# Discovery and Requirements Analysis

# Preliminary Research

In preparation for discovery with the user community, given the short proposal timeline, the Product Manager conducted preliminary research involving an open web search for challenges, risks and disruptions involving FDA and Regulated Industry. Five articles were reviewed, resulting in the definition of a candidate context for persona and user scenarios.

## Research articles:

1. “Makers of Generic Drugs Challenge F.D.A. Plan for Updated Warnings”, New York Times, March 27, 2015. <http://www.nytimes.com/2015/03/28/science/makersofgenericdrugschallengefdaplanforupdatedwarnings.html?_r=0>
2. “Challenges and Opportunities for the Generic Drug Industry”, Remarks at the GPhA Annual Meeting as Delivered by Margaret A. Hamburg, M.D., FDA Commissioner, February 23, 2012. <http://www.fda.gov/NewsEvents/Speeches/ucm294978.htm>
3. “Regulation in the Medical Device Industry: FDA insider lays out the issues”, Metropolitan Corporate Counsel, March 19, 2015. <http://www.metrocorpcounsel.com/articles/31944/regulationmedicaldeviceindustryfdainsiderlaysoutissues>
4. “Device Tax Repeal, FDA’s Actions Affecting Industry Remain Top Issues”, Bloomberg Bureau of National Affairs, January 1, 2015. <http://www.fr.com/wp-content/uploads/2015/01/Bloomberg-BNA.Medical-Device-Law-Industry-Report.January21.2015.pdf>
5. “Project FDA Report: The Digital Future of Molecular Medicine: Rethinking FDA Regulation”, Manhattan Institute, May 2013. http://www.manhattan-institute.org/pdf/fda\_06.pdf

## Preliminary Enterprise Context:

* **Candidate Capability:** The FDA Administrator needs the ability to readily understand and be able to intuitively demonstrate the impact of enhancing product labels on adverse drug events.
* **Why Important:** FDA has been an advocate of more stringent drug labelling, and is now involved with industry in a controversial effort to impose greater labelling requirements on manufacturers of generic drugs.  The value/cost of this is the subject of debate in the news, so an aid in researching the value of labelling may be of interest to key stakeholders.
* **Candidate Persona/Scenario**: An FDA Analyst wants to determine if there any correlations exist between changes in drug labelling and a reduction in adverse events across all drugs.
* **Candidate Person/Scenario:** A Pharmaceutical Researcher wants to determine if there is any correlation between changes in drug labelling and a reduction in adverse events for a specific drug or related drugs.

## Preliminary Front End Data Analysis and Visualization Activities:

To help orient the Front End Web Designer (Data Scientist) on the problem area, the Product Manager looked at supporting data and developed a set of data analysis and visualization activities to enable data exploration for supporting the candidate scenarios.

* Identifying events where a product label changed for a drug, and extracting the drug and date of change. Notes:
  + The drug label change events are accessible [here](http://dailymed.nlm.nih.gov/dailymed/spl-resources-all-drug-labels.cfm) on an HHS site.  The Open FDA data describing Structured Product Labelling (SPL) is all about establishing a standardized data structure (XML) format for the data, but the data itself in that format doesn’t seem to be on the Open FDA site.  Look at the SPL structure to determine the field that contains the date of change.  On the label site, there appear to be a series of monthly updates to the labels.  Like the SPL data, these files are also large.  There is also a huge (6+ GB) aggregate file, but not the historical archives.
  + Drug safety label changes by month are listed by the FDA [here](http://www.fda.gov/safety/medwatch/safetyinformation/safety-relateddruglabelingchanges/default.htm) for the past 7.5 years. This will likely need to be scraped to get it into usable form (e.g. CSV or EXCEL) for data analysis.
* Identifying adverse events on the same drug, and attempt to determine if there is a correlation between the date and aggregate number of adverse events (positive or negative) and the date of the label change. One might expect that the label change has some positive effect (e.g. reduction in adverse events) going forward, but you never know!
* Visualizing the results in aggregate and allowing drill down to specific drugs.  If there is a drug category, this might be used as a filter to select or omit certain categories.
* Filtering based on other key fields like generic/non-generic, drug manufacturer, state/location of the adverse event, type of correlation, magnitude of correlation.
* More complex analysis might involve aggregating a drug with its generic and manufacturer-specific brand variants, assuming that data is readily available.

# Initial Meetings with User Subject Matter Experts (SMEs)

The Product Manager and Front End Web Designer (Data Scientist) then engaged in discovery with a couple of user SMEs. The meetings and related notes are documented below.

## -----Original Appointment-----

From: Robert Damashek

Sent: Thursday, June 18, 2015 9:39 AM

To: Robert Damashek; Chuck Rehberg (chuck\_r@trigent.com); Gail Chen

Subject: Discuss Opportunities to leverage FDA Open Data from an industry perspective

When: Thursday, June 18, 2015 4:00 PM-5:00 PM (UTC-05:00) Eastern Time (US & Canada).

