NATIONAL INSTITUTE OF FOOD AND AGRICULTURE
Data Summit: Changing the Face, Place, and Space of Agriculture

RESULTS OF "IDEAS ENGINE" STAKEHOLDER INPUT

The U.S. Department of Agriculture's National Institute of Food and Agriculture (NIFA) invited over 1000 stakeholders who are interested in various aspects of data to participate in an "Ideas Engine" stakeholder-input mechanism to help identify priorities for NIFA's new Food and Agriculture Cyberinfrastructure and Tools (FACT) initiative. The Ideas Engine system, by Codigital, is an online web-based platform for large-scale collaboration, allowing many ideas to evolve and fuse into concise collaborative output in real time. The system enables entire groups' creative suggestions to emerge, and opinions to be revealed, in a dynamic, democratic, online process. Participants continually and collaboratively process each other's points of view—ideas flow, cross-fertilize, and evolve. At the end of the process, the best ideas float to the top, all input is captured, and analytics describing meaningful patterns in the data around the process are produced.

Features of the system are:

- Quality, representative output
- Easy to use
- Efficient
- Fun
- Self moderating

The process involves:

1. Adding ideas
2. Voting on ideas
3. Editing ideas
4. Voting on edits

Six groups, comprising 40 to 75 participants each, collaboratively developed their best ideas on the most promising opportunities for the USDA NIFA FACT initiative to address data needs in agriculture. Each group was posed an open-ended question about opportunities for NIFA in data-related activities. Those questions were:

- What are the most promising opportunities for data-driven advances in agriculture and the food-production system?
- What are the most promising opportunities for enhancing cross-sector advances in data applications?
- What are the most promising opportunities for data-driven advances to address societal well-being and consumer demands?
- What are the most promising opportunities to address challenges of various facets of data management and application?
- What are the most promising opportunities to ensure future generations of data expertise?
- What are the most promising opportunities for big data in communication, property rights, and communities?

The Ideas Engine was enabled in conjunction with the NIFA *Data Summit: Changing the Face, Place, and Space of Agriculture*. It operated as open access for all participants to contribute—**at their convenience, as often as they desired**—over a period of three weeks, before, during, and after the summit. Individuals provided their ideas, edited and refined ideas, and voted on ideas over that period (in one instance, over 1700 votes were cast). Results of those co-creation and collaborative efforts are presented in the following pages.

USDA United States Department of Agriculture | National Institute of Food and Agriculture

What are the most promising opportunities for data-driven advances in agriculture and the food-production system?

2016-11-01

76	46	28	1724
Contributors	Ideas	Edits	Votes

TOP 3 IDEAS

Rank	Idea	Generation	Score
1	USDA and universities need to develop a platform that enables access to the vast amounts of public and private (e.g. farm / field) data in a way that does not expose (e.g. FOIA) the individual farmer but advances research and learning.	1	65.5%
2	A focus on Open Data FAIR principles: Findable, Acceptable, Interoperable, and Re-usable for all experimental and research data. USDA needs to lead development of standards, best practices and work w/ NIH, NSF, DOE, and NIST (eg Mendeley Data).	4	64%
3	Establish public-private partnerships to create application program interfaces to readily share data and develop user-friendly platforms which facilitate the access to data repositories and customized outputs of bigdata analytics	1	61.9%

ALL 45 IDEAS

Rank	Idea	Generation	Score
1	USDA and universities need to develop a platform that enables access to the vast amounts of public and private (e.g. farm / field) data in a way that does not expose (e.g. FOIA) the individual farmer but advances research and learning.	1	65.5%
2	A focus on Open Data FAIR principles: Findable, Acceptable, Interoperable, and Re-usable for all experimental and research data. USDA needs to lead development of standards, best practices and work w/ NIH, NSF, DOE, and NIST (eg Mendeley Data).	4	64%
3	Establish public-private partnerships to create application program interfaces to readily share data and develop user-friendly platforms which facilitate the access to data repositories and customized outputs of bigdata analytics	1	61.9%
4	New technologies can generate tremendous amounts of information, but it is not easily reuseable. New mechanisms are needed to facilitate data reuse. This will improve research and analysis, verification, and information trust.	3	59.8%
5	The ability to conduct predictive biology and predictive science that marries existing and new data for use in making informed future decisions in genetics or management of agricultural crops and animals.	3	56%
6	Growers need real-time apps that provide decision support for plant and animal timing including planting, breeding, treatment, harvesting, and watering, for the local climate and to inform management practices throughout the season.	3	55.4%
7	To integrate data on climate, cropping, animal production, food and bioprocessing, and distribution systems with data on consumer and community needs and desires so that decisions can be made in context of effects on the entire food system.	3	55.3%
8	Use precision ag technology to generate data in large-scale, on-farm field trials. The data could be used to better understand yield response to inputs, weather, and field characteristics. This is needed to manage well site-specifically.	1	55.1%
9	Sharing of data needs to be rewarded on a par near that of research and profit outcomes. Researchers must have a first right to publish and must cited by subsequent users when data is released. Industries should focus on developing applications.	1	53.9%
10	Vast amounts of potentially useful data reside deep within various private-/public-sector organizations. An Open Data Processing Service would process proprietary/protected data to extract anonymized data for use by app developers and researchers.	1	53.6%
11	Ensure modern farming management services (e.g. Smart Farming) are designed bottom-up (i.e. taking input from farmers at design level) taking advantage of both farmer input and using the latest innovations in crowdsourcing and dynamic data analysis.	3	53.1%

