uidp virtual 2021

Data Use Agreements

April 15, 2021 1:30 - 2:45 PM ET

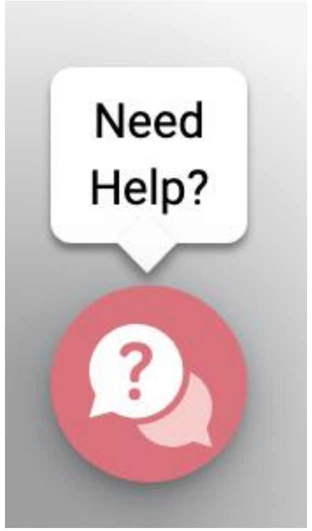
Elaine Brock
UIDP

Paul Lowe
Kansas State University

DISCLAIMER

UIDP materials, which include publications, webinars, videos, and presentations, reflect an amalgamation of the experiences and knowledge of those who participate in UIDP activities. The views and opinions expressed in UIDP materials do not necessarily reflect the official policy or position of any individual organization or the UIDP. At no time should any UIDP materials be used as a replacement for an individual organization's policy, procedures, or legal counsel. UIDP is not a lobbying organization, and UIDP materials are not intended to be used to influence government decisions.

REMO TECH SUPPORT



Having technical problems?

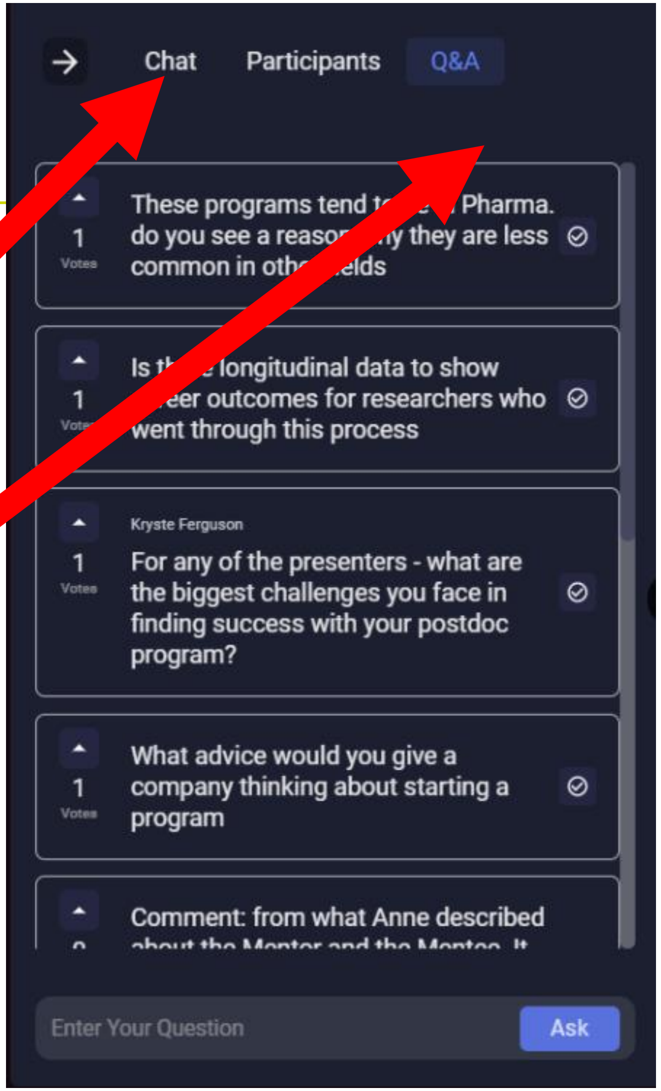
- Please use the pink “Need Help?” button at the bottom left of your screen for live chat support.

how to **PARTICIPATE**

Live Chat and Q&A

At the top right of the screen

- Chat with one another and submit ?s
- Use the Q&A tab at the top right of your screen for polls.



1 Votes — These programs tend t Pharma. do you see a reaso y they are less common in oth elds

1 Votes — Is th longitudinal data to show eer outcomes for researchers who went through this process

Kryste Ferguson
1 Votes — For any of the presenters - what are the biggest challenges you face in finding success with your postdoc program?

