Conceptualization of an NSF Scientific Software Innovation Institute for High Energy Physics (S2I2-HEP)

Peter Elmer - Princeton University
Mark Neubauer - University of Illinois at Urbana-Champaign
Mike Sokoloff - University of Cincinnati
http://s2i2-hep.org

10 April, 2018
URSSI Kick-off Workshop

S2I2-HEP Conceptualization - http://s2i2-hep.org

The primary goal of the S2I2-HEP conceptualization project is to prepare a strategic plan for a potential NSF Scientific Software Innovation Institute (S2I2) to develop software for experiments taking data in the "High-Luminosity Large Hadron Collider" (HL-LHC) era in the 2020s.

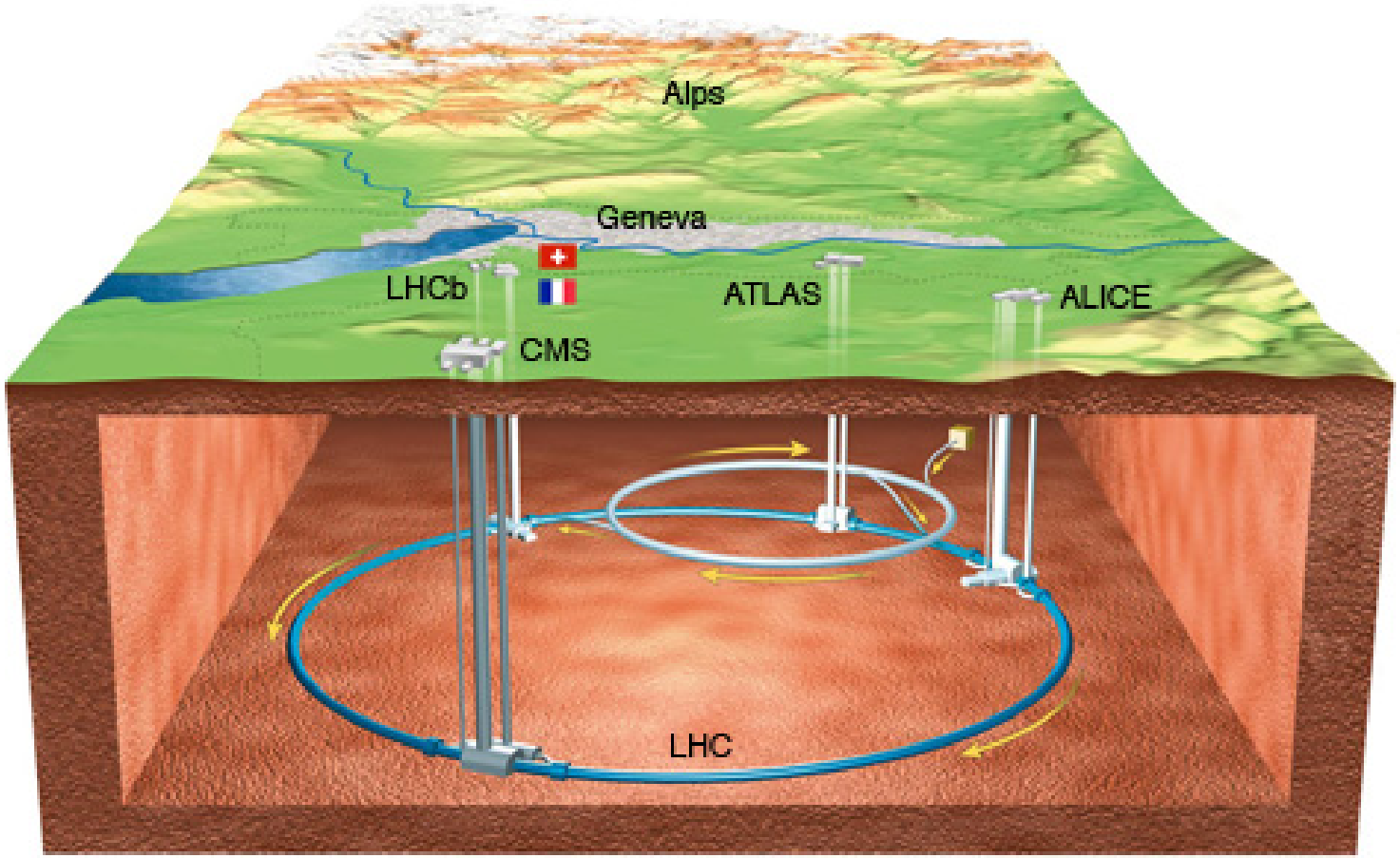
S2I2-HEP sponsors community workshops and conceptual work to "take advantage of the significant data and computing requirements of the Large Hadron Collider as a science driver for next generation high-performance software and sustainability developments."

