"The Subterranean Macroscope: Sensor Networks for Understanding, Modeling, and Managing Soil Processes."

Table of Contents

1. Executive Summary

We have a poor understanding of the physical, chemical, biological transformations and the cycling of soil in the vadose zone, as well as its influence on plant science and food security. We lack adequate soil information at high spatial and temporal length scales, and consequently, our current models are often inadequate. However, revolutionary advances in sensors and nanotechnology, sensor networks, communications and microelectronics technologies, and data analytics are poised to enable scalable and affordable subterranean sensing networks that may revolutionize soil science and plant science itself. This can have a major impact on the environment, food security, and its management. This workshop was dedicated to identifying the key grand challenges in soil and plant science, how they could benefit from a high resolution subterranean sensor network that measured soil parameters at high spatio-temporal resolution, and the underlying engineering and science challenges that need to be tackled in order to create such a subterranean network.

The three grand challenges identified at the workshop were:

A. Understand how the soil microbiome affects plant productivity, water and nutrient efficiency, and soil degradation
B. Create a new generation of accurate terrestrial ecosystem models: build C, N and nutrient cycling models that offer predictive accuracy
C. Understand the root interface between the soil and the plant for efficient use of water and nutrients (Food security).

Based on discussions at the workshop, there is a need for an ambitious program to establish Earth Macroscopes: buried sensor networks that collect high resolution data that are then coupled to, and inform, an intense effort at furthering our knowledge in the three grand challenges described above. Additionally, there is a need for model development at the soil, plant and the soil-plant interface. This model would need to be coupled to experiments and supported by pilot testbeds that will bring convergence between soil scientists, plant scientists, microbiologists, computer scientists, nanotechnologists, and engineers.

Success in developing this vision calls for innovation and discovery science in the engineering and computer science fields, with a focus on four key areas: (i) Sensors and sensor materials for accurate, low energy high spatial and temporal resolution subterranean sensing. This includes innovations in chemical functionalization strategies for the development of cheap and compact optical/electrical platforms, new techniques for root mass and root exudate imaging, compact chip scale polymerase chain reaction (PCR) tools, lysing approaches that are rugged, and field deployable, proxy sensing and analytic inference engines; (ii) Ultra low power draw (< 10nW average power consumption) microelectronic hardware for sensor management, at-node information processing, memory management and communications; (iii) Wireless technology research that leads to robust fully buried low power sensing network with high range and throughput, including research in both the wireless protocol domain and device level implementation; (iv) Computer science research that leads to new methods and algorithms for sensor data model integration. Large-scale aggregation and curation of data across the community and the creation of geographical testbeds with multidisciplinary expertise to enable coordinated R&D on sensors, data solutions, and models.

The workshop discussion was framed by three key topics that build upon NSF's "Big Ideas": (i) Convergence, bringing together the multidisciplinary communities noted above; (ii) Big Data, the efficient transmission, curation, and analysis of dense, *in situ* soil data over time; and (iii) Predicting Phenotype from Genotypes in diverse environments through advances in measuring soil environmental conditions that interact with genotype (G x E interactions).

The workshop was held in downtown Chicago, IL over a two-day period (November 1-2, 2017). There were 68 attendees spanning expertise in the soil and plant sciences, microbiology, genotype/phenotype modeling, nanotechnologists and sensor experts, computer scientists and Internet-of-Things experts, wireless technology researchers and microelectronics experts. Participants came from academia, industry, national laboratories, and a number of government funding agencies.

2. Motivation and Outcomes of Workshop

2.1 Motivation

Science has a poor understanding of one of the most important components of life on earth—the nature of the soil in the vadose zone, the subsurface region of the earth that typically reaches to 0.6 to 3 meters below the surface. Soils provide food, fiber, and fresh water, make major contributions to energy and climate resilience, and help maintain biodiversity and the overall protection of ecosystem goods and services. Despite the importance of soil, we do not fully understand the physical, chemical, and biological transformations and cycling in soils at temporal and spatial scales that are needed to model, visualize, and manage. Most soil models do not incorporate many of the biogeochemical processes needed for prediction and decision support for soil management. This, in turn, has a major impact on our actions to achieve food security and improve soil health and the environment. Our current models are inadequate because, until recently, it has been difficult to gather data at high spatial and temporal resolutions. However, the emergence of new sensing technologies, sensor networks, and data analytics -- driven by breakthroughs in the nanotechnology, wireless technologies, microelectronics, and computing fields, have opened up opportunities for greater monitoring of subterranean properties in a geographically scalable and affordable fashion. Soil science has also progressed to a point where it can begin to leverage this data and bring us to the next level of discovery science. For example, breakthroughs in soil metagenomics have allowed a greater understanding of the diversity and controls on soil processes. However, there are still many remaining gaps in our knowledge. For example, if we could develop a way to map both beneficial and deleterious soil microbial communities, it would revolutionize our understanding of soil biology, what is happening in the soil, and our ability to manage these elements.

2.2 Outcomes

Building on these recent successes across broad science and engineering fields***, this workshop brought together researchers in soil science, (including experts in the biological, chemical and physical nature of soil), dynamic soil modeling expertise, plant sciences with experts in sensor networks, microelectronics and wireless researchers, and machine learning/data analytics to develop a vision fo*****r**:

A. A wide-scale, high-resolution subterranean sensor network with the ability to accurately sense relevant biological, physical, and chemical soil parameters. This vision will also include the development of new sensors and techniques for measuring parameters of importance to the soil science and plant science communities.
B. A trajectory for curating, analyzing, and using the resulting data to develop the next generation of models and management strategies for soil improvement, agricultural intensification, and water conservation.

One of the major outcomes of this workshop is to develop an informed position on the current status and opportunities in sensor technology, an identification of the synergies among the multidisciplinary stakeholders, and an outline of a 10-year research agenda for an intelligent, geographically and temporally scaled subterranean macroscope with the following components: (i) A subterranean sensing network, including the required sensors, the communications technologies, and the low power microelectronics technologies innovations; and (ii) Data analytics, the associated data and management requirements, and the accurate soil and plant modeling systems that will use the data.

Additionally, the workshop enabled us to explore ***three, crosscutting topics that build upon NSF's "Big Ideas"***: (i) Convergence, bringing together the multidisciplinary communities of nanotechnologists, electrical engineers, data scientists, soil scientists, microbiologists, and soil process modeling; (ii) Big Data, the efficient transmission, curation, and analysis of dense, *in situ* soil data over time; and (iii) Predicting Phenotype from Genotypes in diverse environments through advances in measuring soil environmental conditions that interact with genotype (G x E interactions).

The workshop also provided insight and a framework for approaching new questions, informed by the participation of experts from soil science, engineering and biology to computer science and industry.

