

CAS
DATENJOURNALISMUS
2021-2022

CAS DATENJOURNALISMUS 21-22

Sie haben Erfahrung als Journalistin oder Journalist, aber nicht im Programmieren.

Nun wollen Sie lernen, wie man Daten beschafft, strukturiert, analysiert und visualisiert, um damit spannende Geschichten zu erzählen.

Der CAS Datenjournalismus leitet Sie dabei an – praxisnah und effizient. Nach einem 10-tägigen Bootcamp, in dem die Grundlagen vermittelt werden, folgt eine 4-monatige Praxiszeit, während der sie konkrete Projekte entwickeln – gecoacht von den Dozenten. An 9 Vertiefungstagen – verteilt über die Praxiszeit – lernen Sie neues und tauschen sich über die Projekte aus.

- Zielgruppe Journalistinnen und Journalisten,
- _ mit Berufserfahrung (mindestens zwei Jahre Tätigkeit als Freelancer oder auf einer Redaktion)
 - _ die in den Datenjournalismus einsteigen möchten (keine Programmierkenntnisse erforderlich);
 - _ die lernen möchten, datengetriebene Recherchen umzusetzen

- Voraussetzungen
- _ Mindestens zwei Jahre journalistische Praxis;
 - _ Grundkenntnisse im Umgang mit Statistik (oder die Bereitschaft, sich diese vorgängig anzueignen);
 - _ Zeit und Praxisfeld, die es erlauben, konkrete datenjournalistische Projekte zu realisieren und zu publizieren (Absprache mit Redaktion, dass neben den Kurstagen Zeit für Projekte zur Verfügung steht);
 - _ eigener Laptop (Apple- oder Linux-Geräte sind zu bevorzugen), auf dem beliebige Programme installiert werden können (kein Geschäftslaptop ohne Administratorechte) und auf dem Sie sich auskennen. Bitte Geräte, die in den Kurs mitgebracht werden, vorgängig kennenlernen.
 - _ für den Kurs ist keine vorgängige Programmiererfahrung nötig (aber Bereitschaft, vorgängig einführende Tutorials durchzugehen).

- Lernziele
- Der CAS Datenjournalismus ist ein konsequent praxisorientierter Kurs.
- Sie lernen
- _ interessante Datensätze zu erkennen;
 - _ datenjournalistische Projekte zu formulieren und zu dokumentieren
 - _ Daten im Internet zu beschaffen (scrapen);
 - _ Daten zu strukturieren, so dass sie gut analysiert werden können;
 - _ dafür nötige bestehende Programme (etwa Tabellenkalkulationsprogramme, Visualisierungsprogramme etc.) anwenden zu können;
 - _ dafür nötige Programme mit der Programmiersprache Python selbst zu schreiben;
 - _ selbstständig Programmierfehler zu finden oder Probleme im Austausch mit der Open-Source-Community zu lösen;
 - _ strukturierte Daten einfach zu visualisieren, so dass Stories erkennbar werden;
 - _ und vorallem lernen Sie, darauf basierend spannende Stories zu erzählen.

- Methoden
- In einem 10-tägigen Bootcamp lernen Sie von Expertinnen und Experten – alle selbst als Datenjournalisten oder Programmierer tätig – kompakt die Grundlagen mit vielen konkreten Übungen. In diesen zehn Tagen entwickeln Sie auch ein erstes datenjournalistisches Projekt. In der anschliessenden Praxiszeit von vier Monaten realisieren Sie eigene datenjournalistische Projekte, werden dabei von den Dozenten gecoacht und kommen an neun Vertiefungstagen zusammen, um die Projekte vorzustellen, Probleme zu diskutieren und neue datenjournalistische Techniken zu erlernen.

Dauer/Termine Netto 20 Tage (zwischen dem 06. September bis am 16. März 2022) verteilt auf:

- Ein Bootcamp (10 Tage)
- Vertiefungstage in der Praxiszeit (9 Tage)
- Abschlusstag (1 Tag)

Für die Abschlussarbeit bleibt Zeit vom 07. Januar bis am 14. Februar 2022.

Der Kurs umfasst 150 Stunden Präsenzzeit und 120 Stunden selbständige Arbeit.
Der Unterricht (Präsenzzeit) findet am MAZ Luzern statt.

