellung eigener Forschung: Nomisma.org, CN, ...
19.05.2020	Vorstellung Tools und deren Verwendung I
26.05.2020	Vorstellung Tools und deren Verwendung II, einige Kernbegriffe
02.06.2020	Mentoring sessions
09.06.2020	Mentoring sessions
16.06.2020	Mentoring sessions
23.06.2020	Mentoring sessions
26.06.2020	DL for Task 2
30.06.2020	Präsentationen
07.07.2020	Präsentationen
14.07.2020	Präsentationen
28.07.2020	DL for Task 1

How to build teams ...

- In OLAT the members should see each other!
- Use the forum in order to search for partners.
- Next week ZOOM meeting!

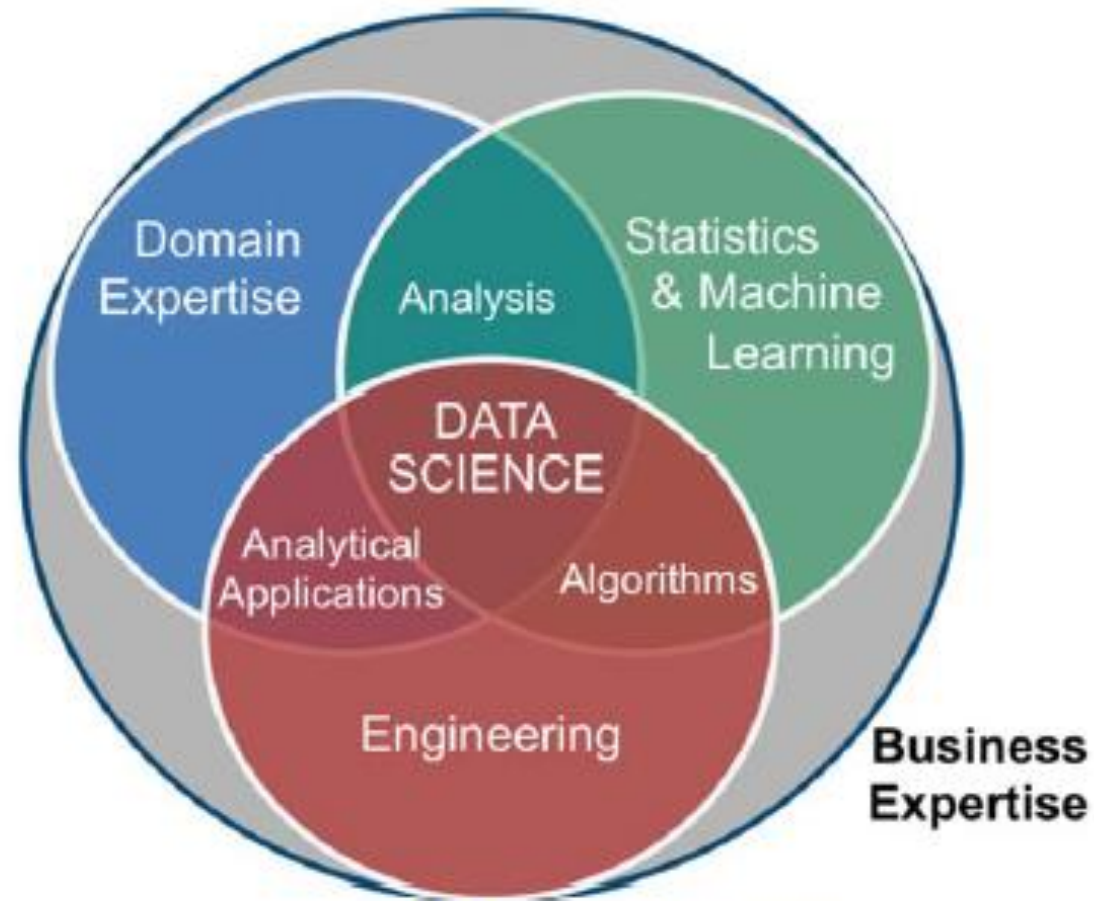


Figure 1: Skills Needed in Data Science

EDISON – EU-Project (2015 – 2017)

Education for **D**ata Intensive **S**cience to **O**pen **N**ew science frontiers

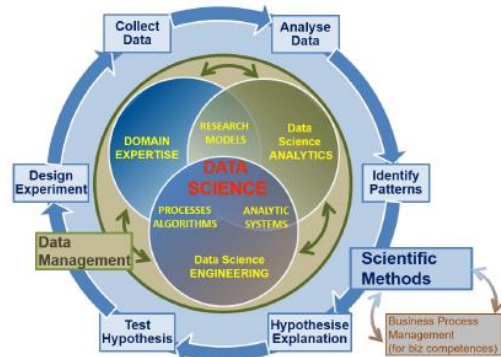
Documents (in OLAT or at <https://zenodo.org>):

- Data Science Competence Framework (CF-DS),
- Data Science Body of Knowledge (DS-BoK),
- Data Science Model Curriculum (MC-DS),
- Data Science Professional Profiles (DSPP)

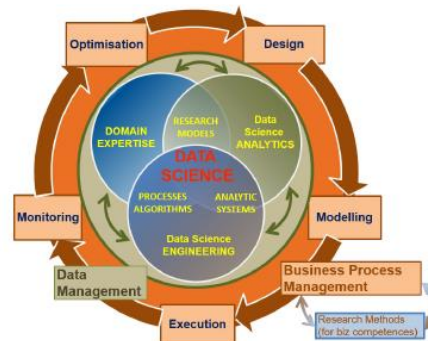
For the next ZOOM meeting (Tuesday 5th of Mai)

Have a look at the EDISON documents and find answers to the following questions:

1. What would be your definition of „Data Science“?
2. Who is the target group of the documents?
3. What are the different professional profiles and what would be your favorite job?
4. What are the differences between the „Scientific Methods” and “Business Process Management” (see below, big versions in the PDFs)?



(a) Data Science competence groups for general or research oriented profiles.



(b) Data Science competence groups for business oriented profiles.

Spezialisierung „Data Science“

- Have a look at the „Prüfungsordnung“ and other documents that can be found under: <http://www.cs.uni-frankfurt.de/index.php/de/studierende-studiengaenge.html?id=949>
- Think of questions you have!

Data Science

Data is increasingly cheap, ubiquitous, and valuable. Powerful technologies are emerging to organize and analyze this avalanche of data.

The rise of parallel processing and cloud-based storage permit real-time, interactive analysis of large-scale data sets.

The Specialization *Data Science* prepares students to be data-professionals. Working with data at scale requires distinctive new skills and tools.