Big Data For Small Organizations

We are a Netherlands-based start-up company that brings the benefits of big data to small organizations. In order to help small organizations remain competitive in the era of big data, we provide access to trustworthy AI, validation of third-party data, and strategies to combat the potentially negative effects of biased corporate sales and marketing algorithms. Our diverse team is particularly experienced in music, the creative industries, and digital humanities.

Our flagship demo projects are the Listen Local ethical music recommendation system based on our Demo Music Observatory data integration and knowledge sharing platform. We have validated our product/market fit in the prestigious Yes!Delft AI+Blockchain Lab

☛ See our services: data curation, open data access, survey harmonization, reproducible research and validated trustworthy AI applications.

⬇ Download our introduction.

🔊 Follow news about us or the more comprehensive Data & Lyrics blog.

📞 Contact us .

Reprex

Big data for small organizations.
Yes!Delft

Big data creates inequalities. Only the largest corporations, government bureaucracies, and scientific organizations have the human financial resources to create large, consistent datasets. In contrast, small organizations with limited resources are at a competitive disadvantage—unless they are able to access and use governments' open data sources and replicable scientific datasets, using the latest innovations in reproducible research to keep data costs low. Because AI requires large datasets to be effective, small organizations face challenges in harnessing the benefits of algorithms—or, worse, biased algorithms created for different purposes will work against their strategic goals.

We help our clients to access datasets of sufficient size, and to build trustworthy scientific, policy, business, or AI applications with the data. Reprex grew out of CEEMID, a project that connected data and pooled surveys across twelve countries and over fifty music industry stakeholders to create thousands of relevant policy, statistical, and scientific indicators for the music and creative industries. (See the Central European Music Industry Report 2020 at ceereport2020.ceemid.edu.)

Reprex offers services to help small organizations access data and use data in a trustworthy way. We encourage our clients to become partners in our permanent data collection and sharing platforms: the data observatories.

- Data Curation: Find the best value data solution for your research or evaluation project.
- Open Data: Access thousands of statistical, business, and policy indicators in the cultural and creative industries and sustainability.

- [Survey harmonization](): Reuse already existing surveys and harmonized question banks; create longitudinal and cross-sectional datasets.
- [Reproducible Research](): Automate data collection, processing, validation, correction, and documentation to support error-prone human work.
- [Trustworthy AI](): Use validated, trustworthy AI applications or identify the biases of algorithms working against your goals.

Our flagship demo projects are the [Listen Local]()ethical music recommendation system based on our [Demo Music Observatory]() data integration and knowledge sharing platform. You can find our globally [open-source statistical software]() and [some publications]()made with our data service on our website. We try to evangelize trustworthy AI on our [Data & Lyrics]() blog.

Cooperation

Clients often contact us seeking affordable data. While data access is an important element of our services, we often find that our partners lack not just data, but information about data availability. Our approach is to help curate the data necessary for a business, policy, or scientific project, replacing an imagined, expensive, or unavailable data source with an available and affordable solution. Consultation about data possibilities and assistance locating the right data sources for a project is, in the end, more valuable than ad hoc access to cheap data.

We prefer to make contracts that are longterm or have a longterm outlook. The secret to affordable data is continuous data collection. Data that is requested ad hoc is usually expensive or of low quality. We can offer significantly better prices and services if our clients partner in one of our permanent data collection programs, the data observatories.

Reprex assists with AI challenges, such as training or validating algorithms, econometric and social modeling, or various business analytics solutions. We believe that small organizations have a competitive disadvantage in this field. Our clients lack the human and financial resources to invest in data pools of the necessary size and in the data processing and normalization required to train algorithms and create good models. Our data observatories offer data not merely as a product, but as a service, where we make the necessary investment in order for your organization to have continuous access to large, well-processed datasets for extended use.

- Not convinced? Challenge us to provide access to data you cannot find in one of our data observatories.
- Contract us for a single dataset or small research program (but we will encourage you to consider an ongoing relationship!).
- If you lack the resources to invest in a permanent data science capacity in your organization, accept our offer for data-as-a-service by joining our data observatories—or for reproducible-research-as-service.

How to join our data observatories?

Partnering with our data observatories is a flexible option. What we ask from partners is a letter of intent in participating in the joint activities listed below, as well as an immediate budgetary or data asset contribution at the time of joining.

Participation in data curation: Help observatory partners to understand data needs and potential data sources. We add value through automated processing and controls, but we need domain specific academic, business, and/or policy expertise to assess the value of the data.

Shared exploitation: Use the observatory data tables and data assets and publicly reference them. Some of the observatory assets are open to the public, while others, due to the nature of their scientific or regulatory origins, will become public after peer review or external audit. We prioritize the embargoes and scientific, business, or other interests of our observatory partners.

Funding: Commit to participate in funding. We can accommodate project-based or permanent funding, and have various budgetary options for academic institutions, trade associations and collective management societies, businesses, and policy consultancies. We give priority to non-project-based funding because, while continuous data collection allows significant cost savings, it also requires an ongoing funding base.

Validation: Through the use of data, critically test and assess the joint data assets and suggest improvements. Our data assets are open to many levels of validation: we give full access to the data and the code that generates the indicators, validating both the data and the code. We believe in ongoing peer review and external audit, and we encourage our users to utilize our data in peer reviewed or externally audited applications. We committed to using the feedback from external reviews promptly for continuous quality improvement.

Seek recognition and external funding: Help us participate in any available EU, UNESCO, or OECD initiatives or funding schemes that could interface with our observatories. This enables a higher level of validation, recognition, and access to base funding for a permanent data program.

Who we are?

Reprex is a for-profit company registered in the Netherlands in 2020. Our diverse team is comprised of the freelancer network of the 2014-19 CEEMID project, as well as new team members with a background in music and data. The co-founding shareholders are Daniel Antal, CFA and Réka Szentirmay. You can find more information about our team on our website. In 2020, we successfully completed a grant-funded project for the Slovak Arts Council, and validated our product/market fit in the prestigious Yes!Deflt AI+Blockchain Validation Lab. You can contact us on various channels via our website.