Where: Skype Meeting

Notes:

We met with Chuck Rehberg, CTO of Trigent Software a company that provides analytical support to researchers in the pharmaceutical industry. Gail Chen, our Front End Web Designer (and Data Scientist), had extracted and loaded a couple of initial FDA datasets from the Open FDA site into Tableau involving both adverse drug events and safety product labels. We walked through available fields with Chuck to better explore this data. This FDA open data involves what is known as Phase 4 - approved drugs.

What’s hot in the pharmaceutical research community is both translational medicine (accelerating getting new products into use) and tracing negative impacts. For the latter, this involves tracing a drug that had toxicity in liver back to what happened, and what might have changed from the expected results. This involves mapping the event from bench to bedside and back including patient conditions, severity of reaction, demographics, and outcome. Researchers look at drug components and drug-gene interactions, relate the event to a body part, organ or body system, then what genes are involved, environmental factors, etc.

Given the Open FDA Adverse Drug Event data, it appears that it would help to first focus on visualizing basic correlations with demographics such as location, occupation, age, and ethnicity. After this, understanding drug-drug interactions is a critical and very valuable area of pharmaceutical research. This includes exploring what other drugs were taken at the time of an adverse event including non-prescriptions drugs, like niacin. The datasets seem to have all drugs found present in an event linked via Case ID. If we can intersect adverse events based on common characteristics and then correlate the results with outcomes (which involve a small set of codes), that could be a very promising contribution to a researcher’s scenario. In the data, indications are patient treatments and the adverse event may have been caused by the disease. Common indications found correlated with events could also be very useful, because physicians don't always know other medications people are taking, particularly non-prescription ones, which could have adverse side effects.

We decided that our team would further explore the art of the possible including:

* what’s in the data and what can we group and correlate based on available data
* what about the demographics is in common such as occupation and age

Chuck offered to validate that whatever we come up with makes sense in supporting a pharmaceutical researcher’s scenario.

-----Original Appointment-----  
**From:** Rose Wang   
**Sent:** Thursday, June 18, 2015 4:33 PM  
**To:** Rose Wang; Robert Damashek; Gail Chen  
**Subject:** Call w/ Zhenson on FDA recall dataset  
**When:** Friday, June 19, 2015 8:30 AM-9:00 AM .  
**Where:** 301-529-1828

## Notes:

We met with Zhenson Huang (known as Robert Zhenson), President of Precise Software, whose company supports several areas of the FDA Office of Research Compliance (ORC), as well as the Mission Accomplishment and Regulatory Compliance System (MARCS) database. Zhenson said we need to clearly understand the audience for the data. FDA’s internal systems capture much more data than is made available externally or publicly, so targeting the FDA researcher or regulators may not produce a very realistic or useable scenario. From a public or physician perspective however, it might be useful to explore drug recalls by company and class, if this is accessible via a unique Product identifier. Finding unique and well curated identifiers is a big challenge in the FDA regulatory community. If useful data is available, it could then be further correlated with adverse events through common demographic characteristics.

# Context, Data Analysis, Product Backlog and User Stories

Based on guidance from the user SMEs and actual user community needs supporting research of adverse drug events, the Product Manager refined the Enterprise Context as well as Persona and Scenarios, and collaborated with the Interaction Designer/User Researcher/Usability Tester to create an initial Prioritized Product Backlog of Epics and User Stories. These were used to provide scope to parallel Pool 1 design and Pool 2 development activities.

## Revised Enterprise Context:

* **Revised Capability:** Regulated entities and impacted health care personnel need the ability to research, analyze and better understand potential causes of adverse events by looking at multiple factors present across a set of similar events.
* **Why Important:** To detect issues in drugs on the market, plan changes in drug warning labelling, plan for product recalls for drugs involved in recurring events with serious outcomes, or plan for new drugs to avoid recurring adverse events.
* **Persona:** Pharmaceutical Researcher responsible for assessing the safety of certain Phase 4 drugs.
  + **Scenario:** The number of adverse drug events for one of a Pharmaceutical company’s drugs has raised concerns.
    - **Use Case:** Explore drug safety by analyzing if one or more common demographics (any or all selected) are present across adverse drug events over a selected time period.
    - **Use Case:** Explore drug safety by analyzing potential drug-to-drug interactions involved in related adverse drug event outcomes over a selected time period.
* **Persona:** Physician concerned that drug-drug interactions in prescribed medications may lead to health and safety issues for patients.
  + **Scenario:** During mandatory continuing education, a Physician learns about the danger of certain drug-drug interactions, some of which involve drugs that have been prescribed to his/her own patients.
    - **Use Case:** Explore if conditions involving any of his/her patients match or are correlated with those observed in adverse drug events over a selected time period.