Rank	Idea	Generation	Score
12	It is better to adopt computing infrastructure similar to big corporations (such as Microsoft, Apple, Oracle, Google, Amazon, Facebook and twitter) for data mining and machine learning applications to inform decision-support	1	52.7%
13	The infrastructure for reading data from modern agricultural technologies has motivated development of open data exchange formats which need to be prioritized to recognize their potential as a framework for delivering research results to producers.	3	52.1%
14	The USDA has to create and fund, with states and universities, the drive to integrate systems for overall efficiency, and encourage and reward publishing of such work as opposed to 'novel scientific findings'. Data needs to drive decisions.	2	51.9%
15	Ag school graduate education programs should include a data management planning (DMP) component as well as training in meta-analysis. Many Ag-School grads (MS, PhD) are unaware of these areas of work.	1	51.7%
16	A new type of AgSchool graduate / program is needed, one that combines agro-ecosystem knowledge with analytical skills in areas such as: machine learning and data mining; statistical and quantitative Analysis; data visualization and problem solving.	1	51.2%
17	Models driven by genetic data would reduce time required to develop new varieties, allowing prediction of performance in target environments. Many high quality datasets are required to characterize dynamic GxExM processes over a range of cultivars.	1	51%
18	There is a need to work on inconsistencies regarding standards, data quality, security, policy, data privacy, confidentiality, and ethics regarding the retaining and sharing/transfer of data	1	50.4%
19	Develop the ability to rapidly obtain, process, and utilize remotely sensed data from small drones to populate prescriptive maps used for precision planting, irrigation, pest and disease control, nutrient/fertilizer management, and harvesting.	1	50.3%
20	The key to effective agricultural data analytics is sharing data across farms, but competition/privacy concerns limit this. Can we develop privacy-preserving analytics for spatiotemporally-correlated fields, e.g. drawing on information theory?	1	49.9%
21	Several TB of heterogeneous data sets exist in public domain whose potential is yet to be explored. Protocols need to be developed to merge these datasets and develop platforms for the benefit of research, extension, education and policy makers.	1	
22	Develop multi-scale early warning network - e.g. linking machine learning and open data with climate/risk indicators to help track vulnerabilities in agriculture regions, and the potential for cascading events across landscapes and markets.	1	
23	Numerous remote sensing applications based on satellite imagery have been developed and tested before (mapping biomass, water use, disease outbreaks, etc.). Developing new tools for accessing and interpreting the data can help production systems.	2	

