Responsible Conduct of Research Scholarship (RCRS) Training Program: 2. Intellectual Property - Data Storage and Ownership

Hyun Min Kang

April 16, 2013


GOOD LUCK ON YOUR NEW JOB!
?
!
DATA
DATA
DATA

Key areas today

Ownership Who owns the rights and responsibility of the data?

Collection Which aspects should be considered in collecting the data?

Storage How should we store and protect the data?

Sharing How should we share the data appropriately?

A Case Scenario

Scenario

John is a graduate student who performed research in Dr. K's laboratory. John is transferring to another university and would like to take the data he generated and collected in Dr. K's laboratory on Dr. K's research funds with him.

A Case Scenario

Scenario

John is a graduate student who performed research in Dr. K's laboratory. John is transferring to another university and would like to take the data he generated and collected in Dr. K's laboratory on Dr. K's research funds with him.

Questions

1. Who owns and controls the data collected in a laboratory?

A Case Scenario

Scenario

John is a graduate student who performed research in Dr. K's laboratory. John is transferring to another university and would like to take the data he generated and collected in Dr. K's laboratory on Dr. K's research funds with him.

Questions

1. Who owns and controls the data collected in a laboratory?
2. What records will John or Dr. K need to prove data ownership?

A Case Scenario

Scenario

John is a graduate student who performed research in Dr. K's laboratory. John is transferring to another university and would like to take the data he generated and collected in Dr. K's laboratory on Dr. K's research funds with him.

Questions

1. Who owns and controls the data collected in a laboratory?
2. What records will John or Dr. K need to prove data ownership?
3. Do computer records pose any unique problems in this case?

Ownership

Who owns the rights and responsibility of the data?

A possible common sense

The person who conducts the research should own the product

Ownership

Who owns the rights and responsibility of the data?

A possible common sense

The person who conducts the research should own the product

In fact

Not always - conditions are typically imposed by

1. Funders
2. Research institutions
3. Data sources

Ownership by different funders

Government funded projects

As an incentive, research institutions (not individual personnel) is given the right to use data collected with public funds. (Bayh-Role Act)

Ownership by different funders

Government funded projects

As an incentive, research institutions (not individual personnel) is given the right to use data collected with public funds. (Bayh-Role Act)

If funded by private companies

Typically retain the right to the commercial use of data

Ownership by different funders

Government funded projects

As an incentive, research institutions (not individual personnel) is given the right to use data collected with public funds. (Bayh-Role Act)

If funded by private companies

Typically retain the right to the commercial use of data

Philanthropic organizations

Retain or give away ownership rights depending on their interests.

Ownership by different funders

Government funded projects

As an incentive, research institutions (not individual personnel) is given the right to use data collected with public funds. (Bayh-Role Act)

If funded by private companies

Typically retain the right to the commercial use of data

Philanthropic organizations

Retain or give away ownership rights depending on their interests.

Private companies

Private companies seek to retain the right to the commercial use of data.

Grants and Contracts : understand the difference

If your research is supported with government funds, make sure you know whether you are working under a grant or a contract. The difference is significant and could determine who has the right to publish and use your results.

Grants and Contracts : understand the difference

If your research is supported with government funds, make sure you know whether you are working under a grant or a contract. The difference is significant and could determine who has the right to publish and use your results.

Grants

Typically control of the data remains with the institution that received the funds

Grants and Contracts : understand the difference

If your research is supported with government funds, make sure you know whether you are working under a grant or a contract. The difference is significant and could determine who has the right to publish and use your results.

Grants

Typically control of the data remains with the institution that received the funds

Contracts

- Require the researcher to deliver a product or service,
- Then usually the outcome is owned and controlled by the government.

Ownership by research institutions

- Support for research is typically awarded to research institutions, not to individual researchers

Ownership by research institutions

- Support for research is typically awarded to research institutions, not to individual researchers
- Research institutions have responsibilities for budgets, regulatory compliance, contractual obligations, and data management.

Ownership by research institutions

- Support for research is typically awarded to research institutions, not to individual researchers
- Research institutions have responsibilities for budgets, regulatory compliance, contractual obligations, and data management.
- This means that researchers cannot automatically assume that they can take their data with them if they move to another institution.