Intellectual Merit: One of the major goals of this workshop was to begin to identify successful shared approaches for the use of spatial, temporal, and multi-modal data, which can improve our understanding of soil dynamics while improving the accuracy of soil and plant models.

Broader Impacts*:* The workshop and the report produced by the workshop will have a broad impact and help to initiative closer collaborations between soil and plant scientists and the engineering/nanotech/computer science community. Moreover, a greater understanding of soil and plant models will also have a profound impact on topics within and beyond academia including agricultural yields, climate models, water and agricultural management, global food security and new sensing materials development.

3. Workshop Organization

The workshop was held in downtown Chicago, IL over a two-day period (November 1-2, 2017). The workshop was co-chaired by Dr. Supratik Guha (University of Chicago) and Dr. Charles W. Rice (Kansas State University). Additionally, a scientific advisory committee consisting of key experts across the fields of plant and soil science, computer science, and wireless assisted in the selection of participants and informed the topics as well as the intellectual direction of the workshop (see Table 1 below).

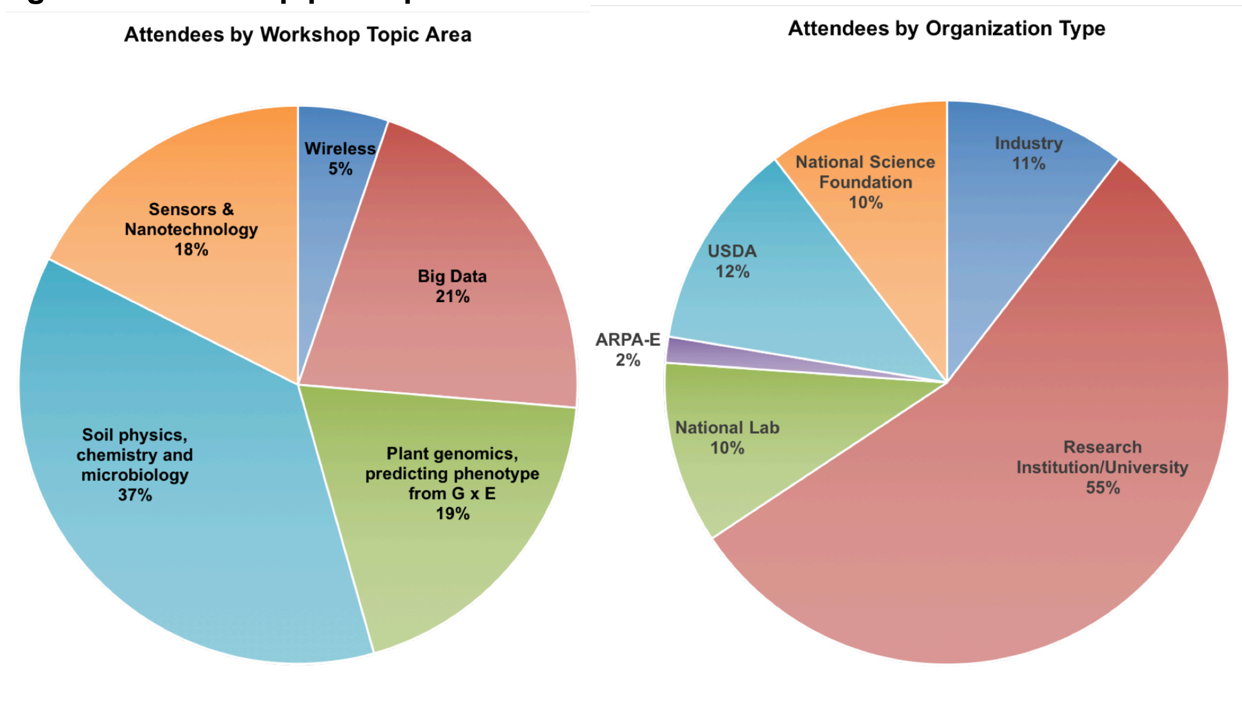
Table 1: Scientific Advisory Committee Members	
Steven R. Evett, Acting Deputy Administrator Natural Resources and Sustainable Agricultural Systems USDA | Agricultural Research Service Ian Foster Arthur Holly Compton Distinguished Service Professor, Department of Computer Science, University of Chicago Distinguished Fellow, MCS Division Senior Scientist, MCS Division, Argonne National Laboratory	Roberto Cesar Izaurralde Professor, Department of Geographical Sciences University of Maryland Ali Mohamed Division Director of Environmental Systems Institute of Bioenergy, Climate, and Environment USDA | National Institute of Food and Agriculture

Monisha Ghosh Research Professor, Institute for Molecular Engineering Associate Member, Department of Computer Science University of Chicago Affiliate, Argonne National Laboratory Supratik Guha (co-chair) Professor, Institute for Molecular Engineering, University of Chicago Division Director, Nanoscience and Technology Argonne National Laboratory	Charles (Chuck) W. Rice (co-chair) University Distinguished Professor Mary L. Vanier University Professorship Chair, Board on Agriculture and Natural Resources, National Academies of Science, Engineering, and Medicine Kansas State University

3.1 Participants

A central objective of the workshop was to not only involve academic experts but also key stakeholders from industry, the national laboratories, and federal agencies who will be collaborators and eventually the end-users of this technology. Workshop participants were selected using nominations from professional societies and the scientific advisory committee members. All proposed participants were discussed and evaluated by the committee.

Figure 1: Workshop participants



Final participant selections were made and approved by the committee to meet the goals of the workshop and to ensure the diverse workshop population (gender, ethnicity, organization type). Overall, there was considerable enthusiasm for the workshop, as evident by the number of attendees and range of organizations represented. There were 68 attendees in total, which exceeded our original estimates of 40-60 participants. Participants included leading practitioners in the fields of soil physics, chemistry and microbiology, biophysical modeling, sensors and nanotechnology, microelectronics, wireless communications, and data analytics and management. Participants were drawn from a diverse group of experts from academia, the national laboratories, federal agencies, and industry. Names and affiliations of participants are provided in Appendix A.

There were 28 distinct academic or research institutions represented, 3 national laboratories, 7 corporations, and 3 federal agencies (USDA ARS/NIFA/NRCS, NSF, ARPA-E). Of the total participants, 14 were women and 54 were men.

An additional goal of the workshop was to include junior researchers such as post-doctoral scientists and students. Student participants not only benefited from the workshop discussions and exposure to diverse organizations but also gained interdisciplinary experience during the final report writing sessions. Seven post-doctoral scientists and students (undergraduate, graduate level) were selected by the scientific committee members to participate and assist the committee member in the workshop in a technical role. These students came from universities across the country including University of Texas at El Paso, the University of Maryland, and Kansas State University.

