

CORRELCON 2021

A WEEKEND FOR DATA4GOOD

NOVEMBER 12-13

Note: all times are given in UTC+1

November 12

17:00–17:30 (Stage) | #data4good, #npo

Keynote: Data and Impact Measurement

In the social sector we often ask (or get asked) “Does this program work and does it create impact?”. To answer it, we need to answer a different question first: “**What if** my program or intervention wouldn’t exist?”. In this talk we will explore what impact evaluation is and how data can help to answer the question “what if?”. Johannes shares his experience in evidence-based impact evaluation and puts it into the context of the non-profit sector.

Speaker: Johannes Müller (CorrelAid e.V.)

17:30–18:30 (Stage) | #data4good, #correlaid, #npo

Project Showcase

In the project showcase, volunteers from our Data4Good project teams show how they have been supporting non-profit organizations with their data challenges. Topics include: report automation, creating an interactive report based on survey data, analysis of community dynamics, predictive modelling and using image data to improve cargo bike routing. The projects are:

- **Hacklab Developer Census:** Hacklab Ghana Developer Census 2020 is the first and most comprehensive survey of people who code in Ghana. We developed an interactive HTML Rmarkdown report that showcases the results covering everything from developers’ favorite technologies and languages to their job preferences.
- **CargoRocket: Classification of Road Quality Using Crowd-Sourced Images:** A cargo bike has its own unique requirements in terms of road conditions, which must be considered when planning a route. Therefore, CorrelAid volunteers developed a Python program that uses machine learning to classify road surface and pavement quality based on crowd-sourced street images.



CORRELAID
GOOD CAUSES. BETTER EFFECTS.

CORRELCON 2021

A WEEKEND FOR DATA4GOOD

NOVEMBER 12-13

Note: all times are given in UTC+1



- **Citizen Science Dynamics:** We've analyzed data of the Citizen Science Platform *iNaturalist*, where users can upload observations of plants and animals, and other users can help identify them. We found out a lot of cool things on how the user base changed over time, what types of users there are, and much more - we're writing a paper about it! I'll also share some lessons learned. Spoiler: Exploratory data analysis can eat a lot of time.



- **Citizens for Europe Report Automation:** Together with Citizens for Europe, we developed a workflow that allows for flexible generation of reports on discrimination and diversity within organizations. We will present how to go one step further than parameterized reports, building R-Markdown templates in a modular, more adaptable way.

- **Science on Stage:** As many other NGOs, Science on Stage collects anonymized data from different sources (online-seminars, website, downloads) to measure the effectiveness of their work. We helped Science on Stage, an NGO supporting STEM-teachers from different backgrounds, to make sense of their data sources by automizing the data collection and cleaning process. In a second step, we visualized the results to create valuable reports in order to facilitate inhouse decision-making as well as keeping the Science on Stage shareholders up-to-date.



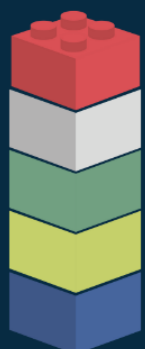
- **The impact of Voting Advice Applications on local politics:** The CorrelAid team supported "Team Tomorrow" with the analysis of data from the "KOMUNAT", an entirely candidate-based voting advice application (most famously known in Germany: the "Wahl-o-mat") for the local town elections in Stuttgart in May 2019, developed to address the traditionally low voter turnout amongst the young first-time-voters. Overall, two major topics had to be investigated: the user statistics for product improvement and the political analysis to assess both the political preferences of the app users and the potential impact of the VAA on local politics.

- **Mare Liberum: Data visualizations to bring attention to illegal push-backs:** Push-backs are a set of state measures by which refugees and migrants are forced back over a border without considering individual circumstances and allowing them to apply for asylum. Mare Liberum gathers data on the illegal practice at the Mediterranean sea and publishes documentation on observed occurrences. Supporting their mission through



CORRELAID

GOOD CAUSES. BETTER EFFECTS.