Ownership by research institutions

- Support for research is typically awarded to research institutions, not to individual researchers
- Research institutions have responsibilities for budgets, regulatory compliance, contractual obligations, and data management.
- This means that researchers cannot automatically assume that they can take their data with them if they move to another institution.
- Data and data books collected by students and postdoctoral fellows on a research project belong to the grantee institution

Ownership by research institutions

- Support for research is typically awarded to research institutions, not to individual researchers
- Research institutions have responsibilities for budgets, regulatory compliance, contractual obligations, and data management.
- This means that researchers cannot automatically assume that they can take their data with them if they move to another institution.
- Data and data books collected by students and postdoctoral fellows on a research project belong to the grantee institution
- Students should not take their data when they leave without making appropriate arrangements. Retaining copies of data is allowed with permission, and although this is not always done, it is certainly good practice.

Example : U of Pittsburgh Guidelines

- Both the PI and the University have responsibilities/rights to use original research data
  - But it belongs to the University of Pittsburgh,

Example : U of Pittsburgh Guidelines

- Both the PI and the University have responsibilities/rights to use original research data
  - But it belongs to the University of Pittsburgh,
- The investigator may be allowed to retain the research records and materials created by him/her.

Example : U of Pittsburgh Guidelines

- Both the PI and the University have responsibilities/rights to use original research data
  - But it belongs to the University of Pittsburgh,
- The investigator may be allowed to retain the research records and materials created by him/her.
- If the investigator leaves the University, an agreement on disposition of research data may be negotiated.
  - However, university has the right of access to all research records and materials after reasonable prior notice

Example : U of Pittsburgh Guidelines

- Both the PI and the University have responsibilities/rights to use original research data
  - But it belongs to the University of Pittsburgh,
- The investigator may be allowed to retain the research records and materials created by him/her.
- If the investigator leaves the University, an agreement on disposition of research data may be negotiated.
  - However, university has the right of access to all research records and materials after reasonable prior notice
- When a collaborative team is dissolved, typically each member of the team should have continuing access to the data and materials .

Ownership by data sources

- Unique resources (e.g. tropical rain forests, rare medical conditions)
  - Countries or research subject may claim ownership of research results based on their data.
  - Or they may want some ownership stake in the end results

Ownership by data sources

- Unique resources (e.g. tropical rain forests, rare medical conditions)
  - Countries or research subject may claim ownership of research results based on their data.
  - Or they may want some ownership stake in the end results
- Make sure you can answer the following questions:
  - Who owns the data I am collecting?
  - What rights do I have to publish the data?
  - Does collecting these data impose any obligations on me?

Data Sources - Approval of Ownership

- In most cases ownership provisions must be approved by the institution and is responsible for the administration of research funds.

Data Sources - Approval of Ownership

- In most cases ownership provisions must be approved by the institution and is responsible for the administration of research funds.
- Researchers therefore should not enter into agreements that affect the control and use of data without getting institutional approval.

Data Sources - Approval of Ownership

- In most cases ownership provisions must be approved by the institution and is responsible for the administration of research funds.
- Researchers therefore should not enter into agreements that affect the control and use of data without getting institutional approval.
- The results could be disastrous and expensive if ownership is disputed later

Summary - Ownership

- Ownership of data substantially varies by the funders

Summary - Ownership

- Ownership of data substantially varies by the funders
- For government funded projects, grants and contracts can make significant difference in ownership

Summary - Ownership

- Ownership of data substantially varies by the funders
- For government funded projects, grants and contracts can make significant difference in ownership
- Research institution (+ principal investigator) typically has the rights and responsibilities for the data

Summary - Ownership

- Ownership of data substantially varies by the funders
- For government funded projects, grants and contracts can make significant difference in ownership
- Research institution (+ principal investigator) typically has the rights and responsibilities for the data
- Special data sources may limit the ownership of data.

Summary - Ownership

- Ownership of data substantially varies by the funders
- For government funded projects, grants and contracts can make significant difference in ownership
- Research institution (+ principal investigator) typically has the rights and responsibilities for the data
- Special data sources may limit the ownership of data.
- Clarify ownership of data, right to publish the data, and imposed obligations before working on the data.

