

Course overview

Data modeling

Marcus Birkenkrahe

August 17, 2021

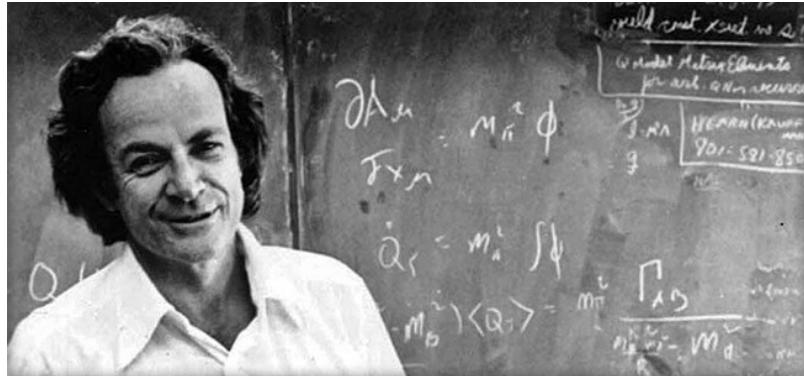
What're you going to learn today?

- Who is your lecturer?
- Who are you and what do you want?
- Which topics will we cover?
- How will we do it?
- What do you have to do to pass?
- What's next?

Who am I?



Science



- Development of WWW
- PhD theoretical particle physics
- 60 research publications
- Assoc. Ed. Int. J. of Data Science
- Ed. Board Int. J. of Big Data Mgmt.
- Scientific member d-cube@Berlin

Industry



- Executive at Accenture & Shell
- Coach and consultant
- Certified psychotherapist
- Startup mentor

Teaching



- Business informatics @HWR Berlin
- Visiting professor of data science @Lyon
- Adviser for CPU @LA
- Internship supervision

Pleasure



- Playing: Assassin's Creed Valhalla (2020)
- Reading: Waugh, Sword of Honour (1952-1961)
- Watching: The Middle (2009-2018)

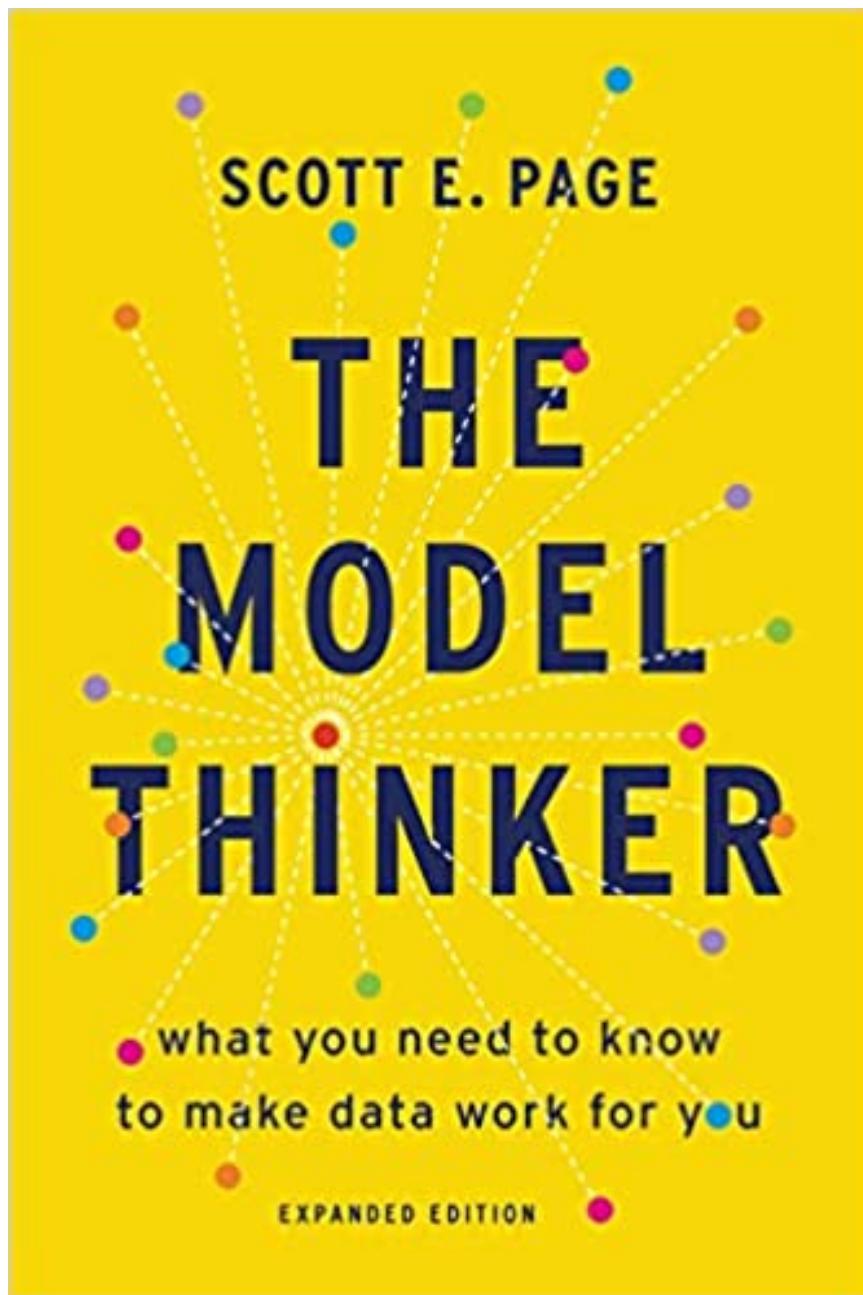
What are your expectations?