1 Votes — What advice would you give a company thinking about starting a program

0 — Comment: from what Anne described about the Mentor and the Mentee. It

Strengthening University-Industry Partnerships

CONTRIBUTORS

Paul R. Lowe, CRA
Associate Vice President for Research and
Director, Office of PreAward Services
Kansas State University

Elaine L. Brock, JD, MHSA
President and Senior Partner

Contracts, Compliance, and Conflict of Interest Authority LLC

Data Issues are multiplying and changing

- Personalized healthcare solutions require access and coordination of large amounts of data
- Other industry sectors use big data in product development and consumer products
- Data access is complicated by:
  - Privacy considerations
  - Social media
  - Emerging types data, .e.g., biometric identifiers
  - Convenience and reliance on data solutions
  - Ability to interconnect data from various sources
- Does social good result from data sharing?
  - Is data a “public utility”?
  - Should non-identifiable data be freely shared for the “right” purposes?

The variety of forms, sources, and uses make defining Data challenging and necessary.

- Technical data, e.g., specifications for of a motor, device, system
- Financial data
- Economic data
- Demographics
- Scientific data, e.g., weather, planetary movements
- Medical/health records
- Biometrics
- Human research subject data and healthcare data governed by HIPAA
- Animal data - health, breeding, migratory patterns
- Agriculture/plants - health, breeding, farming
- Media/pictures/recordings
- Proprietary business information
- Records from governmental agencies or corporations
- Student record information

Data Considerations Stem From These 5 Areas

1. Laws and Regulations
2. Policies, Institutional Rules
3. Contracts and Agreements
4. Ethics, Societal Concerns
5. Practicalities and Logistics

Where is the data from and where is it going?

- Different laws and regulations apply based on the jurisdictions involved, e.g.:
  - Genomic information included in PHI in some states
  - Personal data broadly defined in EU
  - Prisoner data may or may not be public
  - Salary information is public in some states
  - Biometric information is protected in some states
- GDPR - Countries can enact stricter laws and regulations and fill in gaps where the general framework is silent (e.g., for member state implementation see https://www.hhs.gov/ohrp/international/index.html).
- Transfer of data may or may not be physical (the cloud, email)
- Cultural/societal norms and expectations may affect transfer, use retention, destruction of data, IRB policies.

Provider: Can Data Be Used, Shared, Retained ?

- What data is being requested?
- Where did the data come from?
  - Patient records, research tests, interviews, measurements/sensors, analysis of other data, internet
- Who needs the data, for what purposes, and for how long?
  - Research only, clinical care, surveillance, validation, patent support, marketing, product development, integration into a product
  - Status of requestor – individual, organization, for/non-profit, HIPAA status, govt/non-govt
- Is the data confidential, proprietary, or private (identifiable patient/subject/respondent data)?
- What does the consent or other authorization say about use, sharing, and retention of data?
- Are there regulatory requirements covering collection, use, or retention of the data?
  - Human subject data (IRB, FDA, GDPR, CMS); Student data (FERPA); Export controlled data
- Are there any agreements covering data use, sharing, or retention?
  - Sponsored research agreement, MTA, DUA, NDA, data consortia membership, license
  - Website terms of use
- How will the data be transferred, stored, shared, returned, or disposed of?

User: Assess Data Needs

- Identify the data that’s needed.
- Determine who has the data.
- Assess your need for confidential or proprietary information vs. unrestricted data.
- Describe the scope of your requested use, i.e., research only, commercial use, redistribution, publication, other.
- Obtain the right to use the data
  - Website terms of use
  - Sponsored project/ data use agreement
  - Funding for project
- Secure IRB approval or other institutional approvals for use of data.
- Process agreement through authorized institutional office for negotiation and approval as applicable.