HEP Science Drivers - Beyond the "Standard Model"

The *Strategic Plan for U.S. Particle Physics* from May, 2014 (the last community "decadal survey" and planning exercise for US HEP) identified "five compelling lines of inquiry that show great promise for discovery over the next 10 to 20 years." These are the Science Drivers:

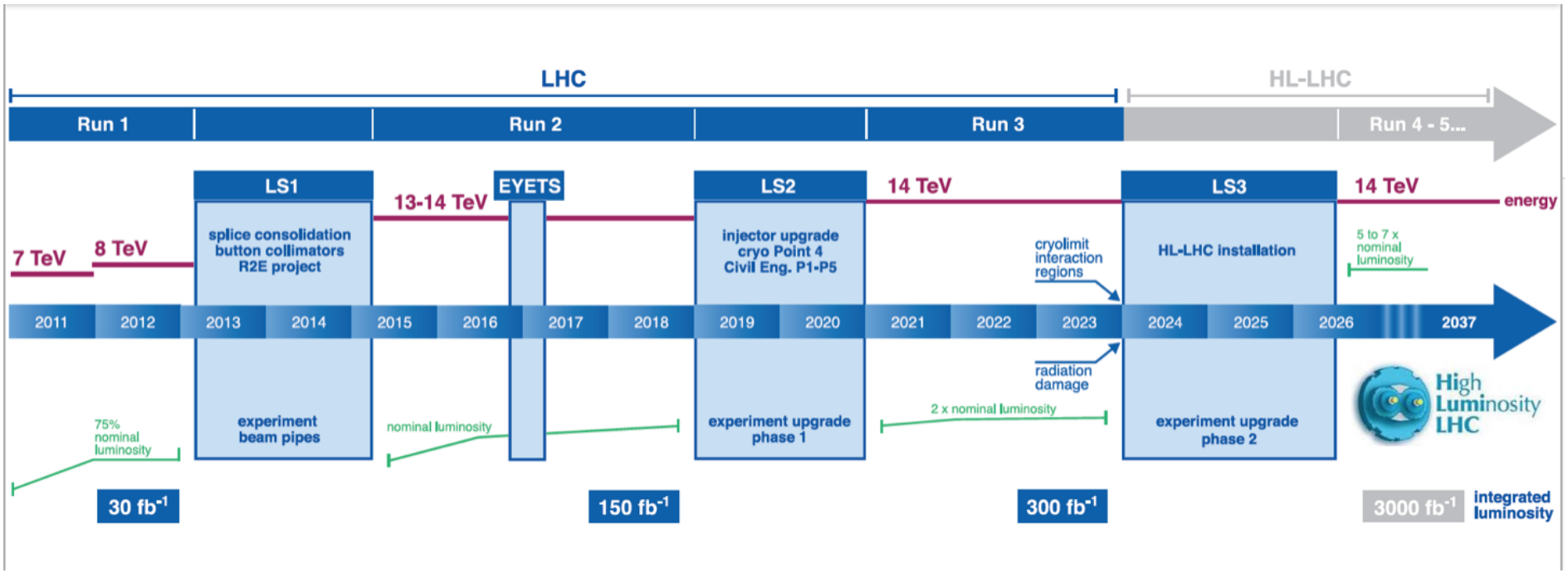
- Use the Higgs boson as a new tool for discovery
- Pursue the physics associated with neutrino mass
- Identify the new physics of dark matter
- Understand cosmic acceleration: dark matter and inflation
- Explore the unknown: new particles, interactions, and physical principles.