3.2 Workshop Themes

The workshop revolved around the following themes and their interplay:

- Theme 1: Soil physics, chemistry, and microbiology
- Theme 2: Plant genomics, predicting phenotype from G x E
- Theme 3: Sensors and Subsystems, including wireless technologies and microelectronics hardware.
- Theme 4: Big Data

Given the multi-disciplinary nature of the workshop, a key goal was to develop a workshop format that encouraged intense interaction and group discussion (see full agenda in section 6). As a result, after an opening plenary session for setting the tone, four-panel discussions were conducted on day one with short presentations by a number of specialists. On day 2, following a plenary presentation, group discussions on the panel topics were conducted, in order to arrive at a consensus on the path forward.

4. Technical Outcome of Individual Panel Discussions and Presentations

In the following section, we describe the outcome and conclusions arising from the presentations, discussion and participation in the four panels that discuss the themes of soil science, plant science, hardware technologies, and data analytics (See Figure 3: Workshop Agenda). Below, we first describe the major grand challenges and opportunities in the soil sciences and the plant sciences that would be greatly enhanced by the development of a new generation of sensors and sensor networks that could provide high resolution spatio-temporal data over a wide range of measurement vectors. Then, we turn our attention to the hardware subsystems research required. We describe the research required in sensors, in wireless technology, and microelectronics hardware engineering research required that would be unique in the context of an "earth macroscope" and would, in addition, lead to significant impact in other fields as well. Following this, we describe the research required and opportunities in Big Data that would be essential for use and analysis. This includes a description, specifically, of testbed opportunities, multi-disciplinary and integrated sites where the scientific expertise is collectively assembled and deployed. This is followed, in Section 5, by a synopsis of key outcomes and recommendations integrated across these four discussion threads. Additionally, in Section 5 we also specifically discuss the role of significant elements of engineering which will play an influential role in future research agendas.

As discussed in earlier sections of this report, there are three grand challenges in the soil and plant sciences area where our understanding is currently limited and severely hampered by the lack high quality, high resolution data, and which would greatly benefit from the development of buried underground sensing systems for soil:

- Understand how the soil microbiome affects plant productivity, water and nutrient efficiency, and soil degradation
- Create a new generation of accurate terrestrial ecosystem models: build C, N and nutrient cycling models that offer predictive accuracy
- Understand the root interface between the soil and the plant (Food security).

4.1 Soil Science and Soil Science Sensor Needs

Important areas in soil science that would be furthered by a new generation of high resolution spatio-temporal data include:

4.1.1 A new class of accurate algorithms to model integrated physical, chemical, and biological processes in soil systems using layer data collected at the appropriate resolution.

Currently, there is a substantial number of mathematical models describing soil processes involving the physical, chemical, and biological domains. These models exist as sole soil process models (e.g. soil organic matter model) or as components of larger ecosystem models (e.g. agroecosystem, forest, grassland models). Depending on the scientific objectives, the development of soil process models has shown various degrees of integration among these domains through the utilization of empirical and process-based algorithms. Laboratory and field experiments have been essential to develop, calibrate, and validate these models at various spatial (e.g. sub-meter to watershed) and temporal scales (e.g. seconds to years). However, experiments are usually incomplete with respect to number and resolution of monitored variables (e.g. transformation and the fate of soil carbon and nitrogen [especially with soil depth], soil water and temperature, soil biology). Workshop participants discussed all issues and identified the need and opportunity to develop accurate algorithms to describe and quantify in an integrated manner physical, chemical, and biological soil processes.

4.1.2 Characterize soil microbiome capacity and expression; community diversity and function. Sensors at rhizosphere scale (high spatiotemporal resolution).

The term soil microbiome is used to describe the collective genome of soil microorganisms (such as archaea, bacteria, viruses, and fungi) living in the soil. Workshop participants discussed the need to characterize the soil microbiome in terms of its capacity and expression using marker gene, genomic and metagenomic analyses applied across space and time scales[1]. This includes the consideration that soil comprises not a single but a wide range of environments such as rhizospheres, soil aggregates (surfaces and pores), and environmental variations that occur at the surface and at depth. Characterizing the capacity and expression of soil microbiomes will require the miniaturization of sensors to monitor and sample environments such as rhizospheres and soil pores that enable genomic and metagenomic analyses as well as deduct direct or indirect influences on plants (e.g. plant health) and soil processes (e.g. biogeochemical cycling).

4.1.3 Develop a basic understanding of soil biodiversity. What species are in the soil and what are their functions?

Soil biodiversity reflects the complexity of organisms living in the soil. The diversity of these organisms are extremely large in terms of taxa (e.g. insects, nematodes, fungi, bacteria, archaea), size (e.g. micron to sub-meter sizes), and function (e.g. heterotrophs, autotrophs, regulation or production of trace gases, influence soil carbon dynamics, nutrient cycling). Workshop participants discussed soil biodiversity and recognized knowledge gaps in the understanding of biodiversity across soil environments and the need to document. A recent example was provided by[2], who analyzed topsoil from 237 locations across six continents and found that almost half of the bacterial communities worldwide could be accounted by only 2% (~500) of bacterial phylotypes. The authors concluded that the bacterial taxa thus identified could be prioritized for genomics characterization to improve understanding of soil microbes and their role in ecosystem functioning.

4.1.4 Understand soil-microbe-root interactions and how they relate to plant production and environmental outcomes.

Workshop discussants also identified the need to improve our understanding of soil-microbe-root interactions and how these related to plant production and environmental outcomes. The rhizosphere is a subterranean volume where plant roots exudate organic compounds, mucilage, and dead cells thus influencing the surrounding soil and microorganisms. In turn, the surrounding organisms influence plant growth through mutualistic, parasitic, or symbiotic associations.

Key sensor needs: CO_2, O_2, N_2O, NO_3, P, H_2O, pH, temperature. Measurement of total soil C and fractionated soil carbon forms.

4.1.5 Integrated understanding of co-evolution of human-soil communities. Attempt to model this understanding and attempt to anticipate how these couplings may change in the future.

There is a need for a holistic understanding of the natural system (e.g. climate, soil, plant, humans, machinery, policy). Coupled Natural and Human Systems requires an understanding of the complex interactions between natural systems and humans. Data layers are needed at the appropriate spatial and temporal resolutions for improved understanding for human decision making in managed ecosystems. The interdependence of food, energy, and water (FEW) systems is challenged by developing ground-based monitoring and modeling at local-to-regional scales for decision making and policy[3] Sensor technology will enable improved decision making in FEW systems.

