Date: February 11, 2013

Description: Brainstorming Additional Problems

Notes:

- Overview of group
  - Materials can be found: [http://isds.wikispaces.com/Technical+Conventions](http://isds.wikispaces.com/Technical+Conventions)

- Use case requirements:
  - An ongoing need for a public health department or other problem owner (involving disease surveillance)
  - Can be stated as a well-defined technical problem
  - Results of a successful method can be recognized, tested in output of proposed methods
  - Requirements of a benchmark dataset understood, and a benchmark dataset may be created with a combination of
    - Authentic data shareable directly or with data use agreements
    - Simulated data based on authentic data characteristics
  - Labeling with effects of interest or simulation of those effects

- Committee roles needed
  - Committee administration
  - Formulation of use cases, rough draft of problem templates (on your own or in consultation)
  - Template refinement calls: problem owners and solution developers
  - Validation of submitted solution methods
  - Communication with stakeholder institutions, individuals external to ISDS
    - Committee purpose and functions
    - Problems, solution methods, publications

- Use problem suggestion:
  - When have special events which lead to population being much different from baseline population (e.g., Republican National Convention, Olympics, etc.), should detection algorithms reflect this? People are concentrated in hotels?
    - Ongoing need (in terms of events)
    - Can be well-defined
    - Need to specify scope
    - Would it be trying to do too much?
    - Important preparedness issue
    - May need to think through more – where the standardized template comes in
    - Delve into business processes, maybe data would not be needed
    - Would it be possible to get real data once this problem is defined?
      - Simulated data may be insufficient
• If cannot share real data, have issue of problem owner not trusting simulated data and solution developer not being able to get access to real data
• Mini dataset to bring to call, so that solution developers could see a preferred dataset format
  ▪ Algorithm development question:
    • Are you looking to reduce number of false positives?
• Access to data
  o To get real data, will likely need MOU – may depend on jurisdiction
  o Traditional software development, usually 2 options
    ▪ Agile solution: Solve one small problem at a time and reiterate and reiterate
  o In this case, agile solution will probably be necessary
    ▪ Start with solving problem of problem owner and then worry about expanding to other situations
• Standardized template is important, but unrealistic to expect problem owners to be able to fill out enough detail for the solution developer
  o Problem clarification may not be possible in a 1-hour call
  o Would possibly need multiple calls
• Scope
  o Each use case might be brought up by a person with a problem that they need to solve and solution developer would work with them to define scope – initiating conversation
  o Could include exploratory data analysis, data quality questions, etc.
• Difference between use case and data requirements
  o Start with use case
  o Once case is better defined, develop set of requirements
• Communicating problems to community
  o Al Ozonoff submitting abstract for APHA
    ▪ Hopefully have one or more use cases to present
  o Howard developing abstract for another conference
  o Define community:
    ▪ International – PAHO and others as potential problem owners, as well as solution developers

Next steps:

• Dave Atrubin (with help of Jeff Johnson and Ian Painter), will complete the problem template and initiate an e-mail conversation to try to define problem better
• Set-up a call to discuss 'Time of arrival problem'