Large Hadron Collider (LHC) and Experiments



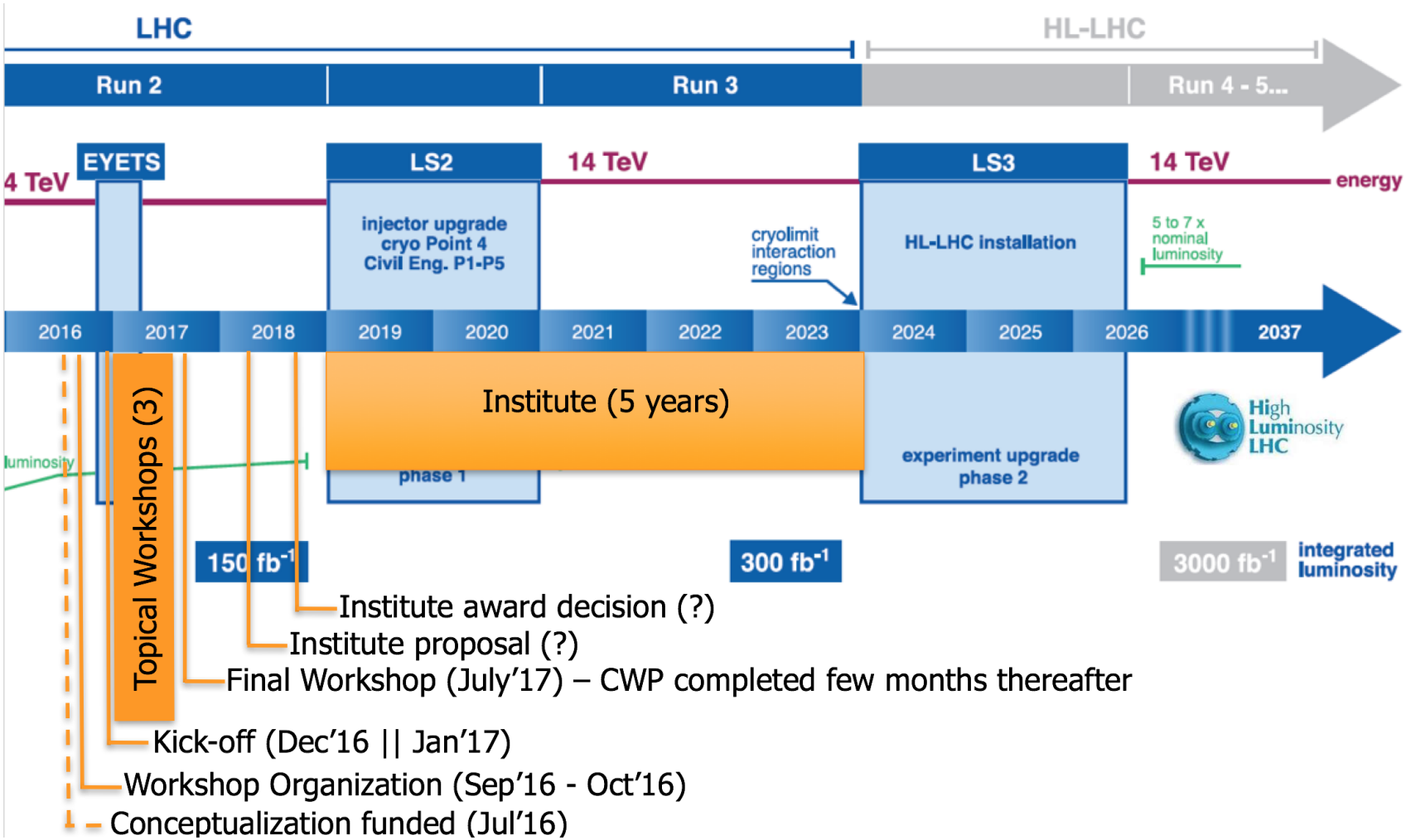
Two very large experiments (Atlas, CMS) with 3500+ people, and two large experiments (Alice, LHCb) with 500+ people

Plans for upgrading the LHC and Experiment Detectors



The High Luminosity LHC Upgrade is a multi-national, multi-agency effort to realize the ultimate physics reach of the LHC. The NSF is participating and preparations are underway for a possible "Major Research Equipment and Facilities Construction" (MREFC) project to begin in ∼2020.

August 2016 vision of Conceptualization Timeline



HEP Software Ecosystem

IgProf

FroNTier

FastJet

RooStats

EvtGen

Examples, definitely incomplete!

Plus 15-20M Source Lines of Code (SLOC) of "experiment specific" codes, as well as dependencies on non-HEP scientific software.