Bootcamp: 06. – 17.09.2021 (10 Tage);
Vertiefungstage 1+2: 04. – 05.10.2021
Vertiefungstage 3+4: 28. – 29.10.2021
Vertiefungstage 5+6: 18. – 19.11.2021
Vertiefungstage 7+8: 06. – 07.12.2021
Vertiefungstag 9: 06.01.2022
Abschlussarbeit bis 14.02.2022
Abschlusstag und Feedback: 16.03.2022

Das Selbststudium umfasst vorbereitende Lektüre und die Arbeit an den eigenen datenjournalistischen Projekten. Ziel ist es, 3-4 solcher Projekte während der Praxiszeit zu realisieren.
Die Abschlussarbeit wird im Anschluss an die Vertiefungstage verfasst. Sie wird unter quasirealen Bedingungen durchgeführt mit dem Ziel, sie in einem Schweizer Medium zu publizieren.

Aufnahmeverfahren Das MAZ entscheidet über die Aufnahme auf Grund des eingereichten Dossiers und bei Grenzfällen nach einem persönlichen Gespräch.

Bewerbung mittels Dossier. Dieses umfasst:

- _ ausgefülltes Deckblatt
- _ Lebenslauf
- _ Kopie des höchsten Abschlusses
- _ Motivationsschreiben (1500-2000 Zeichen)
- _ 2-3 Ideen für datenjournalistische Projekte

Anmeldefrist 06. August 2021 (Berücksichtigung der Anmeldungen nach Eingangsdatum)

Abschluss Wer den Lehrgang erfolgreich abschliesst, erhält ein MAZ-Zertifikat (12 ECTS - European Credit Transfer System).
Erfolgreich abschliessen heisst: Besuch von 90 Prozent des Präsenzunterrichts, bestandene Abschlussarbeit.
Bei Nichtbestehen kann die Abschlussarbeit einmal wiederholt werden.

Leitung Barnaby Skinner, Ressortleiter Visuals & Data der NZZ; Alexandra Stark., MAZ-Studienleiterin

Dozierende Barnaby Skinner, Ressortleiter Visuals & Data der NZZ
Simon Schmid, Datenjournalist, Republik
Thomas Ebermann, Dr. oec., Universität St. Gallen, Programmierer, CH Media

Anzahl Teilnehmer Max. 12

Kosten CHF 8'500.- exkl. Gebühr für Abschlussarbeit (zusätzlich CHF 300.-).
(Die Kosten liegen unter den Vollkosten von 9'500.-, da der MAZ-Innovationsfonds den CAS 21/22 unterstützt.)

Kursort MAZ – Die Schweizer Journalistenschule, Murbacherstrasse 3, 6003 Luzern

Kurs-Nr. J113036

Administration Claudia Meier, claudia.meier@maz.ch, Telefon +41 (0)41 226 33 40

THEMEN UND TERMINE

Grundlagen des Datenjournalismus und Programmieren mit Python Bootcamp (06.09. – 17.09.2021)

Idee: Es kann hilfreich sein, die Abende für allfälliges Weiterarbeiten freizuhalten.

06. September 2021

Themen finden, datenjournalistische Recherchen planen und Github kennenlernen

Barnaby Skinner, Alexandra Stark

Anhand von Best- und Worst-Practice-Beispielen erkennen wir, dass das storyorientierte Denken gerade beim Datenjournalismus, der in aufwändige technische Arbeit mündet, von Beginn weg wichtig ist. Wir lernen, welche Datensätze für journalistische Arbeit interessant sind und wie eine datenjournalistische Story geplant, laufend überprüft und dokumentiert wird. Zudem üben wir die Grundsätze des datenjournalistischen Arbeitens mit Tabellenkalkulationsprogrammen (Google Spreadsheets, Excel, Numbers etc.). Dabei erkennen wir auch die Grenzen dieser Tools und somit den Punkt, wo eigenes Programmieren nötig wird. Zudem wird das Kurs-Repository auf Github vorgestellt.

07., 08. und 09. September 2021

Python: Grundbegriffe

Thomas Ebermann (+ Technical Assistant)

Der Dozent schafft Übersicht über die Programmiersprachen (R, C++, JavaScript), und erklärt, warum Python so gross geworden ist. Darauf führt er das Colab von Google ein und leitet erste eigene Schritte mit Python an. Die Teilnehmenden...