- What do you want to learn here?
- What would you like to avoid?
- What did you take away from another course?
- What did you really not like in another course?

Which topics will we cover?

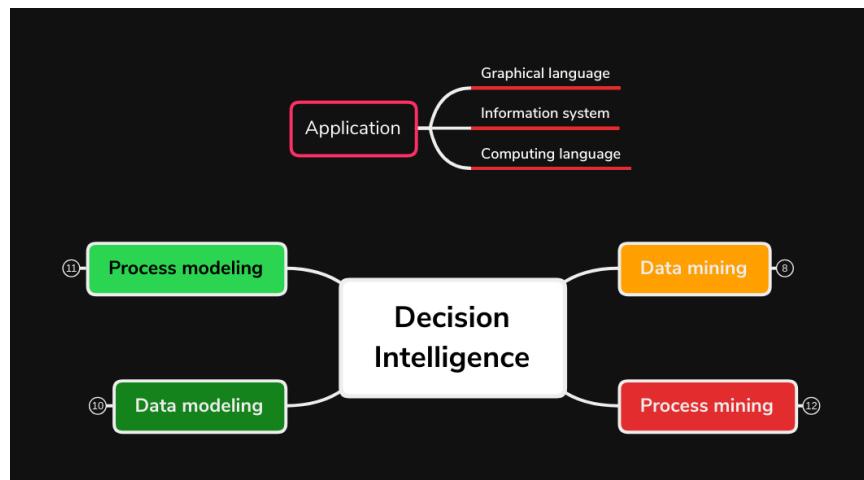
`./img/lavaflow.gif`

Many-model thinking

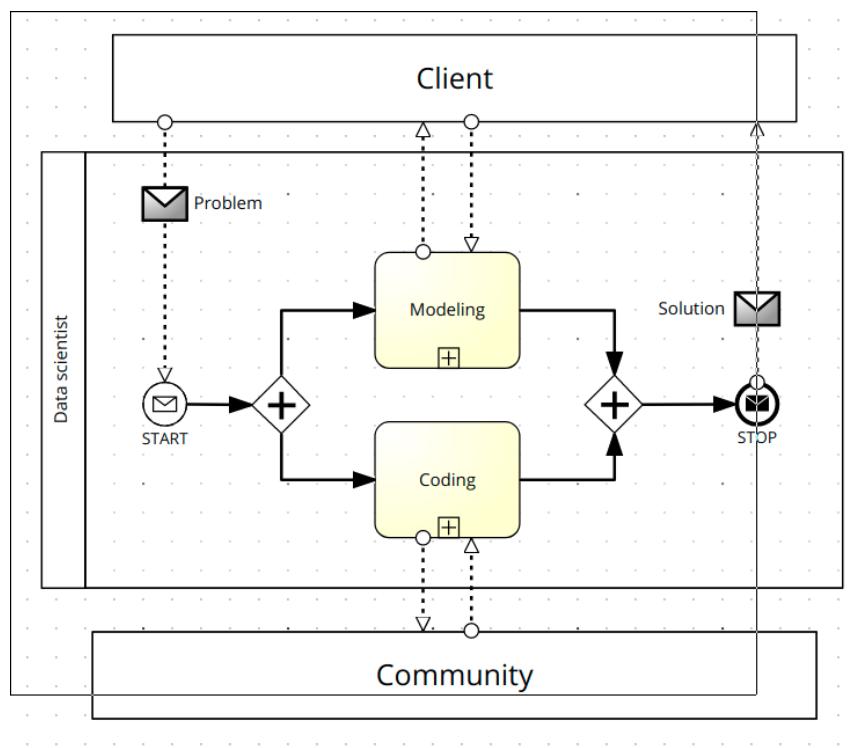


Expanded edition, Basic Books 2021