HEP Software Foundation (HSF)

The HSF (http://hepsoftwarefoundation.org) was created in early 2015 as a means for organizing our community to address the software challenges of future projects such as the HL-HLC. The HSF has the following objectives:

- Catalyze new common projects
- Promote commonality and collaboration in new developments to make the most of limited resources
- Provide a framework for attracting effort and support to S&C common projects (new resources!)
- Provide a structure to set priorities and goals for the work

(Switch to Webpages)

S2I2-HEP (workshops) and papers: `http://s2i2-hep.org`

HSF CWP and WG:
`http://hepsoftwarefoundation.org/activities/cwp.html`

HSF WG Start: `http:`
`//hepsoftwarefoundation.org/cwp/cwp-wg-guidance-sdsc.html`

Impact Criteria (for evaluating each WG topic)

1. **Impact - Physics:** Will efforts in this area enable new approaches to computing and software that maximize, and could potentially radically extend, the physics reach of the detectors?
2. **Impact - Resources:** Will efforts in this area achieve required improvements in software efficiency, scalability and performance and make use of the advances in CPU, storage and network technologies?
3. **Impact - Sustainability:** Will efforts in this area guarantee the long term sustainability of the software through the lifetime of the HL-LHC?

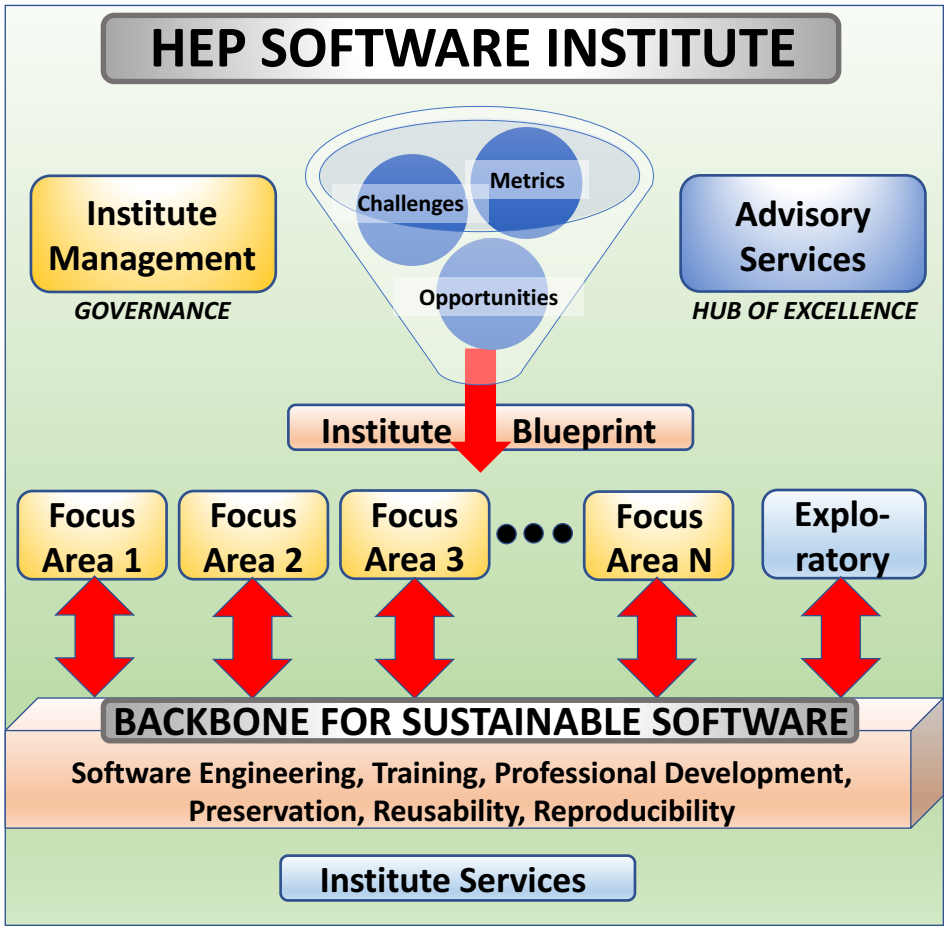
On Impact

"the mighty ships tore across the empty wastes of space and finally dived screaming on to the first planet they came across - which happened to be the Earth - where due to a terrible miscalculation of scale the entire battle fleet was accidentally swallowed by a small dog." - Douglas Adams, The Hitchhiker's Guide to the Galaxy

Additional Criteria (for prioritization)