Summary - Ownership

- Ownership of data substantially varies by the funders
- For government funded projects, grants and contracts can make significant difference in ownership
- Research institution (+ principal investigator) typically has the rights and responsibilities for the data
- Special data sources may limit the ownership of data.
- Clarify ownership of data, right to publish the data, and imposed obligations before working on the data.
- Approval of ownership must be done by the institution, not individual researchers.

Principles of Data Collection - Use Appropriate Methods

- Reliable data are vitally dependent on reliable methods.
- Example : If you didn't observe a side effect of a drug from 5 people, you may not conclude no adverse effect.

Principles of Data Collection - Attention to detail

- A failure to pay attention to detail can result in serious mistakes

Principles of Data Collection - Attention to detail

- A failure to pay attention to detail can result in serious mistakes
- Obviously, it is not possible to avoid all mistakes in research, but learn from others' experiences (errata or colleague's experiences)

Principles of Data Collection - Attention to detail

- A failure to pay attention to detail can result in serious mistakes
- Obviously, it is not possible to avoid all mistakes in research, but learn from others' experiences (errata or colleague's experiences)
- Check common mistakes, for example..
  - Make sure that each figure was correctly labeled
  - Were the calculations double checked?
  - Authors were properly listed?

Principles of Data Collection - Attention to detail

- A failure to pay attention to detail can result in serious mistakes
- Obviously, it is not possible to avoid all mistakes in research, but learn from others' experiences (errata or colleague's experiences)
- Check common mistakes, for example..
  - Make sure that each figure was correctly labeled
  - Were the calculations double checked?
  - Authors were properly listed?
- Researchers have a responsibility to make sure their work is carefully undertaken and reported.

Duke Suspends Researcher and Halts Cancer Studies

(by Natasha Singer, New York Times, July 20, 2010)

Duke Suspends Researcher and Halts Cancer Studies

(by Natasha Singer, New York Times, July 20, 2010)

- Duke University School of Medicine has suspended a researcher and stopped patient enrollment in three cancer studies upon learning of reports that the researcher had overstated his academic credentials. The lead researcher, Dr. Anil Potti, was placed on administrative leave ... while it investigates allegations that Dr. Potti falsely claimed to have been a Rhodes scholar.

Duke Suspends Researcher and Halts Cancer Studies

(by Natasha Singer, New York Times, July 20, 2010)

- Duke University School of Medicine has suspended a researcher and stopped patient enrollment in three cancer studies upon learning of reports that the researcher had overstated his academic credentials. The lead researcher, Dr. Anil Potti, was placed on administrative leave ... while it investigates allegations that Dr. Potti falsely claimed to have been a Rhodes scholar.
- The Cancer Letter, a weekly publication for cancer specialists, reported that Dr. Potti, an Assistant Professor of medicine, had padded his resume on occasion. A spokeswoman at Rhodes House at Oxford University confirmed on Tuesday that Dr. Potti had not received a (Rhodes) scholarship.

Duke Suspends Researcher and Halts Cancer Studies

- The American Cancer Society suspended payments on a five-year, $729,000 grant awarded to Dr. Potti to study the genetics of lung cancer. The society issued the grant based in part on Dr. Potti's resume, which included a Rhodes scholarship, said Dr. Otis W. Brawley, the chief medical officer of the cancer society. Dr. Potti had been featured in several promotional videos for the research at Duke.

Duke Suspends Researcher and Halts Cancer Studies

- The American Cancer Society suspended payments on a five-year, $729,000 grant awarded to Dr. Potti to study the genetics of lung cancer. The society issued the grant based in part on Dr. Potti's resume, which included a Rhodes scholarship, said Dr. Otis W. Brawley, the chief medical officer of the cancer society. Dr. Potti had been featured in several promotional videos for the research at Duke.
- The studies involve a prediction model in which genetic analysis is used to determine which chemotherapy drugs would work best for particular cancer patients. ... A physician should be able to biopsy a patient's lung tumor, for example, and, using the prediction model to determine the most effective treatment, choose the chemotherapy drugs for a patient.