## Data Analysis

The Front End Web Designer (Data Scientist) analyzed the datasets applicable to the identified scenarios. In the process, several gaps and characteristics of the underlying data were identified including:

* Patient occupation data was not available, only Reporter occupation
* Patient ethnicity data was not available
* Patient treatment data was not available
* The Adverse Drug Event table contains 30,000 records just for one quarter of one calendar year
* Adverse Drug Events are identified by a Case Id, and include date occurred and date reported
* The Case Id is a common Foreign Key in all related tables
* The Demographics table contains patient data on patient age, gender, occupation (codes)
* The Outcomes table contains data on patient outcomes (small set of codes)
* The Drugs table contains data on drugs being taken by a patient at the time of the event

## Initial Product Backlog (in priority order) and Preliminary User Stories:

1. **EPIC**: As a Pharmaceutical Researcher I would like the ability to extract one or more common factors that could be contributing to an adverse drug event so that I can better assess the safety of certain Phase 4 drugs.
   1. **User Story:** As a Pharmaceutical Researcher I would like to determine if past adverse drug events can be correlated to a specific period of time within a year, or if they are seasonal correlations such as the high point of the allergy season.
   2. **User Story:** As a Pharmaceutical Researcher I would like to determine if past adverse drug events can be correlated to a specific age range of patients.
   3. **User Story:** As a Pharmaceutical Researcher I would like to determine if past adverse drug events can be correlated to the gender of patients.
   4. **User Story:** As a Pharmaceutical Researcher I would like to determine if past adverse drug events can be correlated to a specific weight range of patients.
   5. **User Story:** As a Pharmaceutical Researcher I would like to determine if past adverse drug events can be correlated to the country where the event occurred.
   6. **User Story:** As a Pharmaceutical Researcher I would like to determine if past adverse drug events can be correlated to the drug(s) patients are receiving.
   7. **User Story:** As a Pharmaceutical Researcher I would like to determine if past adverse drug events can be correlated to documented patient outcome(s).
   8. **User Story:** As a Pharmaceutical Researcher I would like to determine if past adverse drug events can be correlated to the reporter’s occupation.
2. **EPIC** (stretch): As a Pharmaceutical Researcher I would like the ability to determine if drug-to-drug interactions may play a role in certain outcomes so that I can better assess the safety of certain Phase 4 drugs.
   1. **User Story:** As a Pharmaceutical Researcher I would like the ability to determine if drug-to-drug interactions and the resulting patient outcomes can be correlated to the age range of the patients.
   2. **User Story:** As a Pharmaceutical Researcher I would like the ability to determine if drug-to-drug interactions and the resulting patient outcomes can be correlated to the gender of patients.
   3. **User Story:** As a Pharmaceutical Researcher I would like the ability to determine if drug-to-drug interactions and the resulting patient outcomes can be correlated to the weight range of patients.
   4. **User Story:** As a Pharmaceutical Researcher I would like the ability to determine if drug-to-drug interactions and the resulting patient outcomes can be correlated to the country where the event occurred.
   5. **User Story:** As a Pharmaceutical Researcher I would like the ability to determine if drug-to-drug interactions and the resulting patient outcomes can be correlated to a specific period of time within a year, or if they are seasonal correlations, such as the high point of the allergy season.
   6. **User Story:** As a Pharmaceutical Researcher I would like the ability to determine if drug-to-drug interactions and the resulting patient outcomes can be correlated to a designated drug found to be common in a group of drugs patients were receiving.
3. **EPIC** (alternate): As a Physician I would like the ability to explore if symptoms demonstrated by patients match symptoms observed in adverse drug events to better ensure the safety of prescribed medications.
   1. **User Story:** As a Physician I would like the ability to determine if symptoms demonstrated by patients can be correlated to the symptoms documented in adverse drug events based on the age range of the patients.
   2. **User Story:** As a Physician I would like the ability to determine if symptoms demonstrated by patients can be correlated to the symptoms documented in adverse drug events based on the gender of patients.
   3. **User Story:** As a Physician I would like the ability to determine if symptoms demonstrated by patients can be correlated to the symptoms documented in adverse drug events based on the weight range of the patients.
   4. **User Story:** As a Physician I would like the ability to determine if symptoms demonstrated by patients can be correlated to the symptoms documented in adverse drug events based on the country where the event occurred.
   5. **User Story:** As a Physician I would like the ability to determine if symptoms demonstrated by patients can be correlated to the symptoms documented in adverse drug events based on the drugs patient are taking.
   6. **User Story:** As a Physician I would like the ability to determine if symptoms demonstrated by patients can be correlated to the symptoms documented in adverse drug events based on documented patient outcomes.
   7. **User Story:** As a Physician I would like the ability to drill down on any individual adverse drug event so that I can analyze each individual event in greater detail.