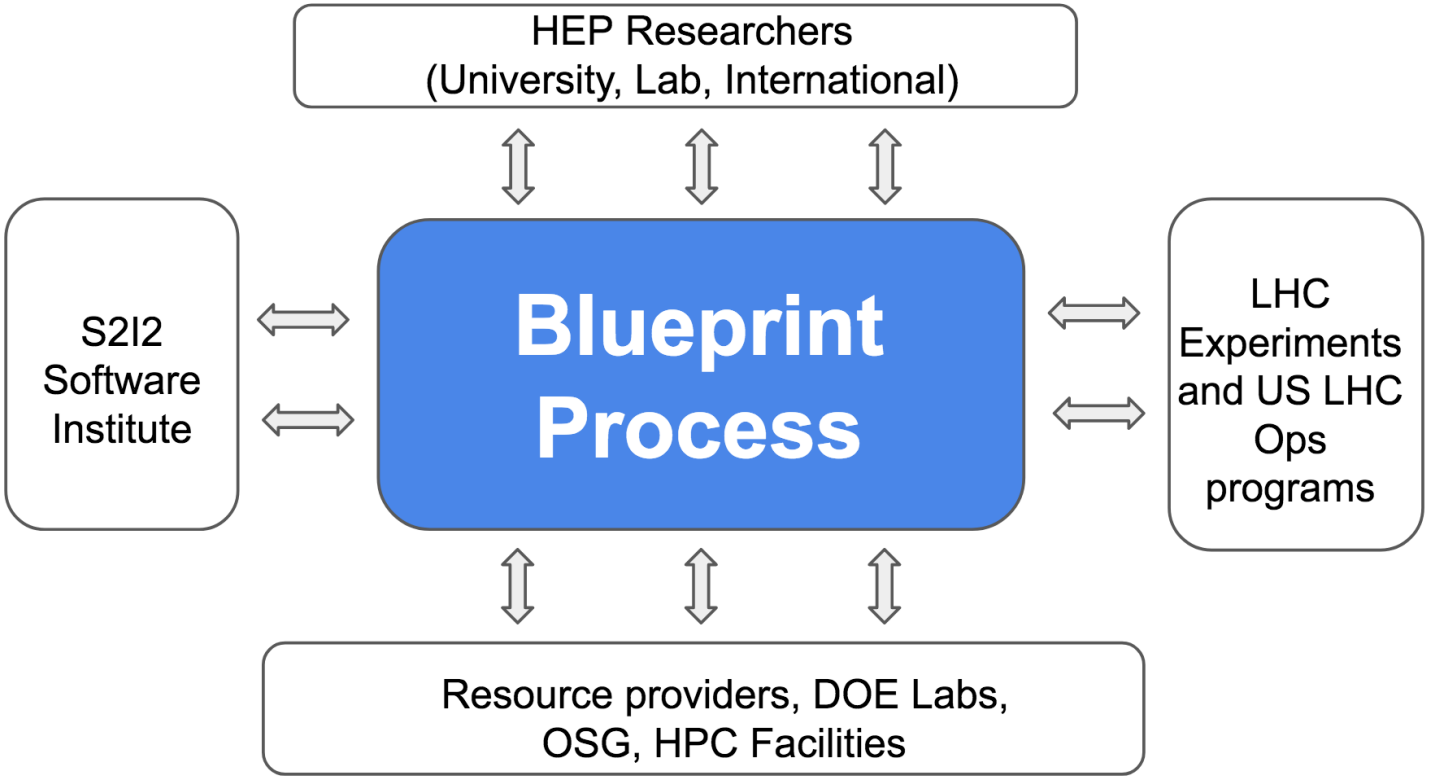
1. **Interest/Expertise:** Does the U.S. university community have a strong interest and expertise in the area?
2. **Leadership:** Are the proposed focus areas complementary to efforts funded by the US-LHC Ops programs, DOE or international entities?
3. **Value:** Is there potential to provide value to more than one LHC experiment and to the wider HEP community?
4. **Research/Innovation:** Are there opportunities for combining research and innovation as part of partnerships between the HEP and Computer Science communities?

Institute Structure



4 focus areas (of 10+ Working Groups) were prioritized. (See the Strategic Plan if you care about the details.)

Blueprint Process



The Blueprint Process will be a primary means of developing a common vision with the major partners. Blueprint activities will likely happen 3-4 times per year, typically with a focus on a different specific topics each time.

Extra thoughts

I made some set of (cryptic) notes as to what we did during our conceptualization that helped us. Many already seem integrated into this project, but several that aren't (yet) visible (food for thought):

- Design "closed loop" processes for meetings and workshops. Capture the vision. Publicize questions to answer from a meeting in advance and get those answers. Outcome-oriented. Designate people to drive/herd meetings, and to own the process of culling out the answers and agreements/disagreements. Iterate.
- Clearly separate "What" from "How" in discussions. Agreeing on what a community wants to do (science driven) is as important as how it does it. The "Vision" is not the "Work Plan". (Blueprint)
- Ask questions that encourage more than incrementalism. What if we actually had extra resources?