## Research Data Management Case: Characterizing a Component of a Rocket Engine used to Control Satellites in Orbit

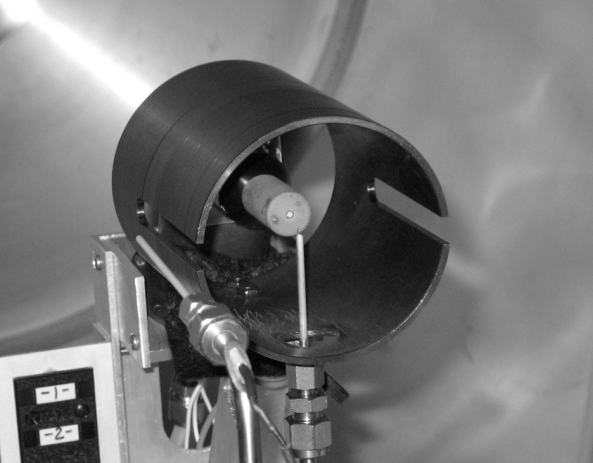
 

Fig. 1 50” x 72” vacuum chamber used for cathode research (WPI Higgins Laboratory)

Fig.2. Hollow cathode and surrounding anode installed in vacuum chamber. Also visible are Langmuir and emissive probes (small white rods).

Data Issues Addressed (Modules):

* Data Management/Ownership
* Data Life Cycles
* Data Storage
* Data Security
* Data Privacy/Restrictions

Suggested Curriculum Levels:

* Undergraduate, years 3-4
* Graduate

Suggested student populations:

* Engineering
* Lab Experimentation – Physical Sciences

Scenario:

A faculty researcher in Aerospace Engineering studies electric propulsion for spacecraft control, in other words a type of rocket engine that uses electricity to ionize and accelerate a gas to produce thrust. He is concerned about data storage and security in his lab and is looking for a standard protocol that could be used by all and would comply with any data storage and security requirements imposed by his research sponsors.

The goal of one of his current projects is to study and characterize a component of an electric thruster being used by NASA, the Air Force, and private companies to control satellites in orbit. This work enables researchers to build more robust thrusters that will have a longer service life than current models, thus enabling longer and more ambitious space missions. The students in the lab have been experimenting on a particular engine component, called a “hollow cathode,” to characterize the plasma it generates. This data will help researchers understand where energetic ions are produced that erode surfaces and limit the cathode lifetime. They perform experiments using Langmuir and emissive probes to collect data from two different cathodes tested in a vacuum chamber.

The two cathodes used in the lab have different restrictions on their use. The first is from a private company, was developed with Air Force funding, and is the same model as a unit which has been used operationally in orbit. Because it is identical to flight hardware, work with this cathode must comply with International Traffic in Arms Regulations (ITAR) requiring that no foreign nationals have access to any aspect of the research. ITAR establishes strict controls on the use and dissemination of information related to defense articles. Some equipment and lab notebooks are kept locked up in order to comply.

The second cathode is from NASA, is a laboratory use model (i.e. not “flight hardware”), and is not subject to the same ITAR restrictions. However, NASA does place some restrictions on publication related to the cathode design, such as specific geometry and dimensions. This makes it difficult to publish this type of research and data since the experiment would not be reproducible by other researchers unless they have access to the same devices.

Raw data is generated from the cathodes during experiments and is downloaded onto a laptop. The experimental conditions are recorded in the laboratory notebook. Students review the data and discard any that is not useful. For useful data, the students produce code using MATLAB and create appropriate graphs and charts displaying the useful data points. They share findings with the researcher so that all can discuss and determine whether there are reportable results that will be useful to the aerospace community. The MATLAB code is the intellectual property of the student and faculty advisor who create it. If code is written primarily by a student, then the advisor will usually try to contact the student (who may have graduated) for permission before distributing it to another research group. Permission is rarely granted for sharing this code outside the research group.

Over the past two years the researcher has had four graduate students work on this project. Two of the students have graduated and are no longer part of the team.

When students graduate they are required to provide the researcher with a CD containing the MATLAB code used for any analysis they completed in the lab. These CDs are then stored in the researchers office in case of a future need. This code is not always sufficiently documented so that it could be easily used in the future. Some students do a better job than others at describing how their code works and the researcher is interested in learning about how to standardize this practice, although he has never had a problem in the past.

Laboratory notebooks are generally kept in the laboratory. Notebooks which include any information on projects which are ITAR restricted are kept in a secure location, either in the researcher’s office or in a locked storage cabinet located in a lab. The archive CDs are stored in the researcher’s office and there is no backup plan for either the CDs or the notebooks. There is also a minor concern over the fact that laboratory notebook entries are lacking some information and may need to be more standardized. In some cases there are no initials indicating who completed an entry in the notebook. Other descriptive elements may be missing as well.

The only data that is retained is that which has been used to generate results presented in either a thesis or a publication. This data is stored on the researcher’s computer which is automatically backed up nightly through his institution.

**Module 1 (Overview module) discussion question:** What issues need to be addressed on this project related to the 7 segments of the data management plan components?

**Discussion Questions for Other Modules:**

1. Types of data

* 1. What types of data are being collected for this study?
  2. How will the data be captured?
  3. How will you ensure all research staff used the same data sources and data definitions?
  4. What needs to be in the plan related to the data captured from testing the two different cathodes?
  5. What analytical methods and mechanisms will be applied to your data either prior to or post integration?
  6. What type of outcome data will be generated?

1. File Formats and Contextual details
2. What file formats and naming conventions will be used for the separate data sources and the MATLAB code? What impact would the naming conventions have on later data access?
3. What other contextual details would you specifically need to document to make your data meaningful to others?
4. In what form will you capture these details?
5. Data Storage, Backup, Security
6. Where and on what media will the data from each data source be stored?
7. How will you manage data security across research staff on the study for each data source?
8. What concerns are there regarding the security of the data that is kept on a CD, laptop, or in the lab notebook?
9. How long following the completion of your study will you store the data?
10. Data protection/privacy
11. How are you addressing any ethical or privacy issues?
12. Who will own any copyright or intellectual property rights to the data from each source?
13. How will the data be associated with a study ID?
14. Policies for reuse of data
15. Will you need to create a de-identified copy of the data?
16. Will the data be restricted to be re-used only for certain purposes or by specific researchers?
17. Are there any reasons not to share or re-use data?
18. Policies for access and sharing
19. Will some kind of contribution or fee be charged for subsequent access to this data?
20. What process should be followed to gain future access to your study data?
21. Archiving and preservation
22. What is the long-term strategy for maintaining, curating and archiving the data?
23. What data will be included in an archive?
24. Where and how will it be archived?
25. What other contextual data or other related data will be included in the archive?
26. How long will the data be kept beyond the life of the project?