Rank	Idea	Generation	Score
24	Must address the various facets of risk vs incentive, recognition and reward for data sharing for different types of data (e.g., Experimental vs Funder Program) by different types data producers (e.g. Researchers vs Private Sector).	1	
25	Food manufacturing, distribution and storage are critically dependent on food physical, chemical and microbiological properties data. Predictive ability for safety/risk/quality of foods are enabled by such properties that need to be made available.	1	
26	Create opportunities for the conversations aligning the interests of those developing data management systems and analysis with the values and existing knowledge of practitioners (e.g. farmers) to sustain healthy, reliable, food sources for all.	3	
27	The fed funders MUST work together, not in separate silos: USDA, NIH, NSF, DOE, NIST. A common proposal format is needed.	2	
28	New sensing technologies with integration measured data at the cloud level to allow for data-driven insights to be generated for management down to the plant and/or animal level.	1	
29	Monitor plant (i.e., agricultural crops) and animal diseases, identify potential hotspots for plant and animal diseases, and predict the spread of the disease(s) at the local, county, state, country, and global levels.	1	
30	There is need to establish massive computation resources (GPUs instead CPUs), new robust and efficient computational and visualization tools to improve data linkages between diverse/heterogeneous datasets and generate new insights/associations	1	
31	Facilitate true collaboration between academic/government research and industry for development of products and technology transfer -needs to be more than industry funding academic research.	1	
32	Basic biology and ecology is as important as genomics, automated management, and HT phenoyping in advancing agriculture. For example, little is known about the diversity of beneficial insects that can do the work of many of the agricultural inputs.	1	
33	Cultivate youth capacity to be data visionaries, producers, disseminators and consumers with respect to agriculture and food production, especially among the most under-represented populations and places in this field.	1	
34	A web-based soil test calibration database for North America that promotes open analysis of crop response to nutrient additions, allows multi-objective recommendations adjusted for location and management system, and includes sustainability factors.	2	
35	Engage librarians as information science experts (esp those in Ag domain) along with computer science and data science experts.	1	

Rank	Idea	Generation	Score
36	Big data in livestock species provides unprecedented opportunity for the discovery of genes and genetic mechanisms underlying phenotypes and for livestock genetic improvement using genomic data.	1	
37	Bibliometrics/Altmetrics and Text & Data Mining are powerful approaches to uncover new findings and linkages between research areas.	1	
38	Bibliometrics/Altmetrics (Scopus APIs + metrics) and Text & Data Mining (ScienceDirect APIs and CrossRef Text & Data Mining) are powerful approaches to uncover new findings and linkages between research areas.	1	
39	Must consider the role of library & information science experts, i.e. librarians, along with the role of engineering and math researchers in computer science and data science to address this issue.	1	
40	the opportunity to increase efficiency of animal and crop improvement and production by means of more specific genetic selection using multiple traits	2	
41	A smart idea is more than a set of analyses and the interpretation of machine data. Connecting farmers from idea generation to science and ideas to enable a market inquiry and introduction into production rapidly .What model and algorithms ?	2	
42	Efficiency in the farm office: (1) a paperless farm can be realized for all farms in only a few years, and (2) data can be used for benchmarking and big data analysis.	2	
43	There is need to develop more pipelines with broad bandwidth and better security system (with advanced encryption and de-identification algorithms) for data transfer from location to another	1	
44	Until university systems for tenure and promotion change from grants and papers to results and impacts, the data will not be used to their potential. There is useful data unused because of traditional funding and publishing measures.	2	
45	Creation of innovative physical products at once under the request of the situation. such as an increase in temperature on earth more than 1.5 deg Experimental products exposed to the market if necessary for production.	2	

New Ideas
(not enough vote data for ranking)

Relegated Ideas

United States Department of Agriculture | National Institute of Food and Agriculture

What are the most promising opportunities for enhancing cross-sector advances in data applications?

2016-11-01

42	27	13	648
Contributors	Ideas	Edits	Votes

TOP 3 IDEAS

Rank	Idea	Generation	Score
1	A focus on Open Data FAIR principles: Findable, Accesible, Interoperable, and Re-usable for all experimental and research data. USDA needs to lead development of standards, best practices and work w/ NIH, NSF, DOE, and NIST.	2	62.8%
2	Better support of multidisciplinary training of students and projects that combine ag experience with data management and mining expertise	1	60%
3	Federal funding agencies MUST work together, not in separate silos: USDA, NIH, NSF, DOE, and NIST.	2	57.3%