Duke Suspends Researcher and Halts Cancer Studies

- But the Duke researchers' data analysis, published in influential medical journals like The New England Journal of Medicine and Nature Medicine, has come under increasing scrutiny by some outside experts.

Duke Suspends Researcher and Halts Cancer Studies

- But the Duke researchers' data analysis, published in influential medical journals like The New England Journal of Medicine and Nature Medicine, has come under increasing scrutiny by some outside experts.
- Last year, two biostatisticians at the University of Texas M.D. Anderson Cancer Center published an article in the Journal Annals of Applied Statistics in which they identified errors in Duke's data analysis and said they had not been able to reproduce Duke's results. ... The article said that the errors in the Duke research might result in doctors assigning wrong therapies to patients, potentially putting them at risk.

Duke Suspends Researcher and Halts Cancer Studies

- But the Duke researchers' data analysis, published in influential medical journals like The New England Journal of Medicine and Nature Medicine, has come under increasing scrutiny by some outside experts.
- Last year, two biostatisticians at the University of Texas M.D. Anderson Cancer Center published an article in the Journal Annals of Applied Statistics in which they identified errors in Duke's data analysis and said they had not been able to reproduce Duke's results. ... The article said that the errors in the Duke research might result in doctors assigning wrong therapies to patients, potentially putting them at risk.
- ... The Defense Department paid for the breast cancer study. The National Cancer Institute financed one of the lung cancer studies and Eli Lilly, the drug maker, paid for the other.

Principles of Data Collection - Check permissions

- Many types of data collection need to be authorized before they can proceed.

Principles of Data Collection - Check permissions

- Many types of data collection need to be authorized before they can proceed.
- Typically permission is needed to use for..
  - human and animal subjects in research;
  - hazardous materials and biological agents;
  - information with copyrights (library, database, web site, photographs)

Principles of Data Collection - Check permissions

- Many types of data collection need to be authorized before they can proceed.
- Typically permission is needed to use for..
  - human and animal subjects in research;
  - hazardous materials and biological agents;
  - information with copyrights (library, database, web site, photographs)
- Researchers have a responsibility to know when permission is needed to collect or use specific data in their research.

Principles of Data Collection - Check permissions

- Many types of data collection need to be authorized before they can proceed.
- Typically permission is needed to use for..
  - human and animal subjects in research;
  - hazardous materials and biological agents;
  - information with copyrights (library, database, web site, photographs)
- Researchers have a responsibility to know when permission is needed to collect or use specific data in their research.
- If you are not sure whether permission is needed, check before proceeding with data collection.

Principles of Data Collection - Record as much as possible

- Physically retain record the data in some type of
  - notebook (hard copy),
  - computer file (electronic copy),
  - or other permanent "record" of the work done.

Principles of Data Collection - Record as much as possible

- Physically retain record the data in some type of
  - notebook (hard copy),
  - computer file (electronic copy),
  - or other permanent "record" of the work done.
- Purpose of any record is to document what was actually done and the results that were achieved.

Principles of Data Collection - Record as much as possible

- Physically retain record the data in some type of
  - notebook (hard copy),
  - computer file (electronic copy),
  - or other permanent "record" of the work done.
- Purpose of any record is to document what was actually done and the results that were achieved.
- Keep two simple rules in mind to avoid problems later:
  - Hard-copy evidence using a numbered, bound notebook.
    - Do not use loose-leaf notebooks or collect pages of evidence in a file.
    - Do not change records in a bound notebook without noting the date and reasons for the change.

Principles of Data Collection - Record as much as possible

- Physically retain record the data in some type of
  - notebook (hard copy),
  - computer file (electronic copy),
  - or other permanent "record" of the work done.
- Purpose of any record is to document what was actually done and the results that were achieved.
- Keep two simple rules in mind to avoid problems later:
  - Hard-copy evidence using a numbered, bound notebook.
    - Do not use loose-leaf notebooks or collect pages of evidence in a file.
    - Do not change records in a bound notebook without noting the date and reasons for the change.
  - Electronic evidence should be validated in some way to assure that it was actually recorded on a particular date and not changed at some later date.

