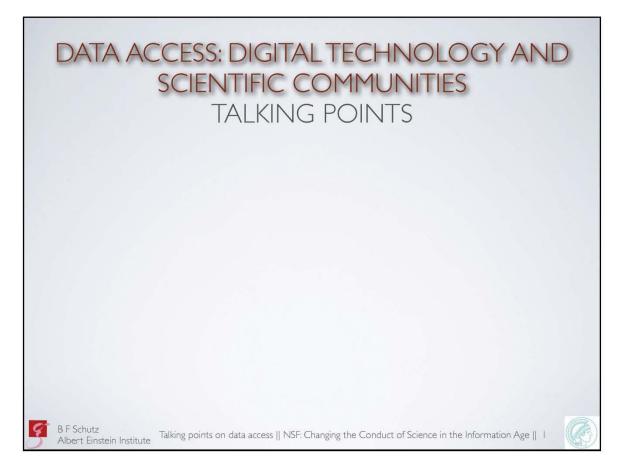
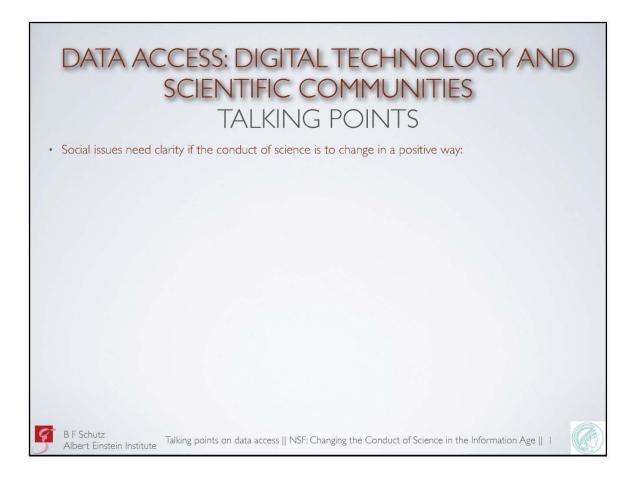
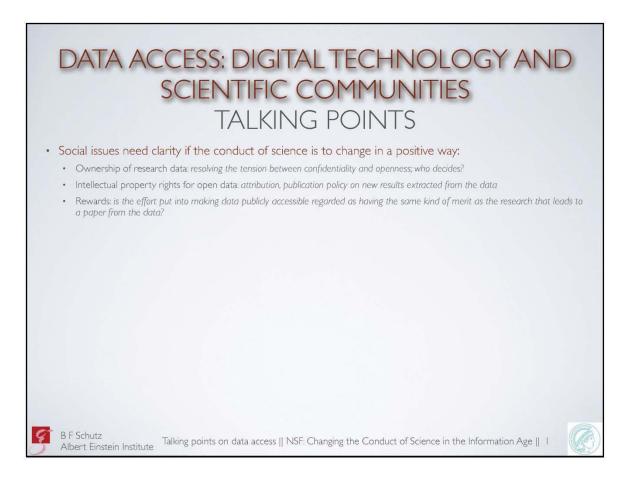
Schutz, Bernard, "Data Access: Digital Technology and Scientific Communities"

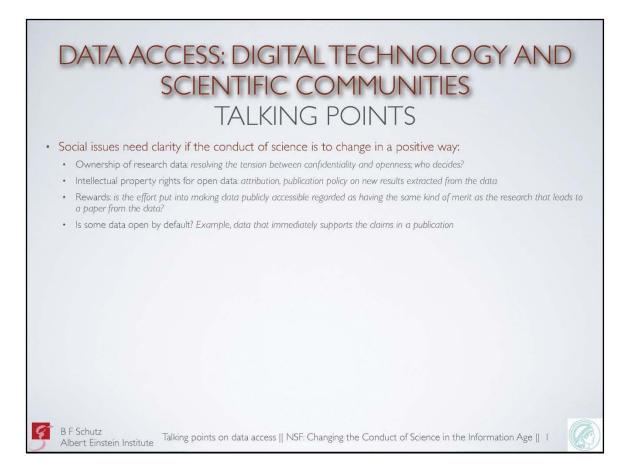














Social issues need clarity if the conduct of science is to change in a positive way:

- Ownership of research data: resolving the tension between confidentiality and openness; who decides?
- Intellectual property rights for open data: attribution, publication policy on new results extracted from the data
- Rewards: is the effort put into making data publicly accessible regarded as having the same kind of merit as the research that leads to
 a paper from the data?
- Is some data open by default? Example, data that immediately supports the claims in a publication
- · Policy issues for data repositories: access to data, access to metadata, preservation policy, data integrity, replication, etc
- Standards: Adoption/creation of international standards (metadata, formats, glossaries, indices) to facilitate search, retrieval, exploitation



B F Schutz

Albert Einstein Institute



Talking points on data access || NSF: Changing the Conduct of Science in the Information Age || 1

Social issues need clarity if the conduct of science is to change in a positive way:

- Ownership of research data: resolving the tension between confidentiality and openness; who decides?
- Intellectual property rights for open data: attribution, publication policy on new results extracted from the data
- Rewards: is the effort put into making data publicly accessible regarded as having the same kind of merit as the research that leads to a paper from the data?
- Is some data open by default? Example, data that immediately supports the claims in a publication
- · Policy issues for data repositories: access to data, access to metadata, preservation policy, data integrity, replication, etc
- Standards: Adoption/creation of international standards (metadata, formats, glossaries, indices) to facilitate search, retrieval, exploitation

• Technical issues must be solved, requires support by research funders:

· Access control and security for archives; protection of confidential data, facilitating automated access



Social issues need clarity if the conduct of science is to change in a positive way:

- Ownership of research data: resolving the tension between confidentiality and openness; who decides?
- · Intellectual property rights for open data attribution, publication policy on new results extracted from the data
- Rewards: is the effort put into making data publicly accessible regarded as having the same kind of merit as the research that leads to
 a paper from the data?
- · Is some data open by default? Example, data that immediately supports the claims in a publication
- · Policy issues for data repositories: access to data, access to metadata, preservation policy, data integrity, replication, etc
- Standards: Adoption/creation of international standards (metadata, formats, glossaries, indices) to facilitate search, retrieval, exploitation

• Technical issues must be solved, requires support by research funders:

- Access control and security for archives; protection of confidential data, facilitating automated access
- · Data integrity, preservation of provenance/attribution



Social issues need clarity if the conduct of science is to change in a positive way:

- Ownership of research data: resolving the tension between confidentiality and openness; who decides?
- · Intellectual property rights for open data: attribution, publication policy on new results extracted from the data
- Rewards: is the effort put into making data publicly accessible regarded as having the same kind of merit as the research that leads to a paper from the data?
- Is some data open by default? Example, data that immediately supports the claims in a publication
- · Policy issues for data repositories: access to data, access to metadata, preservation policy, data integrity, replication, etc
- Standards: Adoption/creation of international standards (metadata, formats, glossaries, indices) to facilitate search, retrieval, exploitation

• Technical issues must be solved, requires support by research funders:

- · Access control and security for archives; protection of confidential data, facilitating automated access
- Data integrity, preservation of provenance/attribution
- · Long-term preservation of data and access to it



Social issues need clarity if the conduct of science is to change in a positive way:

- Ownership of research data: resolving the tension between confidentiality and openness; who decides?
- Intellectual property rights for open data: attribution, publication policy on new results extracted from the data
- Rewards: is the effort put into making data publicly accessible regarded as having the same kind of merit as the research that leads to a paper from the data?
- Is some data open by default? Example, data that immediately supports the claims in a publication
- · Policy issues for data repositories: access to data, access to metadata, preservation policy, data integrity, replication, etc
- Standards: Adoption/creation of international standards (metadata, formats, glossaries, indices) to facilitate search, retrieval, exploitation

• Technical issues must be solved, requires support by research funders:

- Access control and security for archives; protection of confidential data, facilitating automated access
- Data integrity, preservation of provenance/attribution
- Long-term preservation of data and access to it
- Metadata (the richer the better!): distribution, creation of indexes, creation of federated data sets



Social issues need clarity if the conduct of science is to change in a positive way:

- Ownership of research data: resolving the tension between confidentiality and openness; who decides?
- Intellectual property rights for open data: attribution, publication policy on new results extracted from the data
- Rewards: is the effort put into making data publicly accessible regarded as having the same kind of merit as the research that leads to a paper from the data?
- Is some data open by default? Example, data that immediately supports the claims in a publication
- · Policy issues for data repositories: access to data, access to metadata, preservation policy, data integrity, replication, etc
- Standards: Adoption/creation of international standards (metadata, formats, glossaries, indices) to facilitate search, retrieval, exploitation

• Technical issues must be solved, requires support by research funders:

- Access control and security for archives; protection of confidential data, facilitating automated access
- Data integrity, preservation of provenance/attribution
- Long-term preservation of data and access to it
- · Metadata (the richer the better!): distribution, creation of indexes, creation of federated data sets
- Bandwidth and replication: the physical location of data when it is used

9

Social issues need clarity if the conduct of science is to change in a positive way:

- Ownership of research data: resolving the tension between confidentiality and openness; who decides?
- Intellectual property rights for open data attribution, publication policy on new results extracted from the data
- Rewards: is the effort put into making data publicly accessible regarded as having the same kind of merit as the research that leads to a paper from the data?
- Is some data open by default? Example, data that immediately supports the claims in a publication
- · Policy issues for data repositories: access to data, access to metadata, preservation policy, data integrity, replication, etc
- Standards: Adoption/creation of international standards (metadata, formats, glossaries, indices) to facilitate search, retrieval, exploitation

• Technical issues must be solved, requires support by research funders:

- Access control and security for archives; protection of confidential data, facilitating automated access
- Data integrity, preservation of provenance/attribution
- Long-term preservation of data and access to it
- Metadata (the richer the better!): distribution, creation of indexes, creation of federated data sets
- · Bandwidth and replication: the physical location of data when it is used
- Local or remote computing (Grid, Cloud): the physical location of the computing resources that use the data

9

Social issues need clarity if the conduct of science is to change in a positive way:

- Ownership of research data: resolving the tension between confidentiality and openness; who decides?
- Intellectual property rights for open data attribution, publication policy on new results extracted from the data
- Rewards: is the effort put into making data publicly accessible regarded as having the same kind of merit as the research that leads to a paper from the data?
- Is some data open by default? Example, data that immediately supports the claims in a publication
- · Policy issues for data repositories: access to data, access to metadata, preservation policy, data integrity, replication, etc
- Standards: Adoption/creation of international standards (metadata, formats, glossaries, indices) to facilitate search, retrieval, exploitation

• Technical issues must be solved, requires support by research funders:

- Access control and security for archives; protection of confidential data, facilitating automated access
- Data integrity, preservation of provenance/attribution
- Long-term preservation of data and access to it
- Metadata (the richer the better!): distribution, creation of indexes, creation of federated data sets
- Bandwidth and replication: the physical location of data when it is used
- Local or remote computing (Grid, Cloud): the physical location of the computing resources that use the data
- Creation of open community-usable tools to handle data, interface with local computing, perform intelligent search and retrieval

B F Schutz

· Social issues need clarity if the conduct of science is to change in a positive way:

- Ownership of research data: resolving the tension between confidentiality and openness; who decides?
- Intellectual property rights for open data attribution, publication policy on new results extracted from the data
- Rewards: is the effort put into making data publicly accessible regarded as having the same kind of merit as the research that leads to a paper from the data?
- Is some data open by default? Example, data that immediately supports the claims in a publication
- · Policy issues for data repositories: access to data, access to metadata, preservation policy, data integrity, replication, etc
- Standards: Adoption/creation of international standards (metadata, formats, glossaries, indices) to facilitate search, retrieval, exploitation

• Technical issues must be solved, requires support by research funders:

- Access control and security for archives; protection of confidential data, facilitating automated access
- Data integrity, preservation of provenance/attribution
- Long-term preservation of data and access to it
- Metadata (the richer the better!): distribution, creation of indexes, creation of federated data sets
- Bandwidth and replication: the physical location of data when it is used
- Local or remote computing (Grid, Cloud): the physical location of the computing resources that use the data
- · Creation of open community-usable tools to handle data, interface with local computing, perform intelligent search and retrieval
- · User features: ability to "browse" data before extracting it; local help desks; linking repositories with virtual research labs;

B F Schutz