ALL 26 IDEAS

Rank	Idea	Generation	Score
1	A focus on Open Data FAIR principles: Findable, Accesible, Interoperable, and Re-usable for all experimental and research data. USDA needs to lead development of standards, best practices and work w/ NIH, NSF, DOE, and NIST.	2	62.8%
2	Better support of multidisciplinary training of students and projects that combine ag experience with data management and mining expertise	1	60%
3	Federal funding agencies MUST work together, not in separate silos: USDA, NIH, NSF, DOE, and NIST.	2	57.3%
4	Encourage depositing of data sets in trustworthy, sustainable data repositories and publishing of data articles (about research data sets) in open access, multi-disciplinary data journal (e.g. Data in Brief, published by Elsevier).	1	57%
5	It is paramount that data-driven science adhere to the same standards of classical science, in particular with respect to reproducibility. Adherence to this standard would immediately eliminate the ego-driven story-telling so prevalent currently.	1	54.3%
6	Build infrastructure and revamp the reward system in academia and business to integrate systems, share data, and solve problems instead of just get grants and publish papers that are seldom ever read. USDA was created to solve problems.	3	54%
7	Work toward semantic interoperability, building infrastructure (variabe type registries, controlled vocabularies, etc.) that can enable data exchange with shared meaning across agriculture, including industry, government, and academia.	1	53.8%
8	There is a need to work on inconsistencies regarding standards, data quality, security, policy, data privacy, confidentiality, and ethics regarding the retaining and sharing/transfer of data	1	52.7%
9	Take advantage of other agencies (NSF, NIH, NOAA) and their expertise in cyberinfrastructure. Then develop competitive grant program that takes advantage of existing infrastructure to apply to agriculture.	3	52.6%
10	Nothing speaks louder than success. And new data management strategies are best developed through real-world applications. We need to Identify a diverse set of case studies that demonstrate the benefits of cross-sector data applications.	2	51.7%
11	Reach out to experts working in a variety of sectors with different foci, but similar sets of problems (academics to practioners, data vs knowledge, laboratory to real world) at nexus of science and society.	4	51.4%
12	Must address the various facets of risk vs incentive, recognition and reward for data sharing for different types of data (e.g., Experimental vs Funder Program) by different types data producers (e.g. Researchers vs Private Sector).	2	51.1%

Rank	Idea	Generation	Score
13	Support training and online educational opportunities in data mining and data visualization for scientists and researchers	1	50%
14	Precision agriculture can increase efficiency to minimize environmental impacts and maximize farm profits. Integration of GIS, climate data, and biological rates (photosynthesis, water use, etc) are needed on a single, interoperable platform.	1	49.6%
15	Bibliometrics/Altmetrics (Scopus APIs + metrics) and Text & Data Mining (ScienceDirect APIs and CrossRef Text & Data Mining) are powerful approaches to uncover new findings and linkages between research areas.	1	49%
16	Need for support from USDA/Funding agencies to establish Bigdata infrastructure (e.g. NIH BD2K for data submissions, data sharing repositories) which can facilitate a well-defined computational architecture to facilitate large scale data analysis	1	47.6%
17	Offer competitive programs (has to be a CAP to be big enough to attract a cadre of talented scientists) that take a food systems approach to agriculture and integrate the various disciplines.	1	47.4%
18	It is better to adopt computing infrastructure similar to big corporations (such as Microsoft, Apple, Oracle, Google, Amazon, Facebook and twitter) for data mining and machine learning applications to inform decision-support	1	47.3%
19	The food system is a system. We can use integrated data to help make decisions in one sector (crop fertilization) that can have effects through the entire sector (energy use, water need, crop production, animal feeding and efficiency, distribution)	2	45.2%
20	There needs to be an incentive or driver. Find out what motivates people to cross over into other sectors. Is it through standard conferences? Maybe something less formal, such as a social gathering. Some individuals don't need either, why?	1	44.7%
21	Identify a problem (agriculture or food science related). Seek out data available in government, public, and private sectors to address the issue, saving time and money.	1	43.8%
22	Support symposia, workshops, meetings etc. to facilitate transdisciplinary collaborations among researchers, data scientists, computer engineers etc.	1	43.3%
23	Integrate ORCID into the NIFA/USDA grants system and start using ORCID on/at all NIFA/USDA functions to connect researchers with their research, enhance discoverability and impact.	1	37.7%
24	Support the use of big data to train and assist small and medium size producers	1	28.4%
25	Interoperable food ontologies working across the food pipeline, from farm to fork to food waste, allowing optimization for sustainability that explicitly considers uses downstream from agriculture. https://publish.illinois.edu/food-and-data-workshop/	1	

Rank	Idea	Generation	Score
26	Respond to government and public surveys designed to identify data that people need, want, and value collected for agriculture and food sciences.	2	

Relegated Ideas

United States Department of Agriculture | National Institute of Food and Agriculture

What are the most promising opportunities for data-driven advances to address societal well-being and consumer demands?