Data Considerations for Sponsored Projects SOW

- Background (pre-existing) data and related rights
- Need for data to develop SOW
- Need for confidentiality
  - Business sensitive information/data
  - Proprietary data
  - Personally Identifiable Information and data
    - Privacy of individually identifiable information, incl. health information (HIPAA)
    - Extended time limit of confidentiality protection
- Personal data, participant/subject consents - generated or provided
- Possibility, advantages of data segregation and de-identification
- Deliverables, datasets, reports, research tools to use data
- Gov’t or other funders data rights, requirements, restrictions
- Intent to publish data, research results, analysis methods, models derived from data
- Third party access – addressed in an NDA, DUA, MTA, sublicense, subcontract/sub-award
- Export Controls, Classified Data, other legal restrictions

Human Data - purposes the data will/may be used for

- Consent and consistency with terms in agreement
- Comprehensive data use clauses
- Include “primary” and “secondary” uses consistent with consent documents, applicable laws and regulations.
- “Can we?”/“Should we?”
  - Decision weighs provider’s risk threshold.
  - Restrictions and limits may be imposed to protect data subjects and parties involved.
  - Is there a cost to gather and prepare the data for transfer (time, $$, technology)?
  - Liability/litigation – risks, costs, avoidance
  - Society and ethics

Regulations Affecting Research Involving Human Data

- Common Rule (HHS 45CFR46, subpart A)
- HIPAA (Health Insurance Portability and Accountability Act)
- FDA
- 21st Century Cures Act
- FERPA (Family Educational Rights and Privacy Act)
- GDPR (General Data Protection Regulation, (EU) 2016/679)
- MACRA (Medicare Access and CHIP Reauthorization Act of 2015)

De-Identification plus…..

Common Rule - The amended Common Rule expressly defines "identifiable biospecimen" and "identifiable private information" as that for which the identity of the subject is or may readily be ascertained by the investigator or associated with the information/biospecimen.

HIPAA - De-identified health information neither identifies nor provides a reasonable basis to identify an individual. There are two ways to de-identify information; either: (1) a formal determination by a qualified statistician; or (2) the removal of the 18 specified identifiers of the individual and of the individual's relatives, household members, and employers is required, and is adequate only if the covered entity has no actual knowledge that the remaining information could be used to identify the individual.

GDPR - Pseudonymisation: the processing of personal data so that the data cannot be attributed to a specific subject without the use of additional information, if such additional information is kept separately and measures taken to ensure that the data are not attributed to an identifiable natural person. (i.e., data set may still be regulated under the GDPR if it could be re-identified with reasonable effort, e.g., coded data) [cf. Honest Broker system for Common Rule and HIPAA data]

GDPR – Anonymisation: The process of turning data into a form when it cannot be identified by any means "reasonably likely to be used ... either by the controller or by another person". (Anonymized data does not fall under GDPR)

Poll

Does your organization have a framework for dealing with GDPR and other data privacy regulations?

Yes

No

Not applicable

Mandates to Share Data…..

- FDA - Registration of Clinical trials that began enrolling participants after 1/1/19 must include a data sharing plan in the trial’s registration.
- NSF - Investigators are expected to share with other researchers, … the primary data, samples, physical collections and other supporting materials created or gathered in the course of work under NSF grants. Grantees are expected to encourage and facilitate such sharing.
- NIH - NIH expects that in drafting Data Management and Sharing Plans, researchers will maximize the appropriate sharing of scientific data, acknowledging certain factors (i.e., legal, ethical, or technical) that may affect the extent to which scientific data are preserved and shared.
- Gates Fdn - Each accepted article must be accompanied by a Data Availability Statement that describes where any primary data, associated metadata, original software, and any additional relevant materials necessary to understand, assess, and replicate the reported study findings in totality can be found.
- Check data sharing policies of funders https://v2.sherpa.ac.uk/id/funder/961 - Sherpa Juliet is a searchable database and single focal point of up-to-date information concerning research funders' policies and their requirements on open access, publication and data archiving.

Some journals require that the data be accessible for validation of results by other researchers.