Decision intelligence

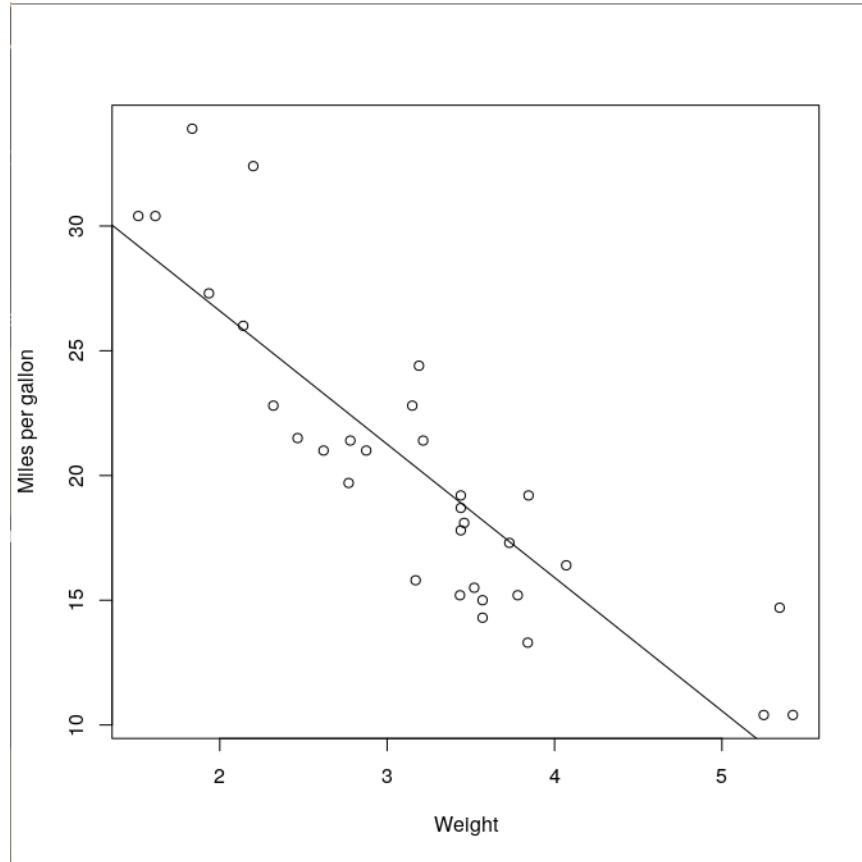


Process Modeling



Source: Signavio / 19 May test lecture

Linear models



Source: R plot

Linear regression in R

```
x <- mtcars$wt  
y <- mtcars$mpg  
plot(x,y,xlab="Weight",ylab="Miles per gallon")  
lm_model <- lm(y~x,data=mtcars)  
abline(lm_model)
```