CORRELCON 2021

A WEEKEND FOR DATA4GOOD

NOVEMBER 12-13

Note: all times are given in UTC+1



data, CorrelAid has helped to conduct data analysis and provide visualisations for their annual report and consulted them on data collection and storage.



- **Addressing deforestation in Kosovo:** Correlaid Netherlands, the Institute of GIS and Geospatial Technology of Kosovo (IGGTK), and Sustainability Leadership Kosova (SLK) worked on a joint project to address the deforestation issue in Kosovo. The main goal of the web platform (<https://www.slkosova.org/mosemerrmalin-welcome>)? Displaying maps of the deforestation events that occurred in the last 20 years to educate and raise awareness of this massive issue affecting Kosovo. #MoseMerrMalin
- **Gmoji:** G-Moji is developing a mental health support app for adolescents and wanted CorrelAid's help in assessing the feasibility of mood prediction from its pilot phone sensor and survey data.
- **Kulturen im Kiez:** Description will be added soon.



Speakers: François Delavy (CorrelAid volunteer and data analyst), Sofia Thai (CorrelAid volunteer and data analyst), Alina Dallmann (CorrelAid volunteer), Frank Schlosser (CorrelAid volunteer), Christine Hedde-von Westernhagen (CorrelAid volunteer), Simon Schölzel (CorrelAid volunteer), Tim Fangmeyer (CorrelAid volunteer), Andreas Neumann (CorrelAid), Aylin Kallmeyer, Lukas Gröninger (CorrelAid volunteer), Federico Franciamore (CorrelAidX NL volunteer), Wen Qiang Toh (aka Wally, Data science volunteer at CorrelAidX NL), Daniela Mondorf (CorrelAid volunteer)



18:30–18:40 (Networking) | #offline

Break

Catch a beverage and snack of your choice before heading to one of our topic sessions!

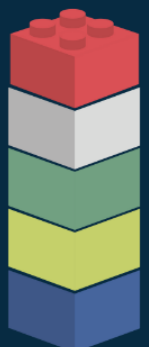
Topic Sessions

18:40–19:40 (Sessions) | #npo, #security

Introduction to Information Security for Non-Profits



CORRELAID
GOOD CAUSES. BETTER EFFECTS.



CORRELCON 2021

A WEEKEND FOR DATA4GOOD

NOVEMBER 12-13

Note: all times are given in UTC+1



Ransomware, leaks, data loss, and what to do about it (in COVID times): In this talk, Ilja will introduce us to the pillars of Information Security, provide an overview of organisational and individual threats, and outline mitigation strategies - all with a focus on NGOs. The talk will include an actionable plan for your team to get started.

Speaker: Ilja Sperling (Freelance Dev & former NGO Tech Lead)

18:40–19:40 (Sessions) | #npo, #education

Data Science Education



Data science is a broad field. Together, we will explore which skill paths a data enthusiast for good may choose and what is needed to excel both professionally and ethically. We will then look at how we can translate those skills to foster data literacy in the civic sector.

Speakers: Nina Hauser (CorrelAid e.V.), Vik Paruchuri (Dataquest)

18:40–19:40 (Sessions) | #npo, #data4good

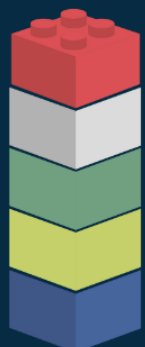
Daten im sozialen Sektor: Chancen und Hürden



