



CORRELAID
GOOD CAUSES. BETTER EFFECTS.

Algorithmen sind nicht alles

was Datenprojekte (nicht) brauchen

Quellenangabe/Lizenz (sofern nicht anders angegeben): [Algorithmen sind nicht alles - was Datenprojekte \(nicht\) brauchen](#), 2022-10-06, Frie Preu und Nina Hauser, CorrelAid e.V., [CC-BY 4.0](#)

Sofern nicht anders angegeben, stehen alle hier verlinkten Materialien stehen auch unter CC-BY 4.0: [Titel], 2022, CorrelAid e.V., [CC-BY 4.0](#)

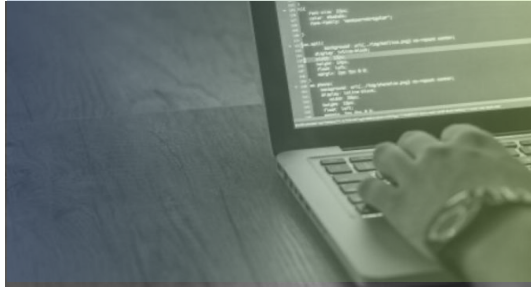
CorrelAid



Wir sind ein deutschlandweites Netzwerk von über 2,200 Data Scientists, die die Welt durch Data Science verbessern wollen.

#MetaWeltretter

Unsere Mission



DATA4GOOD PROJEKTE

Wir führen pro-bono Datenanalyseprojekte für gemeinnützige Organisationen durch.



EDUCATION

Wir vernetzen engagierte sozial denkende Datenanalyt:innen und bieten ihnen Möglichkeiten ihr Wissen anzuwenden und zu erweitern. Außerdem vermitteln wir engagierten Menschen von gemeinnützigen Organisationen grundlegende Data Literacy Skills.



DIALOG

Wir treten in den Dialog über den Wert und Nutzen von Daten und Datenanalysen für das Gemeinwohl.

Über mich: Frie Preu



- Studium der Politikwissenschaft, dann M.Sc. Social and Economic Data Analysis
- 2 Jahre IT Consulting
- Mitbegründer*in von CorrelAid, seit 2020 hauptamtlich

Über meinen ersten R* Kurs 2012/13: “Der Einstieg in R war für mich **sehr schwer**. Ich war vorher **kein "Nerd"** gewesen: Ich hatte angefangen, Politikwissenschaft zu studieren, weil ich ein wirklich politischer Mensch war und weil ich mich für Politik interessierte, **nicht weil ich komplizierten Code schreiben wollte**. Deshalb hatten ich und R einen schwierigen Start: Ich habe **kaum etwas verstanden**. Tatsächlich habe ich nach diesem ersten Kurs noch 1-2 Jahre lang Stata statt R benutzt, bevor ich R eine weitere Chance gab und es lieben lernte.”

*R ist eine Programmiersprache, die in Data Science verwendet wird



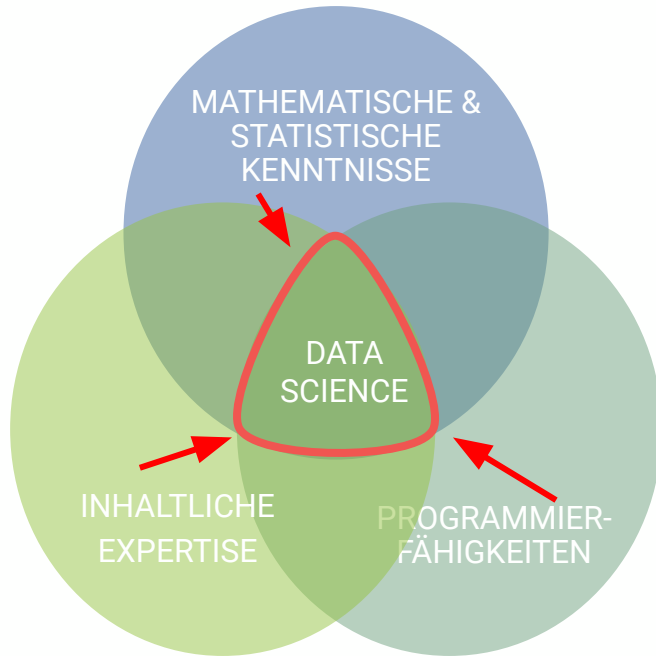
CORRELAID

GOOD CAUSES. BETTER EFFECTS.

