



ZBIW Zertifikatskurs Data Librarian

Modul 2: Daten strukturieren - beschreiben - wiederauffinden

2024-01-08

Björn Engelmann, Fabian Haak
TH Köln - Cologne Information Retrieval Group

Modul 2 - Daten strukturieren - beschreiben - wiederauffinden

Fabian Haak M.Sc.

Teilmodul 1: Strukturierte Daten und Metadaten

8.1. - 22.1.

Daten Strukturieren und Aufbereiten – Wir suchen uns Daten, machen sie sauber und bringen sie in Form!



Björn Engelmann M.Sc.

Teilmodul 2: Suchmaschinentechnologie und Information Retrieval

22.1. - 5.2.

Daten Beschreiben und wiederauffinden – Wir bauen uns eine Suchmaschine!



Termine

- Virtueller Modulauftakt:
08. Januar 2024, 16.00 - 17.30 Uhr
- Virtuelle Fragestunde:
22. Januar 2024, 16.00 - 17.30 Uhr
- Präsenztage vor Ort; TH Köln, Ubierring 48, Raum 211:
05. Februar 2024, 10.00 - 16.00 Uhr

An abstract painting depicting a dense, chaotic scene of electronic waste. The composition is filled with various shapes and colors, primarily in shades of teal, blue, and orange. The central focus is a large, rectangular object with a bright orange front panel, which appears to be a piece of electronic equipment, possibly a computer monitor or a server component, that has been dismantled or is in a state of decay. This central object is surrounded by numerous smaller, fragmented pieces of electronic components, including what looks like a keyboard, a mouse, and various circuit boards. The background is a complex, layered texture of paint splatters, brushstrokes, and overlapping shapes, creating a sense of depth and movement. The overall effect is one of overwhelming complexity and a commentary on the sheer volume of electronic waste generated in the digital age.

Daten sind überall!

2024-01-08

Björn Engelmann, Fabian Haak
TH Köln - Cologne Information Retrieval Group



Data Science = den Datenberg beherrschen

2024-01-08

Björn Engelmann, Fabian Haak
TH Köln - Cologne Information Retrieval Group

Data Science Process

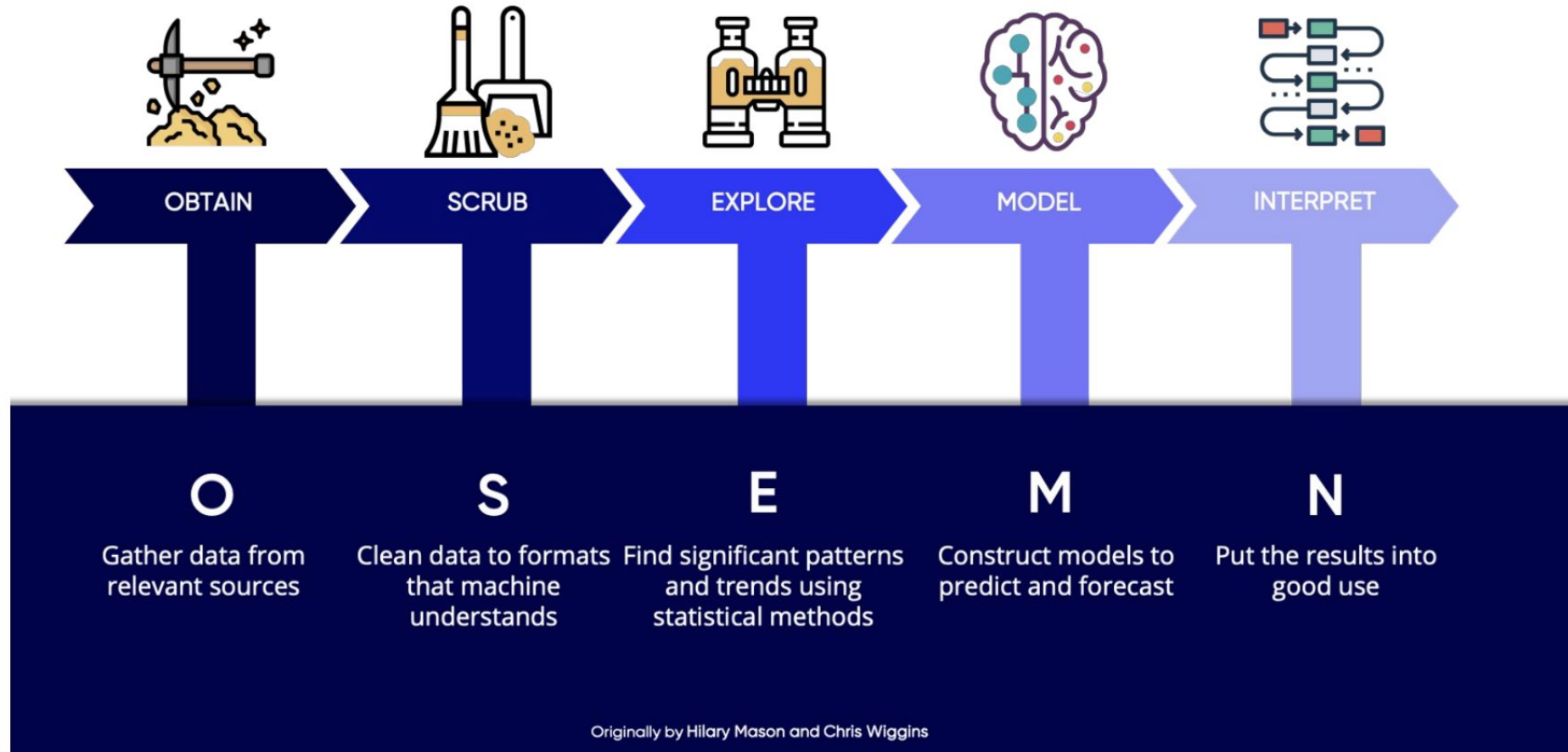


Image source: <https://towardsdatascience.com/5-steps-of-a-data-science-project-lifecycle-26c50372b492a>

2024-01-08

Björn Engelmann, Fabian Haak

TH Köln - Cologne Information Retrieval Group

Seite 6



BibSonomy



pandas



Erwartungen an Euch!

Wir möchten von euch, dass...

- ... ihr mit Hilfe der Tutorials **erste Gehversuche** in Python, Pandas und Pyterrier unternimmt, und
- ... mit “**echten Daten**” einmal den Prozess durchläuft damit ihr
- ... ein Gespür für den **OSEmn-Workflow** bekommt
- ... und dabei eure Fähigkeiten im Einsatz von **State-of-theArt-Werkzeugen** etwas weiter festigt.

Erwartungen an Euch!

Uns ist bewusst, dass...

- das alles sehr viel und überwältigend sein kann
- noch kein Meister vom Himmel gefallen ist

Wir helfen euch da durch!

- Schritt für Schritt, Wiederholung der Basics aus Modul 1
- Hilfe (zur Selbsthilfe) gerne im **Forum** und in der Sprechstunde.
- **Präsenztage** mit gemeinsamen Übungen, Wiederholungen, Fragestunde und vielen Tipps und Tricks!