Während sich in der Wirtschaft in den letzten Jahren ein Verständnis für digitale Technologien entwickelt hat, werden Gesellschaft und zivilgesellschaftliche Organisationen hier bislang größtenteils abgehängt. Deutlich wurde das im Digital-Report 2020: 72 Prozent der 3,400 befragten Organisationen haben eine Organisationsstrategie, aber nur 38 Prozent erheben und/oder nutzen (externe) Daten zur Strategieentwicklung. Eine vollständig durchdachte Wirkungslogik haben nur 30 Prozent der Organisationen und nur insgesamt 28 Prozent messen tatsächlich ihre soziale Wirkung. Es fehlen finanzielle und personelle Ressourcen, um die neuen Möglichkeiten durch Daten und Datenanalyse erfolgreich nutzen zu können. Mitarbeiter:innen schätzen ihre Fähigkeiten und Kenntnisse im Umgang mit Daten als zu gering ein. Dabei können Daten und ihre Analyse ein wichtiger Hebel für gesellschaftliche Wirkung sein. Moderator Johannes Müller diskutiert mit Katarina Peranić Deutschen Stiftung für Engagement und Ehrenamt, Joshua Kwesi Aikins (Citizens for Europe), Georg Förster (AWO Deutschland) und Frie Preu (CorrelAid e.V.) über die Chancen und Hürden der Datennutzung im sozialen Sektor: Was können Daten, Data Science und "KI" für zivile Organisationen leisten, was nicht? Was hindert große und kleine Organisatio-



CORRELAID
GOOD CAUSES. BETTER EFFECTS.



CORRELCON 2021

A WEEKEND FOR DATA4GOOD

NOVEMBER 12-13

Note: all times are given in UTC+1



nen daran, Daten zur Verbesserung ihrer Arbeit zu nutzen? Und wie können Akteur:innen der Zivilgesellschaft zu mündigen Datennutzer- und analyst:innen empowered werden?

Speakers: Katarina Peranić (Deutsche Stiftung für Engagement und Ehrenamt), Georg Förster (AWO Bundesverband), Kwesi Aikins (Citizens for Europe), Frie Preu (CorrelAid e.V.)

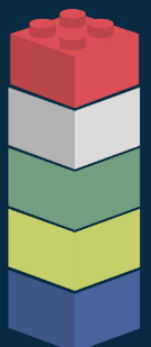
19:40–22:00 (Gather.Town) | #community, #networking

Game Night - organized by CorrelAidX Berlin

Join us on Gather.Town for a fun Friday night to connect and share. We will provide the games but it is BYOB (Bring your own beverage).



CORRELAID
GOOD CAUSES. BETTER EFFECTS.



CORRELCON 2021

A WEEKEND FOR DATA4GOOD

NOVEMBER 12-13

Note: all times are given in UTC+1

November 13

10:00–10:30 (Stage) | #correlaid

CorrelAid in 2021

Frie Preu, Head of Operations, will walk you through the CorrelAid year 2021 with all its milestones, achievements and challenges.

Speaker: Frie Preu (CorrelAid e.V.)

Data Science, Climate Change and the Environment

10:30–10:45 (Stage) | #data4good, #climatechange

Tackling Climate Change with AI

Artificial Intelligence (AI) and Machine Learning (ML) provide powerful tools to tackle climate change in various applications: They can support climate change mitigation, for instance, by helping reduce greenhouse gas emissions within various applications. They can help to adapt to a changing climate and even advance climate science itself. However, AI and ML are not silver bullets and can always only be one part of the solution. This talk provides an overview of the strengths and weaknesses of ML, some example applications and recurring themes.

Speakers: Marcus Voss (Climate Change AI, Intelligence Architect and AI Expert at Birds on Mars)

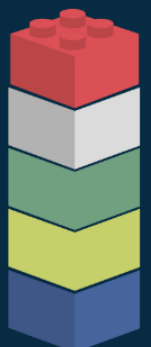
10:45–11:00 (Stage) | #data4good, #sustainability, #usecase

Put your money where your heart is - climate investing for your personal finance

Description will be added soon.



CORRELAID
GOOD CAUSES. BETTER EFFECTS.