Warm-up

Worüber sprechen wir?

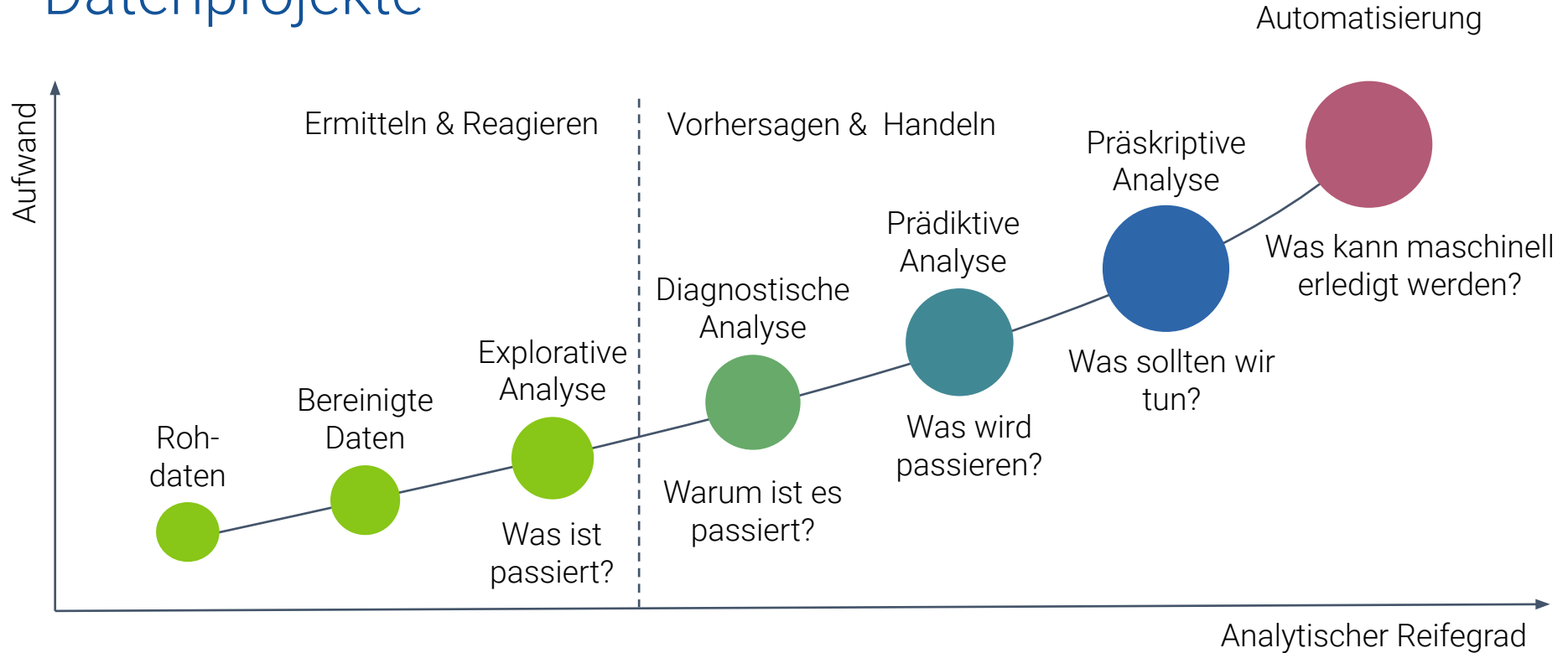
Was ist Data Science?



Data Science ist ...

- Data Science (zu dt. Datenwissenschaft) ist ein Werkzeug, eine Denkweise und ein Kompetenzbereich
- Es ist ein interdisziplinäres Fachgebiet mit zahlreichen Anwendungsfällen
- Mögliche Fragestellungen:
 - Was macht unsere gesellschaftliche Herausforderung aus?
 - Wie können wir bessere Entscheidungen treffen?
 - Wie können wir effizienter und effektiver arbeiten?

Datenprojekte





CORRELAID
GOOD CAUSES. BETTER EFFECTS.