Summary - Data Collection

- Use appropriate methods

Summary - Data Collection

- Use appropriate methods
- Attention to detail

Summary - Data Collection

- Use appropriate methods
- Attention to detail
- Check authorization principles

Summary - Data Collection

- Use appropriate methods
- Attention to detail
- Check authorization principles
- Record as much as possible

Data Storage and Protection

- Data must be properly protected.
  - to confirm research findings,
  - to be reanalyzed by other researchers.

Data Storage and Protection

- Data must be properly protected.
  - to confirm research findings,
  - to be reanalyzed by other researchers.
- Proper storage and protection from accidental damage, loss, or theft:
  - Lab notebooks should be stored in a safe place.
  - Computer files should be backed up in a secure place (physically distant to original data)
  - Samples should be appropriately saved so that they will not degrade over time.
  - Care should be taken to reduce the risk of fire, flood, and other catastrophic events.

Confidentially

- If only authorized individuals have access to the data (e.g. personal information or patent material), care needs to be taken to assure that privacy agreements are honored.

Confidentially

- If only authorized individuals have access to the data (e.g. personal information or patent material), care needs to be taken to assure that privacy agreements are honored.
- Must be stored in a safe place that is accessible only to authorized personnel.

Confidentially

- If only authorized individuals have access to the data (e.g. personal information or patent material), care needs to be taken to assure that privacy agreements are honored.
- Must be stored in a safe place that is accessible only to authorized personnel.
- Using random codes to identify individual subjects, rather than names or social security numbers, can also further protect private information.

Confidentially

- If only authorized individuals have access to the data (e.g. personal information or patent material), care needs to be taken to assure that privacy agreements are honored.
- Must be stored in a safe place that is accessible only to authorized personnel.
- Using random codes to identify individual subjects, rather than names or social security numbers, can also further protect private information.
- Access to these codes can then be restricted to provide a double layer of protection.

Confidentially

- If only authorized individuals have access to the data (e.g. personal information or patent material), care needs to be taken to assure that privacy agreements are honored.
- Must be stored in a safe place that is accessible only to authorized personnel.
- Using random codes to identify individual subjects, rather than names or social security numbers, can also further protect private information.
- Access to these codes can then be restricted to provide a double layer of protection.
- Whatever the method used to protect private or confidential information, the researcher who collects or uses the information has the primary responsibility for its protection.

Period of retention

- Data should be retained for a "reasonable" period of time to allow other researchers to check results or to use the data for other purposes.

Period of retention

- Data should be retained for a "reasonable" period of time to allow other researchers to check results or to use the data for other purposes.
- NIH generally requires that data be retained for 3 years following the submission of the final financial report.

Period of retention

- Data should be retained for a "reasonable" period of time to allow other researchers to check results or to use the data for other purposes.
- NIH generally requires that data be retained for 3 years following the submission of the final financial report.
- Some government programs require retention for up to 7 years.

Period of retention

- Data should be retained for a "reasonable" period of time to allow other researchers to check results or to use the data for other purposes.
- NIH generally requires that data be retained for 3 years following the submission of the final financial report.
- Some government programs require retention for up to 7 years.
- How long is reasonable will vary from field to field and institution to institution.

Period of retention

- Data should be retained for a "reasonable" period of time to allow other researchers to check results or to use the data for other purposes.
- NIH generally requires that data be retained for 3 years following the submission of the final financial report.
- Some government programs require retention for up to 7 years.
- How long is reasonable will vary from field to field and institution to institution.
- Nevertheless, it is important to have a clear retention policy

Summary - Storage

- Researcher who collects or use the information has the primary responsibility for its protection

Summary - Storage

- Researcher who collects or use the information has the primary responsibility for its protection
- Data must be properly protected to ensure safety.

Summary - Storage

- Researcher who collects or use the information has the primary responsibility for its protection
- Data must be properly protected to ensure safety.
- Data must be accessible only to authorized personnel.

Summary - Storage

- Researcher who collects or use the information has the primary responsibility for its protection
- Data must be properly protected to ensure safety.
- Data must be accessible only to authorized personnel.
- Data should be retained for a reasonable period of time (e.g. 3-7 years).