Agile management

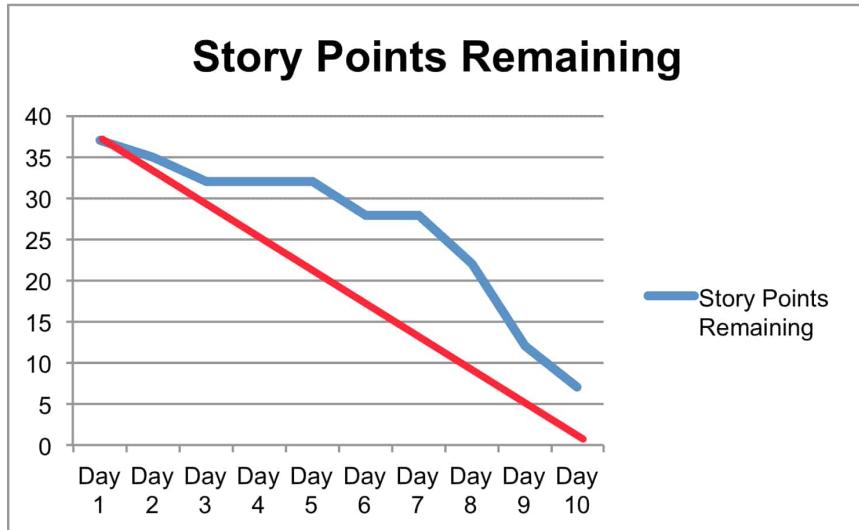


Image: Scrum burndown chart

Robotic Process Automation

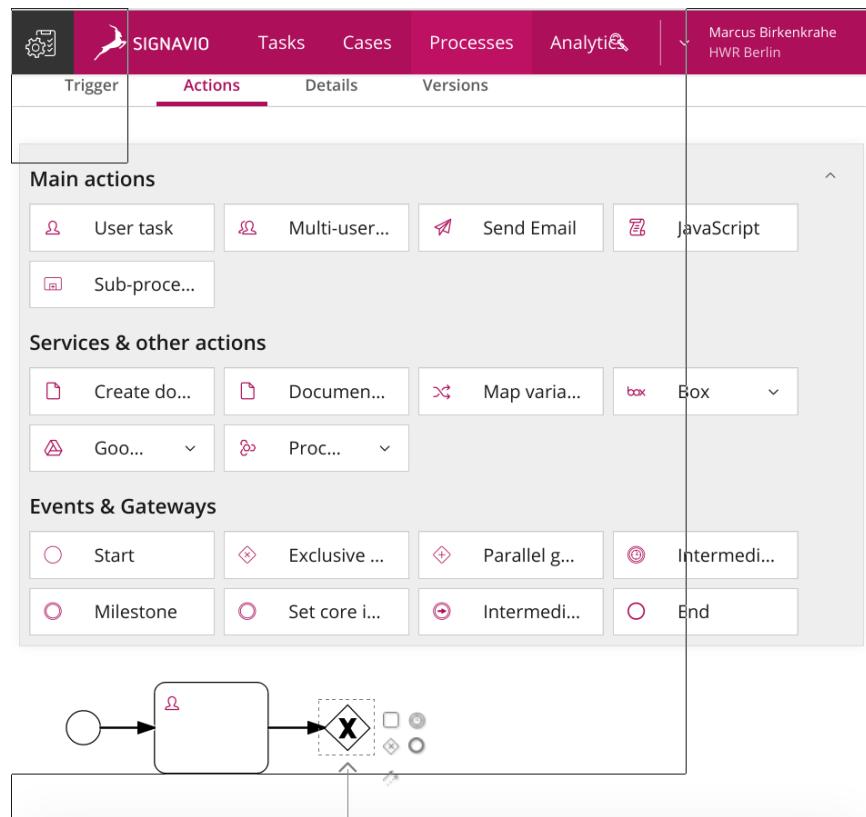


Image: Signavio Workflow Accelerator