Disclaimer: Projekte ersetzen keine Organisationsentwicklung

...können sie aber anstoßen oder katalysieren!*

*Quelle: [Georg Förster, 2022-09-28, Twitter.](#)



Was datengetriebene **Organisationen** brauchen...

 Teammanagement & Kommunikation

Daten"kultur"

- + Unterstützung von Leitungsgremien
- + "Buy-in" der ganzen Organisation
- + Kultur: "Daten" als Selbstverständlichkeit

 **Daten**

  Ressourcen


Personen

Was Daten**projekte** brauchen...

 Ideen und klarer Umfang

 (Daten)*

 Projektmanagement
und Kommunikation

  Ressourcen


Personen

*Datenerhebung kann auch Gegenstand des Projekts sein

Was Daten**projekte** brauchen...

 **Ideen und klarer Umfang**  (Daten)*

 **Projektmanagement
und Kommunikation**

  Ressourcen


Personen

*Datenerhebung kann auch Gegenstand des Projekts sein



CORRELAID

GOOD CAUSES. BETTER EFFECTS.

Ressourcen: Zeit & Geld

Wie sieht das bei euch aus?



CORRELAID

GOOD CAUSES. BETTER EFFECTS.

Ideen & Umfang*

Sinnvolle Datenprojekte definieren

*en: Scope

Aus unserer Arbeit: CorrelAid Projekte

Verantwortung

Pragmatismus

wirkungsvoll

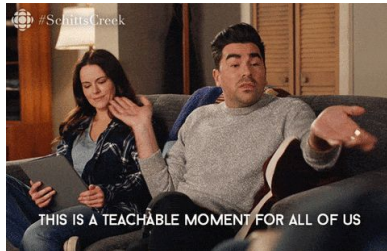
Was wir tun hilft der NPO mit ihrer Mission und stärkt ihre Wirkung.

nützlich und nachhaltig

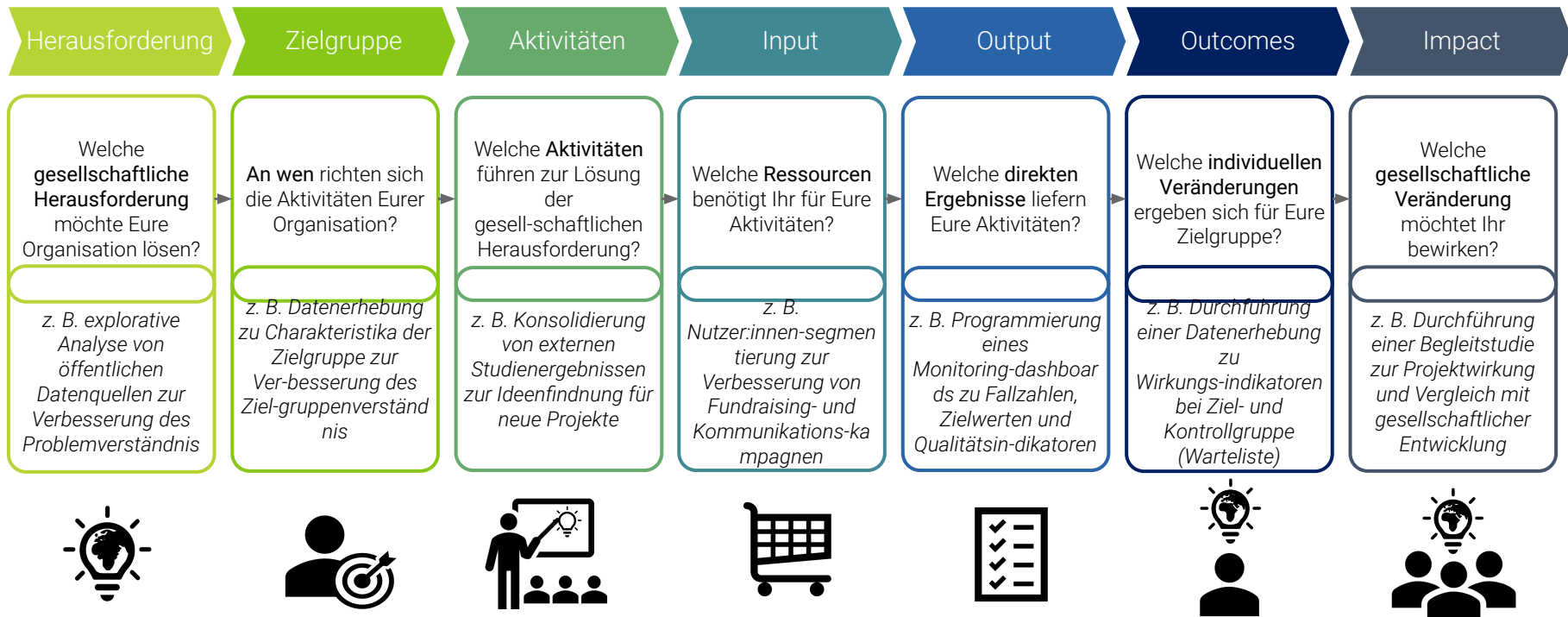
Die NPO kann die Ergebnisse des Projekts nutzen. Das Projekt hat eine Wirkung über den Projektzeitraum hinaus.