Data Sharing

- Research data should be shared,

Data Sharing

- Research data should be shared,
- Researchers are not expected to and in most instances should not release preliminary data
  - Except for when preliminary data that could potentially benefit the public. (e.g. unexpected side effects from a drug)

Data Sharing

- Research data should be shared,
- Researchers are not expected to and in most instances should not release preliminary data
  - Except for when preliminary data that could potentially benefit the public. (e.g. unexpected side effects from a drug)
- Other types of data (e.g. basic scientific discovery) can be held until the researcher is confident that the results will stand.

Data Sharing

- Researchers can withhold confirmed or validated data until publication
  - Keeping data confidential prior to publication is a commonly accepted practice that most researchers and funding agencies accept.
  - Upon publication, all the information about that experiment, should be freely available

Data Sharing

- Researchers can withhold confirmed or validated data until publication
  - Keeping data confidential prior to publication is a commonly accepted practice that most researchers and funding agencies accept.
  - Upon publication, all the information about that experiment, should be freely available
- In some federally funded research data must be made available in response to Freedom of Information Act (FOIA) requests.

Data Sharing

- Researchers can withhold confirmed or validated data until publication
  - Keeping data confidential prior to publication is a commonly accepted practice that most researchers and funding agencies accept.
  - Upon publication, all the information about that experiment, should be freely available
- In some federally funded research data must be made available in response to Freedom of Information Act (FOIA) requests.
- Researches should provide considerable support for sharing data with other researchers and the public unless there are compelling reasons for confidentiality.

Summary - Sharing

- Publication quality research data should be shared.

Summary - Sharing

- Publication quality research data should be shared.
- Data should typically be shared upon publication
  - But some federally funded research data should be shared before publication.

Summary - Sharing

- Publication quality research data should be shared.
- Data should typically be shared upon publication
  - But some federally funded research data should be shared before publication.
- Support for sharing data with other research is necessary.

Future Considerations : Complexity

- Big Data era : data generation outstrip to capacity to score (e.g. genomics or image data)

Future Considerations : Complexity

- Big Data era : data generation outstrip to capacity to score (e.g. genomics or image data)
- Even when researchers want to, it is not always clear how they should go about collecting, storing, and sharing data responsibly.

Future Considerations : Control

- Control of data is important particularly when financial interests are at stake.

Future Considerations : Control

- Control of data is important particularly when financial interests are at stake.
- In multi-site clinical trials, for example, right to retain some control over their own data vs. centralized data management
  - This issue is currently unresolved and the subject of intense public debate.

Future Considerations : National security

- Possible use of data from publicly supported research by terrorists and nations that could pose a threat to national security.

Future Considerations : National security

- Possible use of data from publicly supported research by terrorists and nations that could pose a threat to national security.
- Efforts are underway to address these concerns through voluntary policies and new Federal regulations

Future Considerations : National security

- Possible use of data from publicly supported research by terrorists and nations that could pose a threat to national security.
- Efforts are underway to address these concerns through voluntary policies and new Federal regulations
- Researchers whose work could be affected by these concerns should keep abreast of ongoing policy development and regulation.

Future Considerations : National security

- Possible use of data from publicly supported research by terrorists and nations that could pose a threat to national security.
- Efforts are underway to address these concerns through voluntary policies and new Federal regulations
- Researchers whose work could be affected by these concerns should keep abreast of ongoing policy development and regulation.
- Ask yourself how someone funding your research would feel if he or she had a chance to take a close look at your data management practices.

Summary - Key areas today

Ownership Who owns the rights and responsibility of the data?

Collection Which aspects should be considered in collecting the data?

Storage How should we store and protect the data?

Sharing How should we share the data appropriately?

Questions for discussion

1. Should research data belong to researchers rather than to research institutions?
2. Should data recording practices be standardized to facilitate sharing and monitoring? What recording practices could be standardized?
3. What interpretation practices could be standardized? How does your laboratory verify the accuracy and validity of data before its disclosure or use in grant proposals and publications?
4. Who should pay the cost of sharing data? Who should have access to the data?
5. How long should researchers be able to withhold data to allow time to protect ownership claims? How long should research data be stored?