---

[1] Fierer, 2017, Nature Rev. Microbiol. 15:579-590

[2] Delgado-Baquerizo et al. 2018, Science 320-325

[3] Scanlon et al. 2017 Science 53:3550-3556.

4.1.6 Soil environment is heterogeneous and varies both temporally and spatially (horizontal and vertical) on a scale of cm to meters. Scales will depend upon the subject of the question and target: roots, microbes, or soil chemistry.

Workshop discussants recognized challenges in measuring/monitoring soil physical, chemical, and biological variables at appropriate spatial and temporal scales. The selection of methods and sensors should be driven by research questions that concurrently allow for model enhancement and knowledge integration. Soils are among the most complex components of the biosphere where properties vary and process occur within micro to macro spatial scales and within periods from seconds to centuries[4]. Soil microbes regulate important biogeochemical processes; this regulation occurs in small volumes of soil called hotspots. There is a strong need to sense and visualize the location, size, and evolution of these hotspots in relation to bulk soil. Large differences in process rates and microbial activities are expected to occur whether the hotspots occur in the rhizosphere or the detritus sphere[5].

4.2 Plant Sciences: Specific Challenges

Subterranean science challenges in the plant sciences are inextricably linked to soil and microbiological sciences as well as to the above ground plant. Plant science challenges that are unique to the rhizosphere have relevance to interactions with the soil, the soil solution, and the soil microbiome. Current plant breeding efforts to develop drought tolerant varieties are often focused largely on the development of root architecture and how, when and where root length density increases with plant growth and in response to soil water, soil strength, nutrient and microbial characteristics. Thus, sensing needs include not only root hair growth and architectural development, but also the relevant soil, water and microbial state variables that mitigate root development. Tolerance to disease and pests also involves root characteristics for which there are currently no in situ sensing capabilities. The sites for relevant sensing are both small, nearly microscopic, and spatiotemporally dynamic because root growth at the root hair growing point and nutrient and water uptake and penetration sites for pests (microbial, viral and fungal) are also predominantly in the root hairs and younger roots not yet suberized. These characteristics pose great challenges for sensing systems.

Some relevant science questions include:

- How do genomic differences in root characters (architecture, root length density variations in time and space, root hair function) influence water and nutrient uptake and use efficiencies?
- How do crop, soil, water and nutrient management affect root characters?
- How much carbon is exuded by roots, what is the chemistry of root exudates, and how do root exudates activate genes in the rhizosphere?
- How does the microbiome interact with plant roots and what are the other signals between the plant and the microbiome?
- How does root density and architecture develop and how can it be sensed in situ and over time?
- In genomics, how can nanosensors (chemical, biological, physical) be integrated within the plant architecture itself (both above ground and below ground)?

---

[4] Kuzyakov and Blagodatskaya 2015 Soil Biol. Biochem. 83:184-199

[5] Kuzyakov and Blagodatskaya 2015.

- Roots are plastic and react to the environment. How do we factor in soil properties to explain root morphology?
- More capable models of plant and root growth and 3D architecture in response to stresses are needed both as formulations of science questions and as means to investigate potential responses to environmental challenges (temperature, moisture, salinity, soil mechanical, faunal, microbial and nutrient variations in response to climate and management changes).
- How can simulation modeling be integrated with big data infrastructure, data fusion and data assimilation from a subterranean sensing system to provide near real time analytics and decision support?

Key parameters to sense include: CO_2, N_2O, NO_3, P, H_2O, different soil carbon forms, soil strength, root turnover (biomass), root architecture, root surface area, root density, exudate quantity and chemistry, microbial composition (genetics) and population. Some larger scale and intrusive sensors are available, such as those for sensing soil water content, bulk electrical conductivity, temperature and soil solution; but small, minimally intrusive sensors are needed for these to match the scale of root activity without overly disturbing the rhizosphere environment.

***4.2.1 Plant sensor specific needs*:** Plants respond to above ground and subterranean environments, which are linked by the plants themselves and by the water, nutrient and energy balances. Needs for sensors are driven by weather extremes and the need to sustainably intensify productivity by increasing water and nutrient efficiencies. A primary need is for noninvasive comprehensive sensor systems for root growth, distribution, and activity as it relates to soil processes and plant productivity, efficiency and resilience. For such a sensing system to be truly useful for understanding root development in tandem with nutrient and water uptake by the plant, sensing systems for soil water, physical and nutrient states are needed at the same spatial and temporal scales. Plants interact with the soil microbiome in the rhizosphere through poorly understood chemical signaling processes and through their uptake and nutrients and water as well as through root exudates. The latter contain chemicals with many different biological activities, including microbial and viral suppression and inimical effects on other plant species. Sensing root exudate chemistry may be key to understanding much of rhizosphere dynamics and interactions with the microbiome, including pest and disease interactions.

Spatial scale is an important characteristic of sensing systems not to be overlooked. While a small scale may be needed to determine and understand mechanistically the relevant biogeochemical processes, whole plant and field scale sensing systems are likely more relevant for managing water and nutrients. Needed temporal scales range from the seconds to daily for development of mechanistic understanding to weekly or longer for slowly developing environmental and plant processes.

***4.2.2 New and Future Technologies*:** The use of plants as sensors already has a strong basis in science since the visible plant is integrative of root-soil processes. Aerial imagery, multispectral and multi-band reflectance sensors (some active) and emission sensors (infrared and fluorescence) are all in use, both as spot and imaging systems. Progress in this realm is rapid and increasing as computer vision concepts and systems, rapid prototyping, and lower cost and more capable sensors are quickly integrated by a large community of scientists and engineers committed to open source hardware and software development. The rapid expansion of high throughput genotyping facilities at several locations in the US and Europe is indicative of the success of these sensing systems in developing understanding and guiding genomic practice, and the equally rapidly expanding commercial offerings for on-farm advisory services is a testament to their utility in agronomic practice.

In the subterranean realm, there are clear needs for research in both the wireless protocol domain and device level implementation, as well as in the identification of useful proxies. New approaches may include bioengineering to develop “indicator” plants by including substances such as green fluorescent proteins and other signals and biomarkers in plants to use as sensors. Potential sensors systems involved microelectromechanical systems (MEMS) in the 0.001 to 0.1 mm size range. These include increasingly compact optical/electrical platforms. MEMS manufacturing is already well established in the industry but has been little applied to the plant and soil sciences. MEMS approaches might be used to develop new imaging technologies for studying dynamics of roots and root exudates. MEMS technology can include microfluidics and could enable sampling of soil solution and identification of its constituents, possibly including metabolites, biomarkers, and specific ions. The inclusion of a sufficiently scaled down PCR system combined with an instant lysing subsystem could allow for microbial identification or identification of disease and pest organisms. Nascent efforts in molecular nanotechnology offer ways to get at surface chemistry if robust sensors can be fabricated and communications pathways established. Another possible avenue entails the use of “proxy sensors” and analytical inference engines for difficult-to-measure quantities using a suite of sensors e.g., to infer root morphology instead of imaging it.