umsetzbar

Es ist realistisch, das Projekt mit einem Skilled-Volunteering Ansatz umzusetzen und gleichzeitig eine gute Lernerfahrung für alle Teilnehmer*innen..



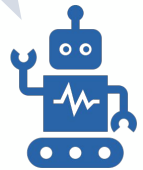
Wirkungsvolle Ideen finden? Orientierung an Wirkungslogik...



...oder aus damit verbundenen strategischen Zielen



Manchmal ist es schwer das derzeitige strategische Ziel zu benennen? Dann macht es Sinn sich zu fragen: "was beschäftigt uns derzeit immer wieder? Wo drückt der Schuh?"



Hinter Euren Datenprojekten steht in der Regel eine strategische Zielsetzung

Ideenfindung für Datenprojekte



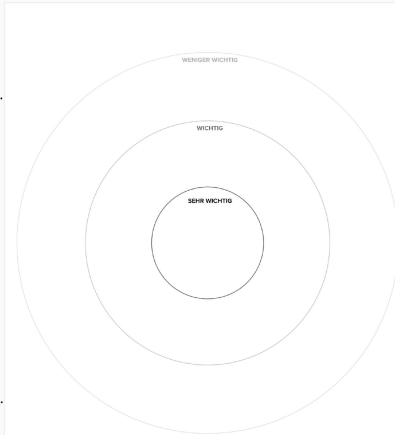
Welche strategischen Ziele verfolgt Ihr derzeit als Organisation? Welche davon können durch Daten unterstützt werden? Wie priorisieren wir die verschiedenen, datenbezogenen Ziele?

1 Sammelt hier strategische Ziele

Strategische Ziele

Diagram showing three colored squares (green, dark green, blue) representing strategic goals.

3 Bestimmt Prioritäten

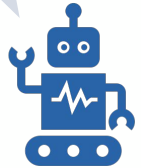


2 Identifiziert ziel- und datenbezogene Herausforderungen

Datenbezogene Herausforderungen

Diagram showing an empty box for data-related challenges.

“Den Wert von Datenanalyse aufzeigen” kann ein strategisches Ziel eines Datenprojekts sein. Idealerweise gibt es aber auch noch weitere Ziele, die das Projekt unterstützt.



