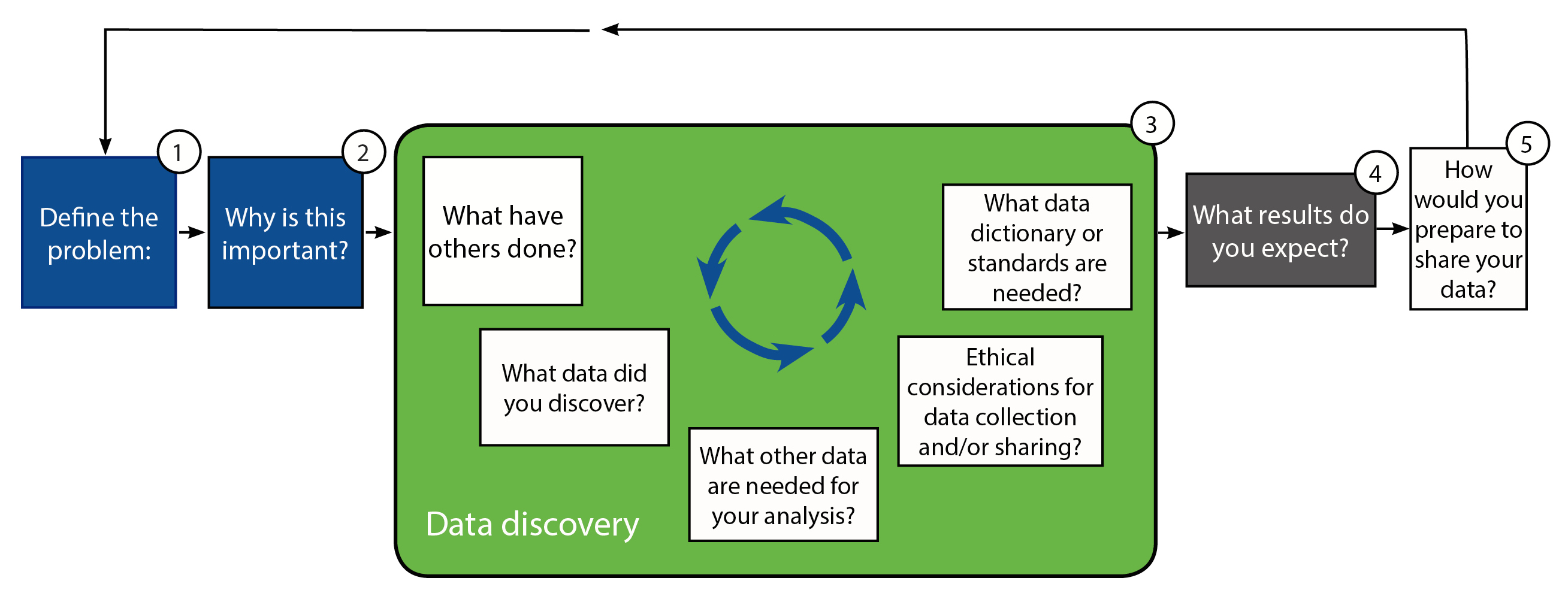
# You are a data scientist! Tell us about it!



**Roles:**

1. **Politician** – How is this going to help my constituents and what about the cost and my reputation? (cost-effective per person)
2. **Multi-national CEO** – What is this doing for my brand?
3. **Jr. Researcher (outside of your specialty)** – How will this change my approach to the problem I am examining?
4. **Higher-Ed Instructor** – What changes will this cause in the curriculum or student engagement?
5. **Business entrepreneur**- What other data is needed to make this solution profitable?
6. **Your mom** – Why is the problem important to the layperson?
7. **NIH Program Officer** – How would you plan your analysis?