4.3 Sensing Science and Sensor Technology Needs

Based upon the summaries of sections 4.1 and 4.2 of this report, the overall sensor requirements related to obtaining data, are needed for:

A. Improved understanding of subsurface processes (e.g. carbon transformations, physical and chemical controls, turnover rates);
B. Sensors for assessing soil microbial processes as it relates to biogeochemical cycling and plant efficiency;
C. Sensing and imaging roots and its associated microbiome, exudates, nutrients, and water;
D. Sensing plants growth-related parameters to sustainably intensify productivity by increaser water and nutrient efficiencies;
E. Noninvasive comprehensive sensor system for root growth, distribution, and activity as it relates to soil processes and plant productivity; and,
F. Understanding the soil-root-plant interaction.

Examples of key specific metrics that were determined by workshop participants to be critical, along with resolution parameters were:

- Monitor soil gases (CO_2, O_2, N_2O, NH_3), ion transport (NO_3^-) and energy exchange
    - Key nutrients: NO_3^-, P, SO_4^{-2} in soil pore water
    - Root exudates (sugars, proteins, peptides)
    - Water activity, soil moisture (average over ~ 1m rather than point-based measurements)
    - Temperature
    - Microbial biomarkers
- Root imaging, root volume, imaging modalities in heterogeneous media (like soil)

- Identification of microbial species in soil
- Spatial Scale
- Mechanistic scale to determine biogeochemical processes (cm)
- Field-scale for environmental and plant efficiency (m to 100 m^2
- Temporal scale
    - Sub-daily for mechanistic understanding
    - Weekly for environmental and plant processes
- Major considerations for utility
    - 5-10 year in-ground lifetime
    - Benchmarking of power consumption needed
    - Compact, portable units
    - Inorganic, organic, biological analytes are increasingly unstable, in that order.

With these requirements, there are three broad areas of sensor science and sensor technology that need to be accelerated. These are areas that will impact the field of sensors in general and will have broad implications beyond soil sensing. They can be classified as:

A. Strategies for chemical functionalization of surfaces for robust, reproducible selectivity and specificity. These approaches involve surface chemistry, biochemistry, molecular design, fluorophore chemistry, and nanotechnology research.
B. Compact optical/electrical platforms: (i) with the decreased cost /increased functionality promise of semiconductor-based light sources (lasers and LEDs) and detectors from the ultraviolet to infra-red; (2) and highly sensitive ion-sensitive field effect transistor such as FINFETs. It is now beginning to be possible to put together highly sensitive and sophisticated electrical and optical detector architectures for compact, cheap sensing.
C. Increased use of algorithms and software for data being gathered. For instance, “proxy sensors” that infer a model for difficult to measure quantities using multidimensional correlation from a suite of sensors. Examples, include inferring root morphology instead of imaging or inferring bacteria concentrations from measuring a host of immediately measurable parameters. Proxy sensing will require the use of machine learning and data analytics techniques combined with multi-modal measurement techniques.
D. New imaging technologies for root and root discharge was identified as an important area that required out-of-the-box approaches. Imaging of root biomass through soil conveniently using electromagnetic radiation is a difficult problem, alternatives in using distributed sensing around roots to map root exudate concentrations is an area of emerging interest.

There were also two new areas discussed that could be envisioned as grand challenges for sub-terrestrial sensing and which could have a remarkable impact upon the field:

A. Compact, chip scale polymerase chain reaction (PCR) with stable reagents for in-field soil microbial sensing. PCR machines have been scaling at a pace faster than Moore’s Law scaling for silicon microelectronics. Today PCR machines can be purchased for ~$1K. However, reagents are expensive, have a low shelf life in the field, and easy to use inline technology for lysing remain as barriers. Development of small, field deployable PCR machines that can carry out in-soil measurements will revolutionize work towards relating the soil microbiome to the evolution of soil and plants.

B. “Plants as sensors”: Genetically engineering plants to themselves act as sensors via optical (such as fluorescent protein) and other signaling mechanisms.

4.4 Wireless Technology: Challenges and Research Needs

4.4.1 Specific challenges**:** Wireless sensors that can be completely buried in a spatially dense deployment offer the promise of being able to advance subterranean science by providing unforeseen ability to continually and unobtrusively sense various parameters of interest such as moisture, pH, nutrient concentration etc. However, there are a number of significant challenges that need to be overcome in achieving reliable wireless transmission from underground sensors. Most existing subterranean sensing systems using wireless transmission usually have the sensors buried underground but the transmission takes place above ground. Specific challenges in meeting the goal of a completely buried wireless sensing network include:

- Significant attenuation through the soil, leading to reduced transmission range.
- Refraction at the soil-air interface causing directional transmissions and reception.
- Power consumption: the sensors need to be very low-power in order to continue functioning for a long period.

Currently, there is a number of low-power, long-range technologies, such as LoRa, 802.11ah, Sigfox, and NB-IoT that could be adapted to the underground transmission scenarios, but none of these provide a comprehensive, robust solution that will meet the needs of advancing subterranean science.

4.4.2 Research Needs**:** In order to address the above challenges, the conclusions from the workshop regarding the research that needs to be conducted are summarized below:

A. Research into extending range of subterranean wireless networks
    a. Mesh networking is a widely used option for extending the range of wireless networks. However, there are challenges to be overcome if the mesh architecture includes nodes that communicate strictly underground since soils will attenuate the signals considerably. Both physical and higher layer protocols need to be researched in order to address this specific problem.
    b. One option to extending the range from buried sensors is to use lower frequencies of operation, in the range 100 MHz – 400 MHz, since the attenuation is lower. However, antenna sizes increase at these frequencies, hence research into flexible antennas, antenna arrays etc. is required.
    c. A combination of existing technologies to extend range in single-hop, star networks: diversity reception, coding, etc.
    d. Instead of mesh or single-hop architectures, the use of drones as mobile hot-spots to interrogate underground sensors needs to be researched.

B. Research into higher throughput links
    a. Most existing technologies (Sigfox, LoRa, NB-IoT) have data rates ranging from 100 bps to 250 kbps. While this is sufficient for sensors today, future sensors may have higher throughput requirements. Attaining higher data rates without increasing power is challenging in the subterranean environment.