Beispiel: Nützlichkeit, Nachhaltigkeit, Umsetzbarkeit

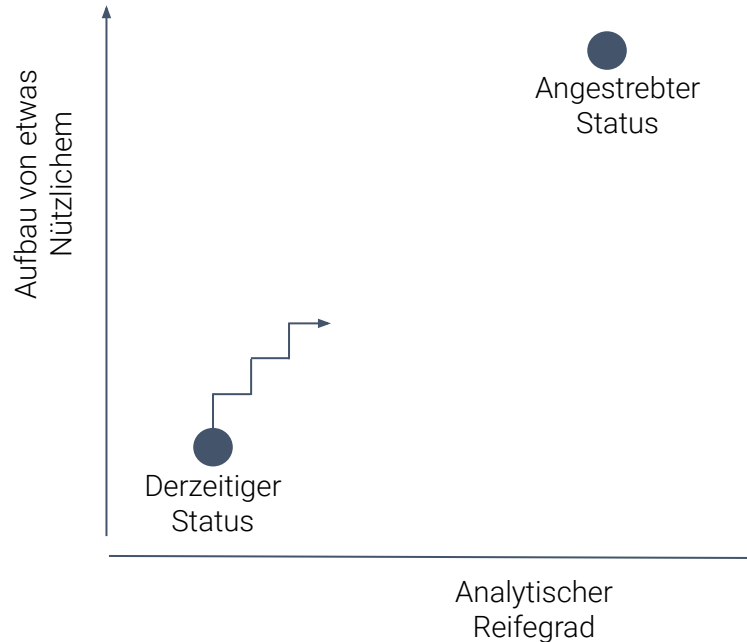
	Dashboard mit live Daten	(teil)automatisierte Analyse (z.B. PDF oder HTML)
wirkungsvoll	Bessere Entscheidungsgrundlage jederzeit	Bessere Entscheidungsgrundlage auf Nachfrage / zu bestimmten Intervallen
nützlich	Interaktivität	Eingeschränkte Interaktivität
umsetzbar	Hoher Aufwand in User Experience (UX) und Design, Hosting-Fähigkeiten nötig, Authentifizierung	UX wichtig, aber einfacher; Datenanalyst*in-Fähigkeiten, ggf. Abstimmung mit IT nötig
nachhaltig	Applikation muss gehostet werden, Wartung	Integration in Prozesse etc., Einrichtung auf Computern von MA; je nach Automatisierungsgrad Person oder Server nötig zum Ausführen

Fragen: Nützlichkeit, Nachhaltigkeit, Umsetzbarkeit

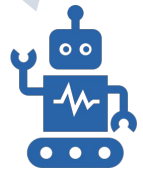
- Wird das Monitoring-Dashboard wirklich täglich angeschaut werden? Oder reicht ein monatlicher Bericht, der im Meeting präsentiert wird?
- Bei Applikationen: Wer übernimmt die Wartung der Applikation? Wer kann befähigt werden, Anpassungen vorzunehmen? Wer updatet die Daten? Haben die Personen dafür die benötigte Zeit?
- Wie werden Serverkosten langfristig getragen?
- Gibt es einen "Buy-In" / ein Commitment von den anvisierten Nutzer*innen? Wie sehr wollen sie das?
- Ist das Projekt mit den verfügbaren Personen umsetzbar? Ist es zu überfordernd? Können wir die technische Komplexität oder den Umfang reduzieren?

Umfang (en: Scope)

- Spannungsfeld zw. "Leuchtturmprojekte" und "klein & realistisch anfangen und schnelle Erfolge"
- Große Projekte inkrementiell gestalten
- Klar inhaltlich begrenzen -> Gefahr des "Feature Creep"



Gestaltet die Inhalte des Projekts so, dass im Prozess langfristig nützliche Nebenprodukte anfallen!



Tipps zum Prozess

- Ideen identifizieren im Gespräch ([Template Scoping \(en\)](#) oder Mural Templates*)
- Voraussetzungen klären: Datenzugang, [DSGVO](#)
- Schlanke [Projektbeschreibung](#) schreiben: Überblick, Partner*innen, Problem und anvisierte Wirkung, Ziel und Aktivitäten, Profil der Teilnehmer*innen, Schritte bis zum Projektstart
- Commitments einholen von Stakeholdern!

*Kommt gerne mal in die [Datensprechstunde](#), wenn ihr Interesse habt!



CORRELAID

GOOD CAUSES. BETTER EFFECTS.

Personen

Data Scientists und wie sie arbeiten

Wer sind diese Data Scientists?



Artificial Intelligence engineer
Machine Learning engineer
Data architect
Machine Learning Ops
...

Tipps:

- Keine Beschränkung auf bestimmte Studiengänge
- Offen für Quereinsteiger*innen und Personen ohne Universitätsausbildung
- Konkrete Projekte > formelle Bildung
- Faires Gehalt, ggf. Kompensation durch Benefits
- Ehrlichkeit bzgl. Aufgaben: wie viel ist "politische" Arbeit / Projektmanagement?

Zur Orientierung im Datenbegriffdschungel:
[Das kleine ABC der Daten](#)

Persönliche Entwicklung

IT / Data Science ist hoch dynamisch → Zeit um am “Puls der Zeit” zu bleiben, sehr wichtig!

Wo/was könnt/wollt ihr investieren? Zeit? Geld?