CORRELCON 2021

A WEEKEND FOR DATA4GOOD

NOVEMBER 12-13

Note: all times are given in UTC+1



Speaker: Constanze Bayer (2° Investing Initiative (2DII))

11:00–11:15 (Stage) | #data4good, #usecase

Using AI to help restore the natural world

Land use and land use change play a critical role in our climate taking up about a quarter of annual anthropogenic emissions of greenhouse gases (GHGs) during 2007-2016. In addition to being a key driver of global warming, careless land use is also destroying valuable ecosystem services and is threatening the livelihood for local populations and a multitude of species. Major conservation and restoration efforts are underway to mitigate and safeguard against these losses, and to highlight the urgency of the issue, 2021-2030 has been declared the "UN Decade on Ecosystem Restoration". However, we cannot preserve what we cannot measure. Machine learning (ML) plays a significant role in responding to this critical call for action and can accelerate the conservation and sustainable use of our natural world.

Speaker: David Dao (GainForest)

11:15–11:30 (Stage) | #datascience, #sustainability

Sustainability of AI

This talk is about the impacts of AI-systems from a sustainability perspective. Based on current work within the project SustAI we present first ideas for a comprehensive sustainability assessment of AI-Systemes. With this effort we want to raise discussions amongst industry, developer communities and society about how to reduce social, economic and environmental risks throughout the whole lifecycle of AI-systems and call for action to make AI-systems more sustainable.

Speaker: Friederike Rohde (Wissenschaftliche Mitarbeiterin am Institut für ökologische Wirtschaftsforschung (IÖW))

11:30–12:00 (Stage) | #NA

Q&A

Ask our experts your questions!



CORRELAID
GOOD CAUSES. BETTER EFFECTS.



CORRELCON 2021

A WEEKEND FOR DATA4GOOD

NOVEMBER 12-13

Note: all times are given in UTC+1



Speakers: Marcus Voss, Constanze Bayer, Friederike Rohde, tba

12:00–13:30 (Stage) | #break, #offline

Lunch Break



Topic Sessions

13:30–14:30 (Sessions) | #careers

Pathways to Data Science



Physics, computer science or political science - all these (and many more) backgrounds can be found among the volunteers of our network. As diverse as our network, however, is the field of data science that connects us all. In this talk, three of our volunteers will tell us about their personal journey into data science. Among other things, this should help prospective data scientists to find their way in an exciting and often bizarre job market.

Speakers: Andrew Sutjahjo, Allison Koh, Sharon Gieske

13:30–14:30 (Sessions) | #careers

Agile2 - Juggling Innovation and Bureaucracy in IT Project Management

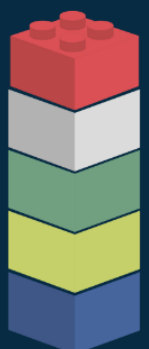


Imagine you created a compelling business case, recruited brilliant team members for your cross-functional project team, and even received generous government funding to build your shiny new software product. Being the modern day project manager you are, you plan to adopt agile methodologies by appointing a product owner for your self-organizing team, creating an emergent product backlog, and writing epic user stories. Before you can turn your ideas into reality, though, you have to convince your most influential non-IT stakeholders – and they are not amused. Who is responsible for controlling time, scope, and cost? Where is this important document called „risk register“ that they were told about? And why the unicorn are team members allowed to create their own tasks? That is not how project management works! Now it is up to you, the project manager, to come up with a solution to this dilemma.



CORRELAID

GOOD CAUSES. BETTER EFFECTS.



CORRELCON 2021

A WEEKEND FOR DATA4GOOD

NOVEMBER 12-13

Note: all times are given in UTC+1



Speaker: Lisa Hehnke (Innovation Project Lead at Tafel Deutschland e.V. | Freelance Consultant & Data Scientist)

13:30–14:15 (Sessions) | #usecase, #data4good

CeMAS

Conspiracy theories have been particularly loudly perceptible since the pandemic at the latest. While news coverage and also science have tended to perceive conspiracy theories with amusement, awareness of them is currently changing and many are recognizing the major problems associated with the spread of conspiracy ideologies: Compliance with pandemic response measures, vaccination preparedness, but also misanthropy and propensity for violence. CeMAS, as an organization, aims to shed more light on conspiracy ideologies, anti-Semitism, and disinformation, and to develop strategies for action. How do we proceed, what are the biggest problems and which technical means do we use to approach this complex of topics.