- erhalten eine Einführung in Variables, data types, lists und dictionaries;
- lernen Zusatz-Module und -Libraries kennen;
- lernen die Arbeit mit for loop; while Loop; if/else/elif; functions und list comprehensions

Sie erkennen, für welche Formen von Daten (strukturierter Datensatz direkt auf Datenträger; unstrukturierter Datensatz auf Datenträger; offen im Internet abfragbarer Datensatz; nur durch Interaktion im Internet abfragbarer Datensatz) diese Programme und Befehle hilfreich sind, und fügen sie in den datenjournalistischen Arbeitsablauf ein (Datenabfrage; Datenstrukturierung, Datenanalyse und Datenvisualisierung). Wir lernen zudem erste Strategien kennen, wie wir Programmierfehler finden (len; print etc.) und wie wir Stackoverflow für die Problemlösung einsetzen können.

10. September 2021

Kennenlernen des Terminals, Installation der eigenen Programmierungsumgebung und Üben der Arbeitsabläufe

Barnaby Skinner (+ Technical Assistant)

Die Teilnehmenden lernen die Grundfunktionen des Terminal kennen und installieren die Software Anaconda. Dies erlaubt es, lokal auf den Rechnern zu programmieren und trotzdem die nötigen Softwarepakete von Python bequem zur Verfügung zu haben. Zudem vertiefen sie die Kenntnisse von Github und verbessern den Arbeitsablauf am Laptop. Als Hausaufgabe gibt es einen strukturierten Datensatz, den die Teilnehmenden übers Wochenende mit einem Tabellenkalkulationsprogramm analysieren.

Wochenende

13. und 14. September 2021

Requests, APIs, Scraping mit Python (Beautifulsoup)

Barnaby Skinner (+ Technical Assistant)

Oftmals machen die eigenen Daten die besten Geschichten: Sie sind exklusiv und man weiss genau wie gut - oder schlecht - die Datenqualität tatsächlich ist. Scraping gehört deshalb zu den fundamentalen Kenntnissen von Datenjournalisten.

Die Teilnehmenden...

- erhalten einen Überblick über die Scraping Module [Beautiful Soup 4](#), [lxml](#), Scrapy;
- werden ins Auszeichnen mit HTML eingeführt und lösen eine erste Scraping-Übung an einem selbst konstruierten HTML-Beispiel;
- erkennen typische Bauweisen von Websites, die fürs Scraping genutzt werden können;
- üben ausgewählte Funktionalitäten der verschiedenen Libraries;
- erkennen, für welche Formen von Daten BeautifulSoup hilfreich ist, fügen es in den datenjournalistischen Arbeitsablauf ein und überlegen, was die Programme ihren Projekten bringen;
- lernen weitere Debugging-Strategien kennen und setzen Stackoverflow für die Problemlösung ein.

15. und 16. September 2021

Daten-Analyse und Visualisierung mit Python (Pandas)

Simon Schmid

In diesem Kursteil vertiefen wir die Datenanalyse. Dazu benutzen wir einerseits die erweiterten Funktionen von Tabellenkalkulationsprogrammen und lernen andererseits Pandas kennen, eine Statistik-Library, die für Börsenmakler entwickelt wurde. Mittlerweile aber von Wissenschaftlern aus der ganzen Welt für ihre unterschiedlichsten Forschungen eingesetzt wird. Unsere Ergebnisse visualisieren wir in Pandas und in weiteren Tools, mit denen sich Charts ins Internet portieren lassen.

Die Teilnehmenden...

- erhalten eine Einführung in Pandas;
- arbeiten die Grundlagen der Statistik auf;
- analysieren eigene gescrapte Datensätze;
- kombinieren Datensätze mit matplotlib lib, um sie zu visualisieren;
- automatisieren den eigenen Grafik-Stil;
- fügen Pandas in den datenjournalistischen Arbeitsablauf ein und überlegen, was das Programm ihren Projekten bringt.