**Choose your role here:** [**http://bit.ly/1l4vxit**](http://bit.ly/1l4vxit)

**Hillary Clinton’s email:**

Throughout 2015, Hillary Clinton has been embroiled in [controversy](https://en.wikipedia.org/wiki/Hillary_Clinton_email_controversy) over the use of personal email accounts on non-government servers during her time as the United States Secretary of State. Some political experts and opponents maintain that Clinton's use of personal email accounts to conduct Secretary of State affairs is in violation of protocols and federal laws that ensure appropriate recordkeeping of government activity. Hillary's campaign has provided their own four sentence summary of her email use [here](https://www.hillaryclinton.com/hillarys-emails-four-sentences/).

There have been a number of Freedom of Information [lawsuits](https://en.wikipedia.org/wiki/Hillary_Clinton_email_controversy#Freedom_of_Information_lawsuits) filed over the State Department's failure to fully release the emails sent and received on Clinton's private accounts. On Monday, August 31, the State Department released nearly 7,000 pages of Clinton's heavily redacted emails (its biggest release of emails to date).

The documents were released by the State Department as PDFs. We've cleaned and normalized the released documents and are hosting them for public analysis. Kaggle's choice to host this dataset is not meant to express any particular political affiliation or intent.

<https://www.kaggle.com/kaggle/hillary-clinton-emails>

http://bit.ly/1W5sR1x

**Mechanosensors that detect and treat lung fibrosis:**

Idiopathic Pulmonary Fibrosis, or IPF, is a terminal disease affecting as many as 500,000 Americans with no FDA-approved therapies capable of stopping disease progression. The disease is characterized by excessive assembly of extracellular matrix (ECM) by activated fibroblasts termed `myofibroblasts'. Recently, studies have demonstrated that tissue mechanics, specifically tissue stiffness resulting from myofibroblasts assembly of ECM and contraction, is capable of driving the differentiation of myofibroblasts and thus disease progression. In short, myofibroblasts are capable of recruiting more myofibroblasts leading to a disease that progresses unchecked. Despite these recent findings we still do not understand how the process is initiated, nor do we have any therapies that effective halt disease progression. Basic molecular mechanisms for how cells "sense" this stiffness ("mechano-sensing") have, however, been developed and identified. In this project we are harnessing the same cellular mechanisms that allow them to sense the increased stiffness toward the development of technology for delivering local, scar-directed therapy with the goal of treating and curing pulmonary fibrosis directly. This approach leverages years of scientific insight into the basis for mechanical sensing by cells and co-opts naturally occurring paradigms to turn the disease state against itself.

<https://projectreporter.nih.gov/project_description.cfm?projectnumber=1R01HL132585-01>

http://1.usa.gov/1UNlMBH

**Is it possible to overcome the resistance of tumors to radiotherapy?:**

Background: Because the area and dose of radiation can be precisely controlled, radiation therapy is an appealing approach for the treatment of diverse types of cancer. However, it is unclear why some cancers respond well and others do not. For example, the local control rate with radiation is nearly 100% for testicular cancer but only 10% for glioblastoma. We do not understand how to explain these differences, whether and how the sensitivities of resistant cells can be modified, or whether there are also mechanisms for acquisition of secondary resistance.

Feasibility: Technologies are available to characterize the genetic and epigenetic makeup of radioresistant and radiosensitive cancer cells, to delineate biochemical and physiological pathways involved in radioresistance (e.g. secretion of cytokines or other stress responses), and to test the impact of various radiosensitizers on radioresistance, such as concordant use of traditional chemotherapies or targeted drugs. Improved imaging methods could be used to follow resistance in small lesions, and new animal models might provide opportunities for understanding radioresistance. Several important questions could be addressed: Do common mechanisms mediate targeted drug resistance, chemotherapy resistance, and radioresistance? Is the basis of radioresistance genetic and/or epigenetic? What adaptive responses are unique to resistant cells? Is there a non-cell autonomous component to radioresistance?

Implications of success: Understanding radioresistance could substantially improve patient outcome for multiple types of cancer, especially if the results led to the identification of well tolerated potentially effective radiosensitizers that could be tested in prospective clinical trials.

<http://provocativequestions.nci.nih.gov/community-dialog/us-workshops/provocative-questions>

http://1.usa.gov/1l4vwuZ