Speakers: Josef Holnburger is a Political Data Scientist and co-director at CeMAS - Center for Monitoring, Analysis and Strategy.

Beginner Sessions

14:30–16:00 (Sessions) | #rstats, #workshop

Beginner session R

In this session you will learn first steps with R and RStudio, like reading data from a csv file, data wrangling and creating simple crosstables and plots. We will mostly use base R and packages from the tidyverse, which are a great starting point for beginners.

Speaker: Ann-Kristin Vester

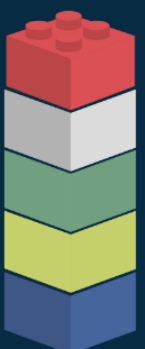
14:30–16:00 (Sessions) | #python, #workshop

Beginner session Python

We take a look at some simple use cases and how we can tackle them using some prominent libraries. We start with some data cleaning using pandas and continue with forecasting using the prophet library and sentiment analysis.



CORRELAID
GOOD CAUSES. BETTER EFFECTS.



CORRELCON 2021

A WEEKEND FOR DATA4GOOD

NOVEMBER 12-13

Note: all times are given in UTC+1



Speaker: Tilman Kerl

Advanced Sessions: Working with Geospatial Data

14:30–15:00 (Sessions) | #geospatial, #dataviz, #mobility

Taking the shortest path - analysis of bicycle infrastructure data in cities

This talk will be all about biking in cities: you will see a brief introduction in tools for scraping and visualizing bicycle infrastructure data of cities. Based on a case study, we discuss a simple method for analyzing bicycle networks and explore how this could provide insights for mobility planners.

Speaker: Fabian Schaipp (TU München)

15:00–15:30 (Sessions) | #geospatial, #mobility, #python

CargoRocket: How to use OpenStreetMap and Graphhopper to create a custom routing for Cargobikes

Navigation apps for cars and bikes fail to properly serve the needs of cargobikes that are allowed to use bike tracks but, e.g., get stuck at too narrow barriers. With the open source project CargoRocket we aim to provide a routing specifically for cargobikes. This talk gives technical details how OpenStreetMap is used to create a "CargoBikeIndex" for every street and how this index is incorporated into Graphhopper, an open source routing engine.

Speakers: Alexandra Kapp (cargorocket, CorrelAid)

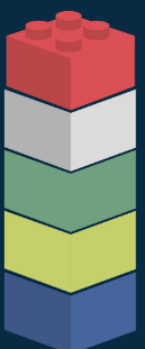
15:30–16:00 (Sessions) | #geospatial

Introduction to analysing spatial data in R

In our (practically oriented) step-by-step introductory presentation, we will walk you through the process of creating maps in R ranging from reading in and preparing data to handling different shapes of spatial data (e.g. points, lines and



CORRELAID
GOOD CAUSES. BETTER EFFECTS.



CORRELCON 2021

A WEEKEND FOR DATA4GOOD

NOVEMBER 12-13

Note: all times are given in UTC+1



polygons). For this we will be using a data set of bike sharing stations and trips in Hamburg.

Speakers: Jasmin Classen, Aylin Kallmeyer, Andreas Neumann



Advanced Sessions: Modelling Corner

14:30–15:00 (Sessions) | #modelling

One Million Posts - Training a neural network to detect hate speech in German text

Natural language processing project based on the one-million-posts dataset using the neural network architecture BERT. More than 3.000 user comments are written each day on www.derstandard.at. Moderators need to review these comments regarding several aspects like inappropriate language, discriminating content, off-topic comments, questions that need to be answered, and more. We have trained a dense layer on top of BERT to classify arbitrary German text regarding the question if it needs moderation.