- Manuscripts submitted to ICMJE reporting results of clinical trials must say:
  - will individual deidentified participant data (including data dictionaries) will be shared;
  - what data in particular will be shared;
  - will additional, related documents will be available (e.g., study protocol, statistical analysis plan)
  - when the data will become available and for how long;
  - by what access criteria data will be shared (including whom, for what analyses, and how).
- ICMJE requires authors of articles of secondary analyses using shared data to reference the source of the data using its unique, persistent identifier to credit those who generated it and to allow searching for the studies it has supported.
- Authors of articles in the Taylor and Francis journals
  - Check for data sharing policies by level of requirement https://authorservices.taylorandfrancis.com/data-sharing-policies/#datapolicies
  - Find journals in earth, space and environmental sciences participating in an open and FAIR data sharing policy as part of COPDESS (The Coalition for Publishing Data in the Earth and Space Sciences).

Data “Clubs”, Consortia, Multi-sponsor, Centers……

- Some research projects are part of larger data sharing groups that control data use and access., e.g., the Genetic Assn Information Network, the Prostate Cancer Consortium, the Robotics Research Consortium

- Institutions often enter into project agreements that pre-determine how data is gathered, stored, shared in a group, e.g., centers, projects with multiple entities involved

- These agreements should facilitate sharing data among a limited group of trusted colleagues under defined conditions.

- Universities and companies should be aware of these agreements both to facilitate their objectives and to avoid entering into conflicting agreements.

Common Data Security Provisions

- **Authorization or privilege management** – identification of individual Users who are allowed to use the Data;
- **Authentication or identity management** - confirmation that the authorized User is really the authorized User; and
- **Monitoring and enforcement** – validation and assurance that use of the Data is consistent with authorized use and conditions of use such as keeping the Data separate from other Data, in a secure location, or not on a linked computer.
- **Data Protection** – instructions regarding any special infrastructure required to store and restrict access to the Data (dedicated and isolated servers and lock-cabinets)[4]; special control processes to protect the integrity of the Data, track the location(s) of the Data, track the release of the Data and the reasons for its release; archiving and/or disposing the Data at the prescribed times.

Ownership vs Control vs Access vs Use

- Ownership, right to control, ability to access, and right to use are not always bundled as rights the provider of Data has. (Remember Data is not likely to be copyrightable unless it is presented in unique compilation that meets the copyright criteria of original work of authorship, creative, fixed in a tangible medium.)

- Ownership of data may be hard to determine particularly if there are many contributors and the collection spans a long period of time.

- Access to and use of data may be controlled based on the value of the data, the privacy of respondents, laws and regulations, e.g., HIPAA, trade secret, restrictions dictated by the owner of the data, as agreed in a contract.

- *Publicly accessible* does not mean *publicly useable* for any purpose.

Data terms should be appropriate to the situation:

- A clear description of Data to be provided.
- Permitted uses of the Data and any regulatory requirements that need to be in place.
- Names (or position descriptions) of individuals who can access or receive the Data.
- Privacy/confidentiality of information about or from subjects/respondents
- The length of time the Data may be retained or used.
- The method and timing of disposal of Data.
- Parties’ rights and obligations regarding new data generated based on the Data originally provided.

Terms (cont’d):

- Security, authorization, disclosure conditions
- Instructions on how Data should be aggregated, encrypted, anonymized, or de-identified.
- Safeguards required to protect confidentiality, privacy, sensitivity.
- Process for review by the Provider of publications resulting from use of the Data.
- Practical aspects of the Data transfer (how to).
- Statement of ownership of the Data if it is proprietary and the provenance and authenticity of the Data if required and known.
- Management of new intellectual property created using the Data.

Practical considerations about use, sharing, retention

- Does the data have inherent value that you want/need to protect?
  - Copyright - may apply to database or data contents (not to data points)
  - Property Rules - ownership of physical data instruments
  - Proprietary software needed to access or interpret
  - Too expensive to reproduce
- A university/company may enable and promote data management and sharing, but the investigator that collected the data, practically speaking, retains most control.
- University/company data sharing infrastructure may support or restrict this local control.
- University/company policies and processes promote common values and regulate behavior.
- Should you share the data even if you can?