17. September 2021

Abschluss mit Selenium und Ausblick auf die Projekte

Barnaby Skinner, Thomas Ebermann (nachmittags)

Will man HTML-Daten scrapen, muss aber mit der Website interagieren, eignet sich Selenium hervorragend. Wir lernen, einen Fake-Browser zu steuern und wenden dazu unser HTML- und Python-Wissen an. Danach machen wir Übungen zu allen in den zehn Tagen gelernten Skills von Python-Functions über Requests, BeautifulSoup, Pandas bis zu Selenium. Dazu setzen die Teilnehmenden einen eigenen Server auf, um künftig Scripts automatisiert auszuführen, ohne das eigene Notebook starten zu müssen. Zum Abschluss planen wir die anstehenden Arbeitsschritte für unsere Projekte – und nehmen dazu das Planungstool des ersten Tages zu Hilfe. Danach stellen wir unsere Projekte und die nächsten Arbeitsschritte dem Plenum vor. Zum Schluss folgt ein Ausblick auf die neun Vertiefungsstage, die Abschlussarbeit und den Abschlusstag.

Vertiefungstage: Eigene datenjournalistische Projekte realisieren. Und: Textanalyse, Machine Learning, Geodaten, Netzwerkanalyse

In der Regel finden jeweils zwei Vertiefungstage zusammen statt. Der erste Halbtage dient dem Coaching der Praxisprojekte der Teilnehmenden. Was für Programmierprobleme sind aufgetaucht? Welche Daten erlauben welche Story? Die Recherchethesen werden justiert, die Datenrecherche wird neu ausgerichtet, geplant und dokumentiert. So reflektieren wir laufend das Storytelling. Am Nachmittag und am Folgetag lernen die Teilnehmenden neue Programme und Techniken kennen und eignen sich diese mit Übungen an. Im Durchschnitt sollte zwischen zwei Vertiefungsblöcken jeweils ein Projekt umgesetzt werden. So realisieren die Teilnehmenden im Laufe der vier Monate 3 bis 4 datenjournalistische Arbeiten – zunehmend selbstständig.

04. und 05. Oktober 2021

Sichtung Praxisprojekte, Vertiefung Pandas (und Auffrischung Statistik)

Simon Schmid

Am Vormittag des ersten Tages stellen wir einander den Stand unserer Projekte vor und helfen mit, allfällige Probleme zu lösen. Dabei üben wir die selbstständige Problemlösung mit Stackoverflow, Github und lernen weitere Debugging-Strategien kennen. Am Nachmittag geht es weiter mit der Vertiefung Pandas, der Abfrage von grossen Datenmengen und Visualisierungsmöglichkeiten. Gleichzeitig werden die Grundlagen der Statistik aufgefrischt.

Am zweiten Tag folgt die Vertiefung der Pandas-Inhalte aus dem Basiskurs. Wir reinigen Datensätze, verbinden komplexe Datensätze, behandeln Zeitreihen und kombinieren verschiedene Python-Module mit Pandas. Wir lernen, wie wir unsere Storythesen anhand von Visualisierungen finden, überprüfen und justieren können. Allenfalls werden erneute Datenabfragen/-Strukturierungen/-Analysen/-Visualisierungen nötig. Wir üben diese Abläufe bis zur optimalen, belegbaren Story.

Am Ende des zweiten Tages stellen alle die Projektideen vor, die sie bis zu den nächsten Vertiefungstagen vorantreiben wollen.

28. und 29. Oktober 2021

Arbeiten mit Geodaten

Simon Schmid

Um einen Datensatz zu verstehen, kann es bei einer Recherche sehr nützlich sein, die Daten räumlich darzustellen; auch wenn das Resultat am Ende gar nicht dem Publikum gezeigt wird. Die Teilnehmenden lernen die Python-Library GeoPandas kennen. Sie macht das Arbeiten mit Geodaten mit Python sehr viel einfacher. Alles, was die teils kostenpflichtigen Webdienste anbieten, funktioniert auch mit GeoPandas: Spatial Joins, Overlay, Choropleth Maps - und noch vieles mehr. Dazu noch ein Ausblick auf [OSMNx](#), eine Python-Library, um direkt aus [Open Street Map](#) Karten zu ziehen. Wer damit arbeitet, beginnt die wahre Kraft von Open Source zu verstehen. Und: Ein Blick auf [Google Maps APIs](#).

Am Nachmittag stellen die Teilnehmenden ihre Ideen für die Abschlussarbeit vor und erhalten von den Dozenten ein erstes Feedback.

18. und 19. November 2021

Textanalyse Vertiefung & Ausblick Machine Learning

Thomas Ebermann (1. Tag nachmittags und 2. Tag); Barnaby Skinner (1. Tag, vormittags)

Wir stellen den Stand unserer Projekte vor (Vormittag), helfen einander bei der Problemlösung und allenfalls beim Weiterentwickeln; der Dozent repetiert oder vertieft je nach Bedürfnis am Nachmittag Unterrichtsstoff für die selbstständige Weiterarbeit. Wir üben dabei, wie wir Probleme mit Stackoverflow, Github und weiteren Debugging-Strategien selbst lösen können.