C. Research into extending battery life and energy scavenging.
   a. Research into protocol designs to reduce energy dissipation in sleep modes of sensors
   b. Device designs to reduce energy consumption.
   c. RF back-scatter methods: these have been shown to work in over-the-air communication. Feasibility in underground devices has not been studied.
   d. Others: microbial fuel cells, remote charging.

D. Alternative technologies
   a. Magnetic induction:
   b. Acoustic transmission.
      Hybrid architectures: combination of magnetic/acoustic and wireless technologies.

4.5 Hardware Technology: Innovation and Development Needs

Subterranean science challenges are fundamentally intertwined with the ability to continuously and precisely measure quantities in a soil environment that is perturbed as little as possible by the sensing modality. Hence a means for recording, storing, and then transmitting the sensor data is needed and this means needs to be miniaturized as much as possible to limit physical interference for the sensor package. Furthermore, the sensor system requires long-term operation allowing longitudinal studies. This requires a new electronics capability that demands new circuit technology innovation at both the circuit architecture and the system architecture level. In particular, what is needed is the following:

A. Ultra-low power draw (< 10nW average power consumption) circuit approaches for acquisition, storage, processing and communicating data. Power consumption is directly related to the size of the sensor. Conventional 10uW sensing system electronics can be readily obtained today. However, with a modest size coin cell for a cm-scale system, such a system would only sustain several days to a few weeks of operation which is clearly insufficient. Hence, instead of growing the battery size, which would lead to undesirable large size, it is required to dramatically reduce the power consumption. With < 10nW, even a small coin cell could sustain operation for several years of operation. However, achieving < 10nW average power consumption is a significant challenge that requires re-architecting of the processor, memory, sensor interface, timers, and power management.
B. Coupled with low power consumption is a need for energy harvesting for long-term, energy autonomous operation. If energy can be harvested at a rate exceeding 10nW, energy autonomy can be established and near-perpetual operation is achievable. However, energy scavenging underground is extremely difficult. Standard approaches rely on light (PV cells) temperature gradients (TEGs) or vibration (Piezoelectric), all of which are in short-supply underground. Chemical needs exploration and in particularly microbial base fuel cells show promise but need to have improved lifetimes and need to be made much more compact to fit into the same sub-cm form-factor of the entire sensing package.
C. A new class of electronics computation needs to be developed that perform “on-sensor” information processing, including inference and feature extraction. Such processing is critical since radio bandwidth is limited and radio transmission energy per bit can quickly dominate total energy consumption of the sensor node. Hence, instead of sending raw-data of radio links, it is critical for overall

power consumption that data is first processed and abstracted to features with much reduced size and hence much lower power consumption. Such "smart" sensors will need new processing capabilities which will need to be ultra-low power and adaptable for long-term operation.

D. Wireless sensors have the additional challenge that they are inherently un-anchored and will move over their multi-year lifetime. This will be especially true in agriculture where implements can introduce significant soil movement. Hence, there is a key need for electronics enabling 3D localization of sensor placement. Since GPS is unavailable underground, solutions should explore localization using dedicated infrastructure in the field. Furthermore, it needs to provide high accuracy resolution (cm range) and long operational lifetime (> 10 year) by being extremely low power. This could be aided by making localization relatively infrequency as sensor movement will be sporadic in nature.
E. Sensor nodes must have a means for wireless communication. Hence, new radio electronic circuits are required. Specifically, these circuits must be designed to support low power *underground* radio communication protocols and antennae. Underground communication tends to operate as low frequencies to enable better ground penetration. Hence, new low power RF oscillators, timers, mixers, power amplifiers, low noise amplifiers in appropriate bands are required. Furthermore, since data rate will be low and link budgets will be challenging, narrow band operation is desired, requiring ultra-stable frequency generation, but uW power budgets to meet instantaneous power supply from small, coin cell batteries.
F. Wireless sensors will be collecting a large array of new sensing modalities for precision agriculture. This could include new chemical sensing and physical metrics (H_2O. NO_3^-, P, CO_2, etc.). Each such sensor needs a specific electronic interface to allow the data to be acquired in digital form. Therefore, an array of new sensor interface circuits is needed to perform this analog to digital conversion. Energy efficiency and accuracy are paramount in these circuits and new circuit architectures to achieve lower power and higher accuracy are required.

4.6 Data Sciences: Specific Challenges

The soil system is extraordinarily complex, encompassing physics, chemistry, and biology at multiple length and time scales, from microscopic interactions at the level of individual root hairs to long-term and large-scale changes in soil structure, chemistry, and microbiology, all of which are subject to influence from weather, climate, plant genomics, and many other factors. Yet data about the soil system proper are sparse, uncertain, and expensive to obtain—both in absolute terms and when compared to other systems of comparable complexity, such as climate and the human body. Those data that exist are frequently collected in different environments, with different methods, and in different formats, and often include highly sensitive information such as soil characteristics and crop yields on individual fields. **Innovations in data science and data infrastructure are thus crucial** to both better understanding of soil systems and improved management strategies for soil evolution, agricultural intensification, and water conservation.

4.6.1 Data Science Specific Needs*:* Science advances via the interplay between observation and theory, with theory helping to plan and interpret observations, and observations used to evaluate models. The combination of system complexity and data sparsity that characterizes soil systems means that progress will require simultaneous improvements in sensors, data management and analysis, and soil and plant modeling. Models are needed to plan, interpret, and evaluate sensor measurements; data from both sensors and other sources (e.g., satellites, drones, weather models, agricultural yields) are needed to evaluate models. Research is needed to bridge the considerable gap that currently exists between experimental and theoretical studies, with the goal of achieving an integrated treatment of sensor design, data analysis and collection, and model enhancements that allows work in each area to inform and drive progress in the others.

New methods are needed to integrate data and models from many sources, at different spatial and temporal scales. Given the complexity of these data integration tasks, it will be essential to be able to capture these data integration processes in reproducible and transparent workflows that can identify potential data and model errors, produce uncertainty estimates, and enable automated updating of results as new data and new models become available. New methods will be required to integrate observational data that are highly diverse and often indirect (e.g., plant growth rates, satellite imagery, field runoff rates) into models.

The use of machine learning methods to create reliable, inexpensive surrogates for complex models is likely to prove fruitful but will require advances in data science methods and considerable experience in the field. Experiments need to conducted with a range of processes and process models for which empirical data (e.g. soil water content) are available and to see whether plant phenotypic data such as rooting depth or biomass, can be linked to data that can potentially be measured with sensors, such as water and nutrient accessibility.

Active learning methods should also be explored, with the goal of maximizing the benefit gained from new sensors and from new measurements by choosing the measurements that best improve the reliability of model predictions. (However, research is required to determine whether sensors can produce data that can be connected to models for active learning.) Transfer learning methods must be developed and applied with the goal of applying data and knowledge from one location, sensor, and problem to other environments. Given sparse, expensive, and sensitive data and the great diversity of soil environments, such methods are likely to be highly important for the soil and plant science problems.