Considerations for Data Related Projects

ISSUE ACTIVITY	Technical -Practical	Costs & Value	Monitoring & Quality	Policies	Security	Laws	Contracts	Unique Issues
Generation, Acquisition								
Characterization								
Storage								
Retrieval, Access								
Sharing								
Curating								
Archiving								
Destruction								

Poll

- Does your company or university have a person or office dedicated to reviewing data agreements?

- Yes
- No
- Wish we did

UIDP Contract Accord 14: Key Principles

1. The Data Provider is responsible for analyzing the source, sensitivity, legal and regulatory aspects of the Data to determine what provisions are needed in the DUA and its related obligations in providing the Data for the User’s intended purpose.
2. The Data User is responsible for assuring that its use is compliant with applicable regulations and it can meet the requirements imposed by the Provider in the Data Clauses.
3. The Data User should clearly explain the intended use of the Data to the Provider.
4. Data Clauses that involve performance of research by a University should include a process to allow the Provider to preview publications before public disclosure, to identify and modify or remove any sensitive Data that the Provider does not want published.
5. The Provider and User should clearly describe any special requirements, e.g., privacy, confidentiality, information security standards, that the User is expected to meet.

Case Study – The K-State 1Data Initiative

- **1Data is a platform** for mining shared data to accelerate the development of human and animal drugs, enhance the regulatory approval process, decrease the use of animal models using in silico virtual animal populations, and many other uses to advance research and technology.

- Multi-Institution initiative, led by K-State with many potential "data contributors" in collaboration with industry

- Provides a framework:
  - For collecting, integrating and simulating animal experiments to address both animal and human health concerns.
  - To enable use of incomplete or "messy" data that is not in a format currently useful for research.
  - For the collection and integration of multiple databases to develop new approaches from experimental, theoretical and computational perspectives.
  - To wed data from animal clinical trials and, eventually, human preclinical animal safety studies to develop a tool to fully simulate animal models and develop virtual models.

- **First attempt at identifying the contractual framework did not succeed.** Why?

Revelation – The K-State 1Data Initiative

- The anticipated data to be provided by industry/private sources deemed to be valuable assets closely held by the data contributors.

- Though data may have been in a dormancy and not being used, data contributors needed to see the ROI, if such data resulted in new discoveries.
  - ➢ Context – Very large and "messy" datasets accumulated over many years and many trials at high-cost R&D and testing and evaluation initiatives.
  - ➢ Data contributors deemed portions of the dataset to be "proprietary", while other portions were not considered proprietary.

- *WHEREAS, the purpose of this Agreement is to ensure the integrity **and confidentiality** of DCO Data, as defined below, as well as to **preserve the rights in** such data and any **resulting inventions** associated therewith*

- The common elements of a standard Data Use Agreement remained, but now needed to expose the normal DUA terms to steroids.

Approach – The K-State 1Data Initiative

- **Data definition**, numerous controls and stewardship approaches required for diverse subsets of Data.

  "Data" is the general reference to the collective data. "Master Data Set", "Searchable Data Set", "DCO Data", "Proprietary DCO Data Set", "Non-Proprietary DCO Data Set", defined e.g., animal health data pertinent to treatment modalities; health outcomes data; research data; care delivery data; pre-clinical data; and/or clinical stages of drug discovery data

- Clear description of the **secured environment** in which the data will reside.

  "1Data Platform" .. data, database structure, APIs, programs, and algorithms created by the KSU and UMKC 1Data Initiative that comprises the 1Data Platform of public, private, and local data that has been processed to create an environment for experimentation and simulation, **as more fully described in Attachment A**.

- Required a deliberate **process for "cleaning the data".**

  "Cleaning/Cleaned/Anonymizing/Anonymized" … processed to **remove corrupt or unverifiable data** and **normalized** for entry into the database. ….. rendered [so] **cannot refer back to any identifiable individual or group**. All Data **is de-identified.**

Approach (Cont’d) – The K-State 1Data Initiative

- To address **data security**, careful consideration of access controls was required.

  “1Data Investigator” **employed by KSU or UMKC**, as **identified in Attachment C,** executed an Individual 1Data **Investigator Confidentiality and Data Use Acknowledgment** form, “Searchable Data Set” shall be the Cleaned, Anonymized Data **selected from the Master Data Se**t defined in the **Data Use Agreement**.