Am Nachmittag geht es mit Regex weiter und der Arbeit mit grossen Textmengen. Gerade in der Verwaltung entstehen viele Textdokumente, die sich von Computern systematisch auswerten lassen. Mit den Python-Modulen glob lernen wir, grosse Textmengen automatisch einzulesen, und mit [Regular Expressions](#) re und dem [Natural Language Toolkit](#), Muster zu erkennen und damit eigene Datenbanken zu bauen. Doch auch hier ist es wichtig, dass wir erkennen, wie wir diese Tools in den datenjournalistischen Arbeitsablauf einbauen und was sie den konkreten Projekten bringen oder welche neue Projekte damit möglich werden. Am Ende des zweiten Tages stellen alle die Projektideen vor, die sie bis zu den nächsten Vertiefungstagen vorantreiben wollen.

06. und 07. Dezember 2021

Netzwerkanalyse

Thomas Ebermann (1. Tag und 2. Tag vormittags), Barnaby Skinner (2. Tag, nachmittags),

Erster Tag: Überblick der relevanten Python basierten Module und Libraries für Netzwerke-Analyse um komplexe Datenstrukturen und Dokumente miteinander in Beziehung zu setzen: networkX et al. Wir formulieren ein mögliches datenjournalistisches Projekt. Am Ende des zweiten Tages stellen alle die Projektideen vor, die sie bis zum nächsten Vertiefungstag vorantreiben wollen. Zudem instruiert der Dozent die Teilnehmenden über die Voraussetzungen ihrer Abschlussarbeit.

06. Januar 2022

Server, Debugging und Projekte

Zweiter Tag, Vormittag: Wir stellen den Stand unserer Projekte vor; allfälliges Coaching und Repetition.

Praxis zur eigenen Problemlösung mit Stackoverflow, Github und weitere Debugging-Strategien;

Nachmittag: Zeit für Repetition; fortgeschrittene Teilnehmende setzen einen eigenen Server auf. Am

Schluss des Tages gibt es einen Informationen zu Abschlussarbeit und Abschlusstag.

Abschlussarbeit

Aufwand ungefähr 5 Tage

Abschlussarbeit, Schlussqualifikation

Barnaby Skinner, Simon Schmid, allenfalls weitere Experten

Die Abschlussarbeit wird im Anschluss an die Vertiefungstage verfasst. Damit zeigen die Teilnehmenden, was sie gelernt haben. Sie realisieren ein datenjournalistisches Projekt unter möglichst realistischen Bedingungen: Sie scrapen, strukturieren, analysieren, visualisieren Datensätze im Hinblick auf eine journalistische Story. Ziel ist es, Daten so aufzubereiten, dass damit ein Artikel/ein Beitrag geschrieben/erarbeitet und veröffentlicht werden kann.

Abgabetermin: Montag, 14. Februar 2022, bis 24.00 Uhr, per Mail an Barnaby Skinner.

Abschlussstag und Feedback

16. März 2022

Barnaby Skinner, Simon Schmid (vormittags), Alexandra Stark

Zum Abschluss zeigen die Teilnehmenden ihre besten Projekte und überlegen, wie sie nach dem Kurs zusammenarbeiten können. Ein Gast aus der Praxis berichtet von seiner Arbeit, die über das Erlernte hinausführt.

Änderungen vorbehalten.

MAZ – DIE SCHWEIZER JOURNALISTENSCHULE

Das MAZ Das MAZ – 1984 gegründet – ist die führende Journalismusschule der deutschen Schweiz. Aus- und Weiterbildung für Medienschaffende sind seine Kernkompetenzen. Die enge Zusammenarbeit mit Redaktionen garantiert Praxisnähe. Kooperationen im In- und Ausland ermöglichen den Blick nach aussen. Hohe Qualität auf allen Ebenen ist der Anspruch. Träger des MAZ sind: der Verband Schweizer Medien, die SRG SSR, die Journalismusverbände, Stadt und Kanton Luzern; weiter wird das MAZ vom Bundesamt für Kommunikation BAKOM unterstützt.