2016-11-01

single organic address research onfarm researchers usda demand agricultural seeds farmers climate big open nifa water make weather require creating program fund examine **data** small change establish product products food needed agriculture apis standards producers practices work collection soil development develop health digital science quality consumers public field

50	35	10	733
Contributors	Ideas	Edits	Votes

TOP 3 IDEAS

Rank	Idea	Generation	Score
1	Establish a joint US Department of Education and USDA NIFA research and education competitive grants program focused on data science.	1	64.3%
2	Develop partnerships that engage producers, ag industry, governmental agencies and academia in on-farm research and data collection and set performance standards for data collection.	1	63.8%
3	Fund Workshops/Convenings focused on building community around digital agriculture and that link public and private partners.	1	62.8%

ALL 35 IDEAS

Rank	Idea	Generation	Score
1	Establish a joint US Department of Education and USDA NIFA research and education competitive grants program focused on data science.	1	64.3%
2	Develop partnerships that engage producers, ag industry, governmental agencies and academia in on-farm research and data collection and set performance standards for data collection.	1	63.8%
3	Fund Workshops/Convenings focused on building community around digital agriculture and that link public and private partners.	1	62.8%
4	Encourage adoption of open, cloud-based API's to make data available to more people capable of using it, and fund open source software libraries to reduce barriers to the use of data.	1	62.3%
5	Availability of interoperable data from a variety of sources will allow evidence-based decision making, e.g., data linking nutrition with health outcomes combined with agricultural production data to support a demand-driven agricultural system.	1	62.1%
6	Funding projects which focus on producing quality public datasets must be a priority if data in ag is ever to progress beyond the novelty stage. There is simply too little data available in accessible form for effective machine learning efforts.	2	58%
7	The fed funders MUST work together, not in separate silos: USDA, NIH, NSF, DOE, NIST. Common grant application formats and protocols are needed.	2	57.2%
8	A focus on Open Data FAIR principles: Findable, Acceptable, Interoperable, and Re-usable for all experimental and research data. USDA needs to lead development of standards, best practices and work w/ NIH, NSF, DOE, and NIST.	2	55.2%
9	Apply the technology and resources to identify genetic potential of plants to mitigate problems that may be encountered regionally from climate change plus management practices.	2	55.2%
10	Help farmers use data to grow crops safely & efficiently: establish the infrastructure (semantics, common formats, reference data APIs, incentive mechanisms, etc.) that can enable principled decision-making in agricultural field operations.	2	55.1%
11	Addressing consumer demands require a 2-way communication avenue where general public can not only view, graph, and download data, but to comment and provide direct feedback on what is useful and what is not.	1	54.1%
12	Engage multiple communities and stakeholders across sectors to develop both descriptive and encoding metadata standards to promote cross-disciplinary analysis. Require their use.	1	54.1%

Rank	Idea	Generation	Score
13	Collect and make available data generated in organic agriculture experiments. It's (almost always) not big data, but should be part of the ag-data ecosystem, and available to both researchers and growers. May require standard(s) development.	1	52.2%
14	We need to know how to make use of what is in the field before adding an input. Band-aids are not going to work for ag systems faced with the realities of climate change and an increasing demand for local food produced with sustainable practices.	1	50.8%
15	NIFA program for data science/ analytic infrastructure like a small CAP around $2 - $5 mil could facilitate linking many agricultural data sets, currently unconnected and allow for more robust analyses and modeling to occur.	1	50.7%
16	Must address the various facets of risk vs incentive, recognition and reward for data sharing for different types of data (e.g., Experimental vs Funder Program) by different types data producers (e.g. Researchers vs Private Sector).	1	50.1%
17	Consumers demand processed foods that do not utilize artificial ingredients. Sponsor research to develop database of how different innovative food process parameters influence pathogenic and spoilage organisms, nutrients, allergens, and toxins	2	50%
18	Bibliometrics/Altmetrics (Scopus APIs + metrics) and Text & Data Mining (ScienceDirect APIs and CrossRef Text & Data Mining) are powerful approaches to uncover new findings and linkages between research areas.	1	48%
19	Agricultural food production systems are rooted in soil and its health. Farmers need a good assessment of on-farm soil health. Digital ag tools need development for farmer use. The tools can also be used to inform consumers about the food produced.	1	47.9%
20	Collect media consumption data to understand the publicâ€™s preferred terminology. Use those data to improve ag message development and delivery methods. Test change in audience knowledge, attitudes, and behaviors after consuming related information.	1	47.8%
21	Fund projects that link 1862s, 1890s, 1994s and Hispanic Serving institutions around digital agriculture.	1	46.8%
22	Using precision ag tech to conduct on-farm field trials, measure yields and nutrient loss, with purpose of minimizing social costs of cleaning up the hypoxic "dead zone" in the Gulf of Mexico and other key watersheds.	2	45.9%
23	Form independent oversight panels to promote discourse, resolve disputes, investigate failures &recognize successes in data driven policies. Human oversight is vital to deal with misguided plans, inadvertent bias, malicious actors & ethical issues.	1	45.8%
24	Big data can increase the transparency of company/product pricing (e.g. health insurance, health care, pharmaceuticals), individuals/ families can efficiently use resources and increase household financial security	1	44%

