

Democratizing AI

Open Data Licences

Providing Access to open training data

Through...



Community



Public entities



Start-Ups, NGOs

Commons-based Peer Production

Individuals working cooperatively for the creation or maintenance of a shared resource. Resources are open, (re-)usable & freely accessible to anyone

Future users must give improvements back to the community.

Can it work?

Yes, but ...

it needs a stable commons management design!



Elinor Ostrom (Nobel laureate)

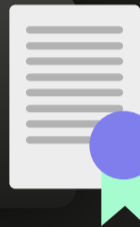
Otherwise...

Tragedy of the Commons

Overloaded and badly managed commons will be unusable

Licenses

give official permission for the person to do something or for the activity to take place



License Types

Proprietary



Work may not be modified or redistributed unless with author permission

Public Domain



Waive All the rights to work under copyright

Recommended for AI Data

Permissive Open Licenses



Open + Minimal requirements about how software can be modified or redistributed

Maximum use + co-licensing for commercial business models

Openness

Unfettered access and easy ways to contribute

Not giving back to sustain the data pool

Responsible AI Licences



Open - but restrictions to prevent irresponsible and harmful applications

No privatisation & foster community

Responsibility

Growing the pool of training data with safety in mind

Some restrictions but giving back a lot to the data pool

CopyLeft Licenses



Open - but you can modify and redistribute work only under same license type

No privatisation & foster community

Perpetuity

Sustaining and growing the pool of training data

Inflexibility and Restrictions

TENSION

TENSION

How to tackle the tension?

Mechanism of openness and access



Mechanism for perpetuity



current

Data Sharing Approches

Data Trusts

Data Cooperatives

Data Commons

Data pools & data spaces

Mandatory Data Sharing

Open Data

Data Markets & Data Exchanges