- Controlled queries by type of “User”.
  - Project level queries to create a product or analyze data for usage **such that intellectual property may be created**
  - Meta-analysis queries for pre-structure analysis e.g., systematic review, cumulative meta-analysis, network meta-analysis, and prospective meta-analysis.
    - A **pre-defined list of data** that meets the criteria of the analysis; and
    - **No intellectual property is likely to result** from the query itself**.**

The Value Proposition – The K-State 1Data Initiative

- Behind-the-scenes **tagging** of all data **to facilitate attribution of discoveries to data contributor.**
    - Proprietary DCO Data Sets shall only be accessible by 1Data Investigators.
    - If a Query by a 1Data Investigator returns results ≥45% based on data from a Proprietary DCO Data Set, DCO will be notified.
    - No Data derived from a Proprietary DCO Data Set will be shared with a non-1Data Investigator without the express written consent of DCO
        - approval of a Research Request/Authorization Form; and
        - execution of a Data Use Agreement
    - Tagging process and results for a specific data contributor's data could be audited by that contributor, but they could not review the "tagging" of other data contributor's proprietary data.

- **Notice to Data Contributor** – if 1Data Investigator(s) or third-party Data Recipients makes or observes any new discovery, improvement or invention ("Invention(s)") relating to the DCO Data or as a direct result of the research

- **Broad NERF** -If patentable invention arises from the use of the DCO Data in the research, then ownership shall follow inventorship, KSU and UMKC hereby grant to DCO a worldwide, irrevocable, royalty-free, nonexclusive license under any and all Patents claiming research Inventions invented solely by the 1Data Investigator(s) or third-party Data Recipients for all purposes."

Master Data Sharing Agreement between KSU, UMKC, Data Contributing Organization

The recitals – the devil is in the details. Puts everything into context. Nice to have these upfront.

1. **Definitions – Data (including definitions of the individual types of data), Types of Users, Process Terms**
2. **Permitted Uses by Data Category**
3. **Data Use and Access by User Group**
4. The Term
5. Additional Data Procedures (Housecleaning)
6. Confidentiality
7. Custodial Responsibility, Data Stewardship, and Assurances
8. Disclaimers (Security of Contributors transmission network)
9. Data environment audits
10. Roles and Responsibilities
11. **Permissible data use, linking and sharing terms**
12. Data transfer terms and processes
13. **Operational costs**
14. Disclaimer of warranties for data or linkage quality
15. Indemnification/liability
16. **Publication**
17. Termination, Modifications, Publicity, Export Control. Governing law, etc., etc.
18. Attachments – Data list, Approved investigators, Data transfer protocols (Technical), Research request, Third-party Data Use Agreement

Data Sharing Cites to Explore

- Merck Procedure on Access to Clinical Trial Data
- Clinical Study Data Request Data Sharing Agreement Template https://www.clinicalstudydatarequest.com/
  - CSDR is a consortium of clinical study sponsors/funders with a mission to facilitate patient-level data through a research-friendly platform with independent review of proposals and patient privacy.
  - Sponsor/funder participants include: Eli Lilly, Novartis, GlaxoSmithKline, and the Bill & Melinda Gates Foundation.
- https://www.clinicalstudydatarequest.com/Documents/DATA-SHARING-AGREEMENT.pdf
- International Association of Privacy Professionals.
- Federal Demonstration Partnership, DTUA template and more
- “Data Sharing: Creating Agreements In support of community-academic partnerships” www.ucdenver.edu/research/.../community.../DataSharingCreatingAgreements.pdf

THANK YOU!

UIDP virtual 2021

- Did you enjoy the session? Rate it in the Attendee Hub!
- You'll receive a survey via email about UIDPVirtual at the end of the week. Please give us your feedback.
- Join us for one of the **next concurrent sessions at 3:25 ET**
  - **Introducing the NSF Engineering Research Visioning Alliance**
  - **Tax-Exempt Bonds & Research Contracts**