Breakout_A-6 (can map to issue name later)

Issue: https://github.com/CFDSI/Kickoff_Workshop/issues/#
Related Issues:

Issue Statement: What are the greatest obstacles to extracting insight from experimental and computational data bases?
Discussion topic:
Moderator: Scott Dawson
Note taker: Patrick O'Leary
Reporter: Scott Dawson
Group Members:

- Assad Oberai
- Scott Dawson
- Patrick O'Leary
- Anya Jones
- Paul Bauman
- Jean Hertzberg
- Guillermo Araya

Please address these topics in your discussion (moderators please make sure that there is enough time to cover all three before the session ends).

1. Describe the problem:

- What is insight? Is it differentiated by perspective? (Quantitative versus qualitative)
    - Insight: qualitative/visualization vs quantitative/algorithms
    - Basic statistics easy, but higher level insight would be better (is there separation). For example, look for "structures" in the flow. What does that mean? What are the tools to extract that?
    - Can quantitative/algorithmic methods be used to enhance qualitative methods, and make definitions more rigorous?
- What is the difference between experimental and computational data? Are they becoming more similar?
    - Starting to produce TB of experimental data, too big there must be interesting features we can't see. How can you leverage human insight to increasingly large datasets?
    - Need the details from experiments (geometry, boundary conditions, …)
    - Experimental data for inverse problems, what to use to drive this process

    - How can the community obtain a good understanding of the effects of spatial and temporal resolution, noise, other computational/experimental uncertainty on both qualitative and quantitative insight-gathering procedures? How should this be accounted for when archiving/publishing data?
    - How can tools to obtain insight influence decisions related to data acquisition/storage?
- How do we integrate insight tools, physics and analytic techniques into education, without losing training in more established theory?
- Best way to generate data, and best ways to analyze. How can these work together?
- How can you find synergistic collaborations between expertise on data generation with expertise on “extracting insight”?
- How to get funding/support for this synergistic step?
- How can individuals tap into the collective expertise of the diverse community?
- How to train students to be brilliant in physics, not just the requisite tools?

2. What are potential solutions?

- Portal of tools and techniques for the non-expert data scientist to reduce the barrier of entry.
- Renaissance education for CFD including statistics, machine learning, … (project-based)
- Allow for the community to recognize, understand, and work on the issues above?
- Engage more people with expertise in computer sciences.

3. What can CFDSI do to help?

- Provide expertise (CS, stats,) to support project-based education and research approaches, in contrast to fixed curricula and pre-planned issues. How to compensate for time?
- Promote and produce tools that are easy to use for both educational and research purposes, on a broad range of data
- Promote Working groups at meetings
- Promote get-acquainted travel. 10 minute presentations at a conference don’t result in honest exchanges
- Documenting standards for data-exchange/sharing

4. Misc ideas so they don’t get lost (e.g., Did you find new issues? If yes, create the issues on GitHub!):

- How can tools to obtain insight influence decisions related to data

acquisition/storage?
- How can we incorporate new techniques and methods in the community (student coursework and more generally) without losing training in more established theory?

5. Summary for report-back (Alternatively, just bold the key points above):

    -