Unified Modeling Language

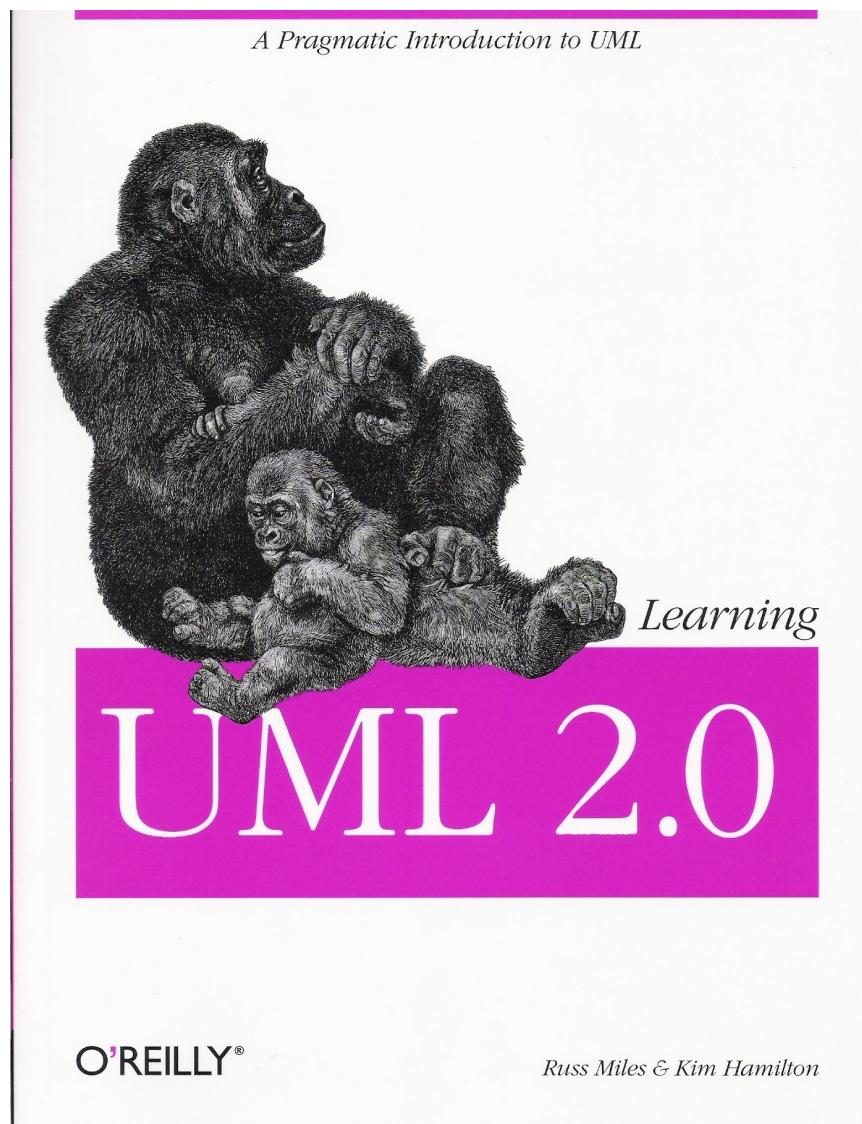


Image: Learning UML 2.0 (2006)

Process mining



Image: Celonis dashboard

Schedule (see Syllabus)



Image: Princeton U.

