

Nadezhda Purtova

Important questions to answer before talking of "data commons"

- What do we mean by "data commons"?
 - Many scholars speak of the new models of data governance in terms of "data commons". But what do we mean by "data commons"? There are two options:
 - * **Data commons = data as a common-pool resource (CPR)** according to the traditional economic resources classification along the axes of excludability and rivalry. Elinor Ostrom proposed this new type of economic resources and founded a branch of neo-institutional economics that studies how CPRs can be managed to avoid a "tragedy of the commons". Classification of data as a CPR implies that data is not (fully) **excludable**, i.e. access to data cannot be prevented, and is **rival**, or subtractable, i.e. enjoyment or other "consumption" of data leads to deterioration of its quality/usefulness. This deterioration is referred to as common social dilemmas.
 - * **Data commons = data "held in common"**, i.e. commonly "owned" and managed. Unlike "data as a CPR", this framing does not imply that data has certain inherent characteristics. Instead, it is more of a normative claim: data should be held in common, as opposed e.g. to by a few tech giants. Ostrom noted a conceptual confusion between the 2 meanings and urged to guard the separation.
 - Why are we talking about data commons? How is this framing useful for us?
 - * The "data as a CPR" framing is useful because studies led and inspired by Ostrom have resulted in several **design principles for governing common-pool resources** (e.g. [here](#)) that could be used for data governance, provided data is a CPR.
 - * The "data should be held in common" frame is useful to make a **normative statement, make a claim on data, talk about data politics, etc.** But it does not provide ready answers as to how data should be governed, other than "in common".
 - * Research that talks about data commons as "data should be held in common" often reach to the commons design principles which belong to the "data as a CPR" frame. This is not correct.

- If we go with the "data should be held in common" frame, what are the normative underpinnings to argue that all data should be common?
 - * It is easier to argue this when we talk about personal data, since individuals have a claim on "their" personal data, and personal data affects groups of people beyond the immediate "data subjects".
 - * But what is the normative ground to claim all data in commons?
- Suppose we want to go ahead with the "data as a CPR" frame and use the design principles. Then other questions arise.
 - * What is the resource that is a CPR? Data is not rival (one's "enjoyment" of data does not lead to its deterioration or depletion).
- Open questions:
 - * (To what extent) are the CPR design principles transferable to the "data held in common" context? Perhaps, they are transferable to a limited extent, since the CPR design principles also deal with problems of cooperation common to both frames. But further research is needed.