Speaker: Dominik Hohmann (Volunteer at CorrelAid)

15:00–15:30 (Sessions) | #modelling, #python, #usecase

Network models for analysis of NGOs

In this talk we will discuss different network models and how they can help us to model some processes in nature and especially from data. We will look into some examples as well as talk about packages for their implementation. We will also discuss particular cases of where network analysis could be useful for some NGOs.

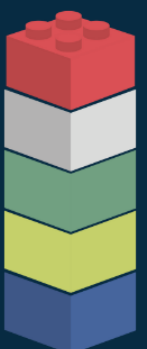
Speakers: Liu Bauer (researcher at CRI, co-leader of CorrelAid Paris, co-founder of City Interaction Lab, Montreuil)

Data Visualization

14:30–15:00 (Sessions) | #dataviz, #rstats



CORRELAID
GOOD CAUSES. BETTER EFFECTS.



CORRELCON 2021

A WEEKEND FOR DATA4GOOD

NOVEMBER 12-13

Note: all times are given in UTC+1



Creating your own themes and colour scales for ggplot2

When producing graphs for others to consume, you may want to have your own distinct visual identity across all graphics, using your own colour palette, your fonts, your logo instead of the built-in `theme_minimal()`, `scale_colour_viridis()` etc. Using the CorrelAid theme and colour scales in the {correltools} package as examples, this talk will show how you can easily create reusable functions to apply your own “brand” to your ggplot2 graphics in a time-saving manner.

Speaker: Long Nguyen (PhD student at SOEP RegioHub/Bielefeld University)

15:00–15:30 (Sessions) | #dataviz, #python



A picture is worth a thousand words - Visual Data Exploration

Data is manifold, which makes it hard to discover its secrets and hidden relationships. Data Scientists can rack their brain on descriptive statistics or choose to explore the data using interactive visualisation tools. While both ways have benefits and dangerous pitfalls, the latter is definitely the more fun. This talk will help you discover your data visually by showing how you can quickly transform simple plots into powerful, interactive dashboards using Panel. Furthermore, it will provide you with a quick overview over Python’s most loved interactive plotting libraries and show how you can combine those in a single, interoperable dashboard.

Speaker: Nico Kreiling (scienceers)

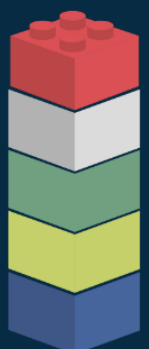
15:30–16:00 (Sessions) | #dataviz, #accessibility

Revealing Room for Improvement in Accessibility within a Social Media Data Visualization Community

Screening of the talk “Revealing Room for Improvement in Accessibility within a Social Media Data Visualization Community” held by Silvia Canelón and Liz Hare at csvconf. Followed by a Q&A and discussion of tips outlined in this blog post. For a description of the talk, see here. License information: “Revealing



CORRELAID
GOOD CAUSES. BETTER EFFECTS.



CORRELCON 2021

A WEEKEND FOR DATA4GOOD

NOVEMBER 12-13

Note: all times are given in UTC+1



Room for Improvement in Accessibility within a Social Media Data Visualization Community” by Silvia Canelón and Liz Hare, available on YouTube, used under CC-BY-SA 4.0

Speakers: Silvia Canelón, Liz Hare



Tooling and Software Development

14:30–15:00 (Sessions) | #tooling

Git internals (or understanding git): The three basic git objects and their relation to git’s way of version control

Git as a version control program can be confusing. This talk aims at helping you to understand git better by knowing some internals of the program: Who does git version control is build around the three basic git objects? Afterwards you hopefully grasp what git trees are made of, that each node in these trees represent a snapshot of your working directory and how branches work. There will be a brief demo of the core concepts. Audience: You use git, but sometimes get confused why something did not work. Hopefully knowing some basics about the internals will help you with these struggles.