Advances in data science methods, and the effective application of those methods to soil science and related problems, also requires new methods for effective large-scale data aggregation, so that data from different locations and sensing modalities—often individually sparse, but certainly “big” in aggregate—can be processed effectively by individual researchers and the community. An overarching goal should be to create a **Soil Science Commons** within which authorized researchers can ask queries of, and perform computations over, large collections of data from field experiments, sensor feeds, and other sources, aggregated on a national scale. Different vectors of data would exist, stacked vertically, that correspond to a “geo-pixel”, or an x-y co-ordinate of the land of small area, defining the spatial resolution of the data. Different layers of data can be employed, in physical models for soil, to create new layers of data—an example, for instance, is using soil moisture, wind speed, solar radiation and vegetation data to create estimated evapo-transpiration data. This Soil Science Commons may be centralized (e.g., in a commercial cloud) or distributed; regardless, it should be integrated at a level that allows for its use as a single logical resource, with methods for curating, analyzing, and using the data that it contains to drive both fundamental research on sensors and soil science and the development of the next generation of models and management strategies for soil evolution, agricultural intensification, and water conservation.

Creating this Soil Science Commons will require a combination of research, data infrastructure development, and community building. Research is required to develop methods for the efficient and secure transmission, curation, and analysis of large quantities of dense, in situ soil data over time; integration of data from multiple sources, of different types, formats, and quality; the organized integration of diverse soil and plant modeling systems to those data; and the development and application of machine learning and active learning methods to evaluate data quality, quantify uncertainty, improve models, and prioritize observations and model improvements.

Soil science and land sensing applications have unique needs with respect to data hosting, curation, privacy, and standardization. Methods for incentivizing data release are required, as many current data are inaccessible on hard drives. These problems are not unique to soil science and land sensing, but the diversity, spatial nature, and sensitivity of those data introduce specialized requirements.

4.7 Field Testbeds and a Soil Science Commons for Research Collaboration and Education

Innovation in sensors, data analysis methods, and other related areas is hindered by a lack of suitable testbeds for the coordinated testing of different sensors, models, and methods in controlled environments. Lacking such testbeds, individual researchers end up devoting much time and resources to the development of evaluation environments. Such one-off deployments are time-consuming and costly to produce, are inevitably limited in scope, and make comparisons across technologies and methods difficult.

To overcome this challenge, we need community infrastructure to enable coordinated testing of different sensors, models, and methods in controlled environments. Therefore it is essential to create field testbeds that contain subterranean sensing systems that are useful for increasing scientific understanding as well as providing data for management decision support.

Field testbeds will provide an environment in which many different sensors, communication technologies, and other innovations could be deployed and operated for extended periods within a limited geographical region, such as a single field. They would include support for the easy integration of new sensors, either individually or at scale, for example via standard interfaces and data collection fabrics. They would enable both short-term and long-term studies in which the same or different sensors, model, and methods are both studied independently and compared to each other.

Fortunately, several existing field networks are already involved in environmental monitoring and could lend themselves to testbed development. In some cases, these network share locations, and often teams are composed of university, federal and private scientists. Some examples are: (i), The Long Term Agro-ecosystem Research (LTAR) network, a joint USDA Agricultural Research Service, university and private institution endeavor that monitors the environment and ecosystems at 18 locations across the US (https://www.ars.usda.gov/natural-resources-and-sustainable-agricultural-systems/water-availability-and-watershed-management/docs/long-term-agroecosystem-research-ltar-network/); (ii), The National Ecological Observatory Network (NEON), a U.S. national observatory network sponsored by NSF and operated by Battelle (http://www.neonscience.org/) with 47 terrestrial sites; (iii), The Critical Zone Observatories (CZO), a U.S. national program funded by NSF to make environmental observations in the critical zone where the atmosphere, ecosystems, water, soil, and rock interact (http://criticalzone.org/national/) with 10 observatories in the US.

The Soil Science Commons discussed above represents another form of community infrastructure. These public platforms will make multi-valued soil data available for easy retrieval, and permit the sharing of both data and software, such as downstream analysis pipelines. The Soil Science Commons could be restricted to only public data but will be more useful if it can also hold private data, with mechanisms for controlling access to both raw and derived data.

Field testbeds and the Commons can and should be integrated, so that data from field testbeds can flow to the Commons for analysis and quality control, and analyses performed in the Commons can inform testbed work. The creation of field testbeds and the Commons is likely to both reduce the cost of experimentation and improve the quality of the science conducted. They can also provide an excellent environment for training students and new investigators. The Commons can host virtual testbeds for education that are then a real sensor array. These efforts can address the need for workforce development to inculcate the new skills required to apply modern data science methods in data science and agriculture.

5. Recommendations

The grand challenges identified in section IV, are reproduced here for convenience:

- Understand how the soil microbiome affects plant productivity, water and nutrient efficiency, and soil degradation
- Create a new generation of accurate terrestrial ecosystem models: build C, N and nutrient cycling models that offer predictive accuracy
- Food security: understand the root interface between the soil and the plant will have deep impact on global sustainability, food security and the environment.

These are science problems of deep significance for the sustainability of mankind, with impact on the evolution of the environment and food security. Our understanding of these phenomena and ability to predict and manage is not at that level it should be. There is the need for an ambitious program in establishing an Earth Macroscope—ultimately a vast buried sensor network that collects high resolution data that is then coupled to and informs an intense effort at furthering our knowledge in the three grand challenges described above. There is a need for model development for the soil, the plant and the soil-plant interface, coupled to experiments, and supported by pilot testbeds that will bring together convergence between soil scientists, plant scientist, microbiologists, computer scientists, nanotechnologists, and electrical engineers.

The convergence of these efforts will create a new generation of accurate terrestrial ecosystem models and build C, N, and nutrient cycling models that offer predictive accuracy; it will understand how the soil microbiome and sub-surface processes (e.g. carbon transformations, physical and chemical controls, turnover rates) affects plant productivity, water and nutrient efficiency, and soil degradation; and it will understand the root interface between soil and plant.