How will we do it?

./img/deer.gif

Classroom sessions



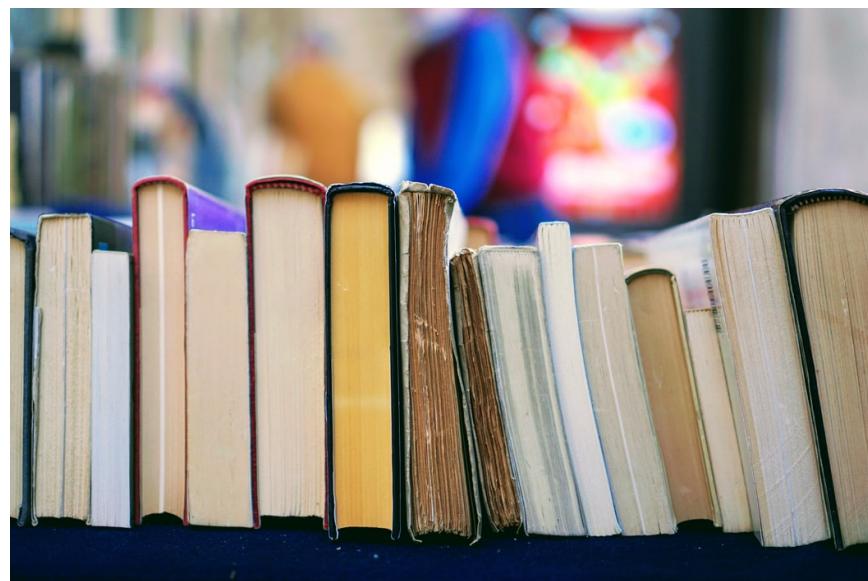
Lecture scripts with exercises (GitHub)

The screenshot shows a GitHub repository page for 'dsc101'. The URL in the browser's address bar is highlighted with a red box and labeled 'URL'. The repository owner is 'birkenkrahe'. Key interface elements highlighted with colored boxes are: 'Raise issues' (orange), 'Watch' (purple), 'Issues' (orange), 'Discussions' (green), 'Discuss' (green), and a file named '1_overview' (blue). The repository description states: 'Repository for DSC 101 - Data science methods and tools'. The README.md file contains the following content:

```
dsc101

• Repository for DSC 101 - Data science methods and tools.
• Contains Emacs Org-mode files (rendered as Markdown), Wiki, Discussion forum.
• First offered: fall 2021.
```

Reading assignments



- Image: Unsplash (@tomhermans)

Lab sessions



- Image: Unsplash (@Emin Baycan)

Stuff you bring to class



- Image: Unsplash (@Evan Demicoli)

What do you have to do to pass?

`./img/oceanrock.gif`

Weekly lab practice (> 50%)



Weekly participation (> 50%)



Final essay (> 50%)



Source: Unsplash (@Aaron Burden)

What constitutes an essay?

- IMRaD structure (video)
- Research question
- Literature review
- Methodology
- Results (e.g. glossary)
- Discussion with limitations
- References

Do you have essay examples?

- Chapters in "Model thinking"
- (Parts of) Research papers

- Scientific or industry blogs

Can I write a scientific essay?

- Keep It Simply Scientific (IMRaD)
- Read and take notes (see FAQ)
- Researchers are beginners

Final exam (> 50%)



Final exam: date TBD

What's next?

[./img/river.gif](#)

In the course

- Lecture "Decision intelligence"
- Lab discussion "many-model thinking"

- Data vs. models (2 articles)
- What is a model anyway?

Your challenges

What?	When?
Read "Many-model thinking"	Aug 19
Complete test challenge	Aug 24
List possible research questions	Sep 2
Check FAQs x 2 in GitHub	n.d.
Ask questions (class/GitHub)	n.d.

**) do this every week until December*

Any questions?

`./img/sip.gif`

This presentation is available online.