Speaker: Henry Weibel (Researcher at Novo Nordisk Foundation Center for Protein Research)

15:00–15:30 (Sessions) | #tooling, #workflow

Introduction to Pipelines with Make

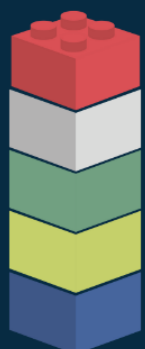
Managing large (research) projects with multiple data processing and computation steps can be hard. However, there is a wide range of tools that can help in such situations. In this talk, Malte will give an introduction to pipeline toolkits, including a brief tutorial on the classic GNU Make utility.

Speaker: Malte Kyhos

15:30–16:00 (Sessions) | #softwaredevelopment, #rstats



CORRELAID
GOOD CAUSES. BETTER EFFECTS.



CORRELCON 2021

A WEEKEND FOR DATA4GOOD

NOVEMBER 12-13

Note: all times are given in UTC+1



Big things come in small packages – Tips and tricks for developing your own R package

R is a great resource for data management, statistics, analysis, and visualization – and it becomes better every day. This is to a large part because of the active community that continuously creates and builds extensions for the R world. During the talk, I will provide you with tips and tricks for contributing to this world by writing packages. If you want to get a quick introduction before the talk, here is a hands-on tutorial on how to write your first R package.

Speaker: Cosima Meyer (Data Scientist)

16:00–16:30 (Networking) | #NA



Break

Enjoy some coffee or tea or go for a quick walk around the block before our CorrelAidX afternoon commences!

16:30–17:15 (Sessions) | #community, #correlaid

Team presentations of NLP-challenge results

Since late September several teams from our different local chapters have taken on this year's CorrelAid-X Challenge. Here we would like to give them the stage to briefly present their results and tell us about the pitfalls and obstacles they faced along the way and what is more, overcame them as a team. We would like to thank Krystel from BuildUp for providing us with the data for this year's challenge.

Speaker: tba

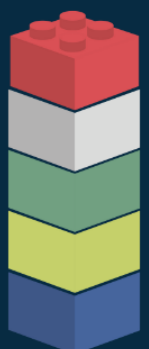
17:15–18:00 (Stage) | #community, #correlaid

CorrelAidX Pecha Kucha

Meet our local chapters through this engaging presentation format and see where you could find a chapter that might work for you.



CORRELAID
GOOD CAUSES. BETTER EFFECTS.



CORRELCON 2021

A WEEKEND FOR DATA4GOOD

NOVEMBER 12-13

Note: all times are given in UTC+1



- Local Chapter Konstanz
- Local Chapter Netherlands
- Local Chapter Berlin
- Local Chapter Switzerland
- Local Chapter Cologne
- Local Chapter Rhein-Main
- Local Chapter Munich
- Local Chapter Bremen



Speakers: Zoé Wolter, Marco Lax, Valentin Geier, Alexis Gillett, Liam Bailey, Freya Moßig (CorrelAid Switzerland and Vice Chair at DataCross), tba, tba, Daniel Manny, Pia Baronetzky, tba

18:00–18:15 (Stage) | #community, #correlaid

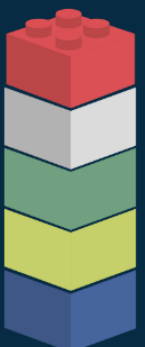
Thank you and goodbye

A warm thank you from us to participants, speakers and sponsors of CorrelCon!

Speaker: CorrelCon organizers



CORRELAID
GOOD CAUSES. BETTER EFFECTS.



CORRELCON 2021

A WEEKEND FOR DATA4GOOD

NOVEMBER 12-13

Note: all times are given in UTC+1

Our Sponsors

Logo sponsors



CORRELAID
GOOD CAUSES. BETTER EFFECTS.

