

Mozilla Data Collective

Create. Curate. Control.



Agenda



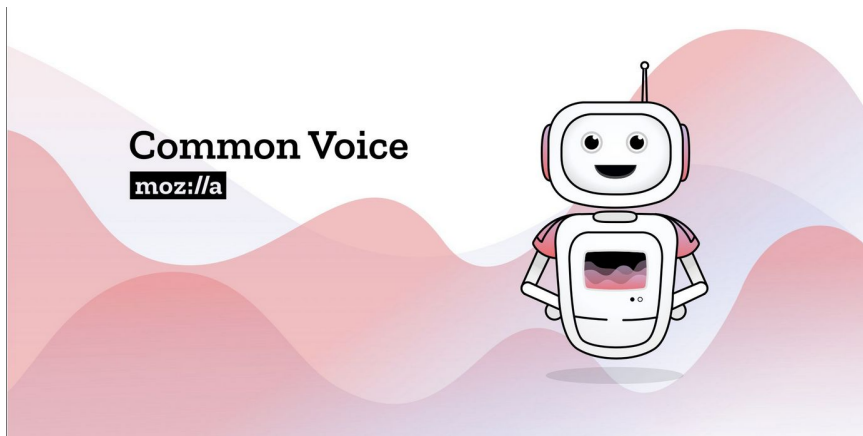
- 01 **Background:**
The MCV experience and the AI data landscape
- 02 **A better way:**
Elevating human agency and fair value exchange
- 03 **Exhortation!**
YES! Mozilla Data Collective





Background

Mozilla Common Voice



- The Mozilla Common Voice experience
- Massively multilingual speech platform – over 300 language communities represented
- Volunteer and community driven
- Public domain licence

What communities want



- **Human agency** – Control over their data
- Collaborative relationships with developers
- Applications that serve them
- **Fair value exchange**

Current data landscape



- Institutional repositories
- Non-institutional repositories
- Data platforms
- Data brokers



A better way

What we are promoting



- Centering communities — both language communities and developer communities
- Integrated creation and distribution of datasets
- Full control over distribution



Mozilla Data Collective

Create, curate, control



- **Create**
 - — Tooling to enable dataset creation, collaboration in data projects
- **Curate**
 - — Collaborate in uplifting datasets to maximise utility.
- **Control**
 - — Choose the licence, set the terms, specify the audience, define what fair value exchange means to you

How it works



- **Uploaders**

- Provenance: Every request to upload and every dataset upload is reviewed by a human

- **Downloaders**

- Browse, search and download datasets, request permission if applicable

Get involved!



- Make an account
- Search for and download datasets
- Request a dataset
- Upload a dataset
- Partner with us

Gràcies



Gràcies als meus col·legues per als seus suggeriments i ajuda amb la preparació d'aquesta xerrada!