**KICKER,
KOSTENLOSE
GETRÄNKE,
SITZSÄCKE...**

**RESSOURCEN
FÜR WEITERBILDUNG,
4+1 MODELL,
80% STELLEN,
AUSTAUSCH**

Programmieren braucht Fokus

Koordinierende/managende Rolle

ethics committee 9:15 – 10:15am
tidytuesday on website 10:15 – 11am
daniel antworten, 11am
emails beantworten, 11:15am
R Lernen Q3 12pm, https://us02web.zoom.us/j/81174653624
datenschutzerklaerung obs 1:30 – 2:30pm

FOCUS



Tipps:

- Meetings in “Randzeiten”
- Meetings mit Zweck (siehe Scrum)
- “Diese Meeting hätte eine Email sein können”
- Asynchronität > Telefon
 - Vorteil: Dokumentation
- **Nachfragen, wann und wie Personen gut arbeiten**



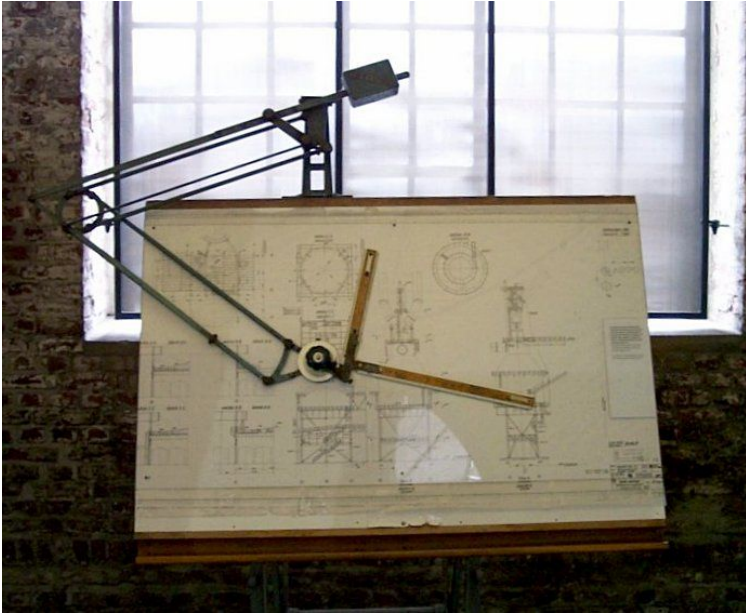


CORRELAID

GOOD CAUSES. BETTER EFFECTS.

Projektmanagement & Kommunikation

Klassisches Projektmanagement und IT Projekte



[Historische Zeichenmaschine mit doppeltem Parallelogrammgestänge und Ausgleichsgewicht](#), [Attribution-Share Alike 3.0 Unported](#)

Aus unserer Arbeit: CorrelAid “macht kein Scrum”

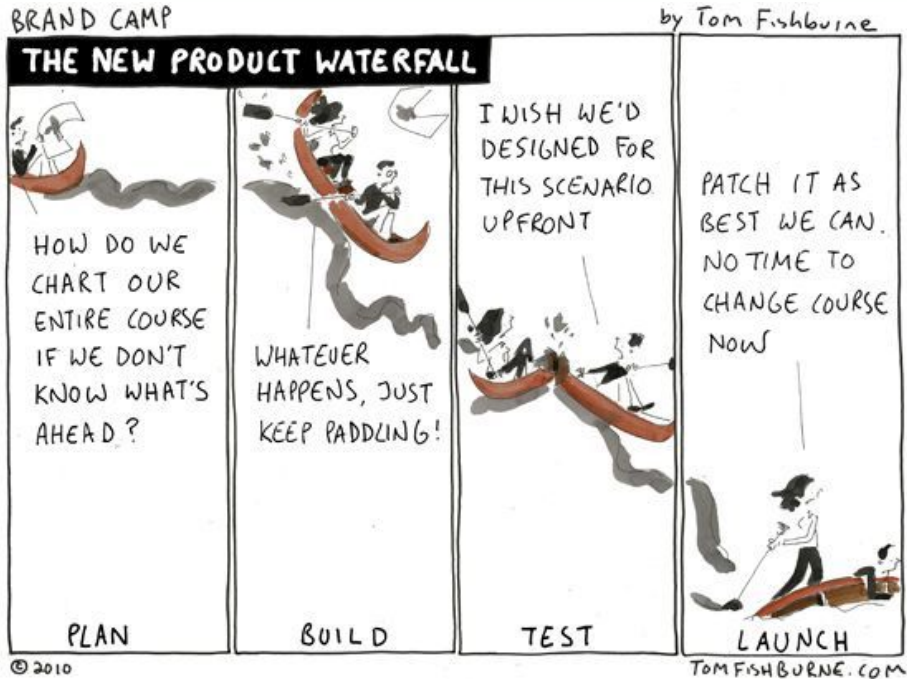
→ **wir “machen nicht Scrum”** aber lassen uns davon inspirieren: “Es ist zwar möglich, nur Teile von Scrum zu implementieren, aber das Ergebnis ist nicht Scrum”.