In order to accomplish the above**, there are critical dependencies on the role of Computer Science and Engineering research**, there are four areas in particular where it is essential for research involvement from the engineering and computer science disciplines. There are specific needs and opportunities here for the following:

A. **Sensors, sensor materials, and sensor device research:** that needs to be accelerated, for accurate, low energy high spatial and temporal resolution subterranean sensing:
    a. Chemical functionalization strategies for robust, reproducible selectivity and specificity
    b. Compact optical/electrical platforms
    c. "Proxy sensors" and analytics inference engines for difficult to measure quantities using a suite of sensors. e.g: infer root morphology instead of imaging
    d. New imaging technologies for root and root exudates
    e. Key parameters to measure: CO_2, N_2O, key nutrients (NO_3, P), root biomass, root turnover, root surface area, structure and density.
    f. Compact, chip scale PCR with stable reagents for in-field soil microbial sensing
B. **Micro/nano electronics research**: Develop a class of long lifetime (> 10 yrs) electronic sensor technology enabled by:
    a. Ultra-low power draw (< 10nW average power consumption)
    b. New energy harvesting capabilities for continuous and *long-term* underground energy scavenging
    a. New class of intelligent electronic sensor nodes that perform "on-sensor" or at-the-edge information processing
C. **Computer science research**:
    a. New methods for sensor+data+model integration
    b. Large-scale aggregation of data across the community
    c. Geographical testbeds with multidisciplinary expertise to enable coordinated R&D on sensors, data solutions, models
D. **Wireless Technology research**: Need for robust fully buried low power sensing network with high range and throughput
    a. low-power IoT radios available today (e.g. LoRa, Sigfox, NB-IoT) do not fully meet the cost and power requirements for a subterranean sensing network.
    b. Clear needs for research in both the wireless protocol domain and device level implementation

6. Workshop Agenda

Day 1: Wednesday, November 1	
Time	**Event**
7:55a – 8:05a	**Welcome and Goals** Supratik Guha, *University of Chicago*
8:05a – 9:20a	**Opening Plenaries** **Sensor Networks for Agriculture's Unchartered Frontier** *Nick Dokoozlian*

	Vice President, Viticulture, Chemistry and Enology *E&J Gallo Winery* **Understanding the Role of Soil in the Genetics x Environment x Management Concept** *Jerry Hatfield* *Laboratory Director and* Supervisory Plant Physiologist *National Laboratory for Agriculture and the Environment* **Research needs for sensing and monitoring biological analytes in buried soil environment** *Rajakkannu Mutharasan* *Frank A. Fletcher Professor of Chemical and Biological Engineering* *Drexel University* Plenary Chair: Chuck Rice
9:20a – 9:40a	**Plenary Q&A**
9:40a – 9:55a	**Break**
9:55a –11:10a	**Opening Plenaries** ***A World Without Soil*** *Jo Handelsman* *Director of the Wisconsin Institute for Discovery at the University of Wisconsin-Madison* *Vilas Research Professor* *Howard Hughes Medical Institute Professor* ***Seeing Beneath the Surface: The Role of Big Data and Computing*** *Ian Foster* *Arthur Holly Compton Distinguished Service Professor* *Department of Computer Science, University of Chicago* *Distinguished Fellow, MCS Division* *Senior Scientist, MCS Division, Argonne National Laboratory* **Agricultural Internet of Things: View from the Field** *Mehmet Can Vuran* *Susan J. Rosowski Associate Professor* *Cyber-Physical Networking Laboratory* *Computer Science and Engineering* *University of Nebraska-Lincoln* Plenary Chair: Chuck Rice
11:10a –11:30a	**Plenary Q&A**
11:30a –12:15p	**Lunch**
12:15p – 1:30p	**Panel 1: Soil physics, chemistry, and microbiology** *Presenters:* *1. Tyson Ochsner, Oklahoma State University* *2. Katalin Szlavecz, Johns Hopkins University* *3. Jennifer Pett-Ridge, Lawrence Livermore National Laboratory* *4. Zoe Cardon, Marine Biological Laboratory* *5 David Brown, Washington State University* *6. April Ulery, New Mexico State University* *7. David Myrold, Oregon State University* Panel Chair: Roberto Cesar Izaurralde

1:30p – 2:15p	**Panel 2: Plant genomics, predicting phenotype from G x E** *Presenters:* *1. Steve Welch, Kansas State University* *2. Chris Topp, Danforth Plant Science Center* *3. Alison Thompson, USDA ARS* *4. Edgar Spalding, University of Wisconsin-Madison* *5. David Baltensperger, Texas A&M University* *6. Steve Evett, USDA ARS* Panel Chair: Ali Mohamed
2:15p – 2:30p	**Break**
2:30p – 3:30p	**Panel 3: Sensors and Subsystems** *Presenters:* *1. Viacheslav Adamchuk, McGill University* *2. Xufeng Zhang, University of Chicago* *3. David Blaauw, University of Michigan* *4. Hongda Chen, USDA NIFA* *5. Michael Haley, University of Oregon* *6. Agnelo Silva, University of Southern California* *7. Raphael Viscarra Rossel, CSIRO* *8. James Krogmeier, Purdue University* Panel Chair: Supratik Guha
3:30p – 4:30p	**Panel 4: Big Data** *Presenters:* *1. Bruno Basso, Michigan State University* *2. Alex Szalay, Johns Hopkins University* *3. Ken Birman, Cornell University* *4. Greg Gandenberger, Uptake* *5. Deb Agarwal, Lawrence Berkeley National Laboratory* *6. Ranveer Chandra, Microsoft* *7. Alok Choudhary, Northwestern University* Panel Chair: Ian Foster
4:30p – 4:45p	**End of Day Wrap Up** *Participants sign up for day 2 breakout sessions*
7:00p – 9:00p	**Conference Dinner** *315 N LaSalle Dr, Chicago, IL 60654*

Day 2: Thursday, November 2	
Time	**Event**
8:00a – 8:10a	**Overview of Breakout Sessions** Supratik Guha
8:10a – 8:40a	**Keynote: Big Data gets Physical** *Hendrik Hamann* *Distinguished Researcher and Research Manager for Physical Analytics, IBM Thomas J. Watson Research Center*
8:40a – 8:45a	**Head to breakout room sessions**
8:45a –9:45a	**Concurrent Breakout: Soil Science** Session Moderator: Roberto Cesar Izaurralde **Concurrent Breakout: Plant Genomics** Session Moderator: Steve Evett
9:45a – 10:00a	**Break**
10:00a – 11:00a	**Concurrent Breakout: Subsystems and Infrastructure**

	Session Moderator: Monisha Ghoish **Concurrent Breakout: How can data analytics help soil science?** Session Moderator: Ian Foster
11:00a –11:30a	**Wrap Up and Develop Summary Presentations** *Moderators develop 1-2 slide summary*
11:30a –1:00p	**Presentation Readout** *10 minutes per group*
1:00p	**Workshop Concludes**
1:30p – 4:30p	**Report-Writing Session** *Advisory committee meets with breakout group leaders to review and create a skeleton draft of the report.*