Inspiration:

- Sprints: reguläre Meetings/Events
- Retrospektive: regelmäßig über Zusammenarbeit reflektieren
- ...

Aus unserer Arbeit

Best Practice: regelmäßig, transparent und offen kommunizieren



Warum ist das wichtig?

- Veränderung ist unausweichlich in IT Projekten (gerade Daten!)
- Regelmäßige Kommunikation hilft dabei, mit Veränderung umzugehen
- Frühe Feedbackprozesse und Anpassungen vermeiden Frust am Ende

Typische Stolpersteine in Datenprojekten

- Zustand der Daten nicht geeignet für Projektziel
- kein/zu wenig Puffer für Datenprobleme
- Zu wenig Budget für Personen
- Keine Schnittstelle zu Nutzer*innen / Wasserfall
- Keine eigene Expertise
- Zu wenig "Buy-In"



Tipps

Do:

- Retrospektiven!
- “Vergabekompetenzen” und Data Literacy für Projektleitungen
- Ggf. Projektmanager*innen im agilen Arbeiten weiterbilden & Kapazitäten schaffen
 - Scrum Master wenn internes DS Team
 - Product Owner wenn externes DS Team

Don't:

- Data Scientists in politischen Kleinkämpfen verheizen
- Bei Outsourcing:
 - Nur einen Plan am Anfang entwickeln und “über den Zaun schmeißen”
 - Gar kein Projektmanagement selbst machen

Was Daten**projekte** brauchen...

 Ideen und klarer Umfang

 (Daten)*

 Projektmanagement
und Kommunikation

  Ressourcen


Personen

*Datenerhebung kann auch Gegenstand des Projekts sein



CORRELAID
GOOD CAUSES. BETTER EFFECTS.

Danke für eure Aufmerksamkeit!

Was sind eure Erfahrungen?

Beschreibungstext

Titel: Algorithmen sind nicht alles - was Datenprojekte (nicht) brauchen

Beschreibung: Künstliche Intelligenz als Schlagwort ist in aller Munde und findet zunehmend auch in der Zivilgesellschaft und im sozialen Sektor Anwendung. Doch was brauchen Datenprojekte abseits von Algorithmen, um strategisch sinnvoll, erfolgreich und nachhaltig zu sein? Frie Preu von CorrelAid teilt Erfahrungen und (Miss-)erfolgskonzepte aus der langjährigen Arbeit des Data4Good Netzwerks mit über 50 gemeinnützigen Organisationen und gibt Einblicke in die Motivationen und Arbeitsweisen von "Data Scientists", basierend auf eigenen Erfahrungen in der freien Wirtschaft.

Format: Vortrag, unterbrochen von Diskussionselementen

Lernziele

- **TN wissen welche “Data Science” Rollen es gibt und verstehen wie Data Scientists gut arbeiten können**
- **TN kennen verschiedene Arten von Data Science Projekten und deren Komplexität**
 - Explorativ / einmalige Analyse
 - Monitoring
 - Prädiktiv: Vorhersagen
- **TN verstehen wie sie sinnvolle Projekte identifizieren und definieren können**
 - Wirkungslogik
 - Projekt template mural
- **TN kennen typische Pitfalls von Data Science Projekten**
 - Keine Schnittstelle -> Wasserfall
 - Kein Puffer für Datenbereinigung / Model fitting
 - Daten nicht in einem akzeptablen Zustand
 - Zu wenig Budget