Rank	Idea	Generation	Score
25	Establish within NIFA AFRI a separate competitive grants program on digital agriculture.	1	42.6%
26	Agriculture needs on-farm, rural-based weather data. A joint standards panel for weather data collection, instruments, proper siting and metadata is needed. Quality weather data is needed for on-farm decision making as part of precision agriculture.	1	
27	Must consider the role of and collaborate with library & information science experts, i.e. librarians, along with the role of engineering and math researchers in computer science and data science to address this issue.	2	
28	First, actually getting data to a larger percentage of consumers...real data, it need not be big (for example there is nothing wrong with gluten, animals are not routinely mistreated, and they are not ruining the planet and causing warming).	3	
29	Creating a reference, Standards on organic healthy foods.Control at all stages: soil, water, seeds, plants, storage, processing, finished product delivery to the table Formation of consumer demand for organic natural products.	1	
30	Examine the difference between the perception of some condition (e.g. GMO corn is in everything) versus measurements of the condition (e.g. my grocery receipt shows X calories are from GMO products)	1	
31	continuing on getting real data, large or small to consumers, meat is not laced with antibiotics, 'consumers' are not clamoring for 'plant based diets', food poisoning is not rampant, meat is not bad for you....	1	
32	By 15: consider the soil, water and seeds. This approach gives a very good result	1	
33	Support medium and small scale producers using data science	1	
34	19. Seen in the context of - the use of antibiotics, resistance to them, food poisoning and the quality of products in the global climate change context	1	
35	Creating a unique platform for organic agriculture.To examine the relationship: soil, water, seeds, plant, storage, processing, final product, food safety. .Creating mixed synthetic modeling a single product, single health, single funding.	1	

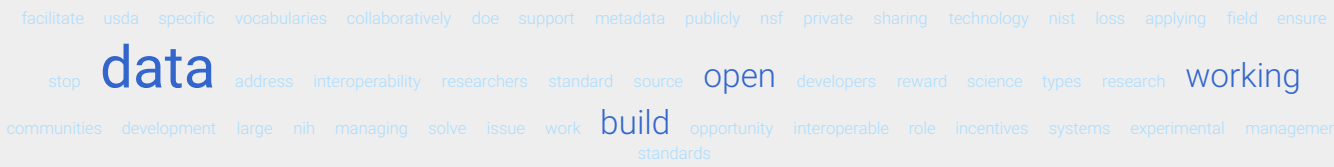
New Ideas
(not enough vote data for ranking)

Relegated Ideas

United States Department of Agriculture | National Institute of Food and Agriculture

What are the most promising opportunities to address challenges of various facets of data management and application?

2016-11-01



38	21	8	462
Contributors	Ideas	Edits	Votes

TOP 3 IDEAS

Rank	Idea	Generation	Score
1	Build systems that support good data management and maintain provenance through the data life cycle, starting at creation/collection. Ensure that those systems support data processing and aggregation without loss to allow later de-aggregation.	1	65.2%
2	A focus on Open Data FAIR principles: Findable, Accesible, Interoperable, and Re-usable for all experimental and research data. USDA needs to lead development of standards, best practices and work w/ NIH, NSF, DOE, NIST, GODAN, CGIAR.	3	65.2%
3	Need to put in place incentives to stop loss of data never submitted to repositories/published. Incentives could include citing use/re-use of data, working with Ag Journal Editors, data sharing plans for USDA grant proposals.	1	57.2%

ALL 19 IDEAS

Rank	Idea	Generation	Score
1	Build systems that support good data management and maintain provenance through the data life cycle, starting at creation/collection. Ensure that those systems support data processing and aggregation without loss to allow later de-aggregation.	1	65.2%
2	A focus on Open Data FAIR principles: Findable, Accesible, Interoperable, and Re-usable for all experimental and research data. USDA needs to lead development of standards, best practices and work w/ NIH, NSF, DOE, NIST, GODAN, CGIAR.	3	65.2%
3	Need to put in place incentives to stop loss of data never submitted to repositories/published. Incentives could include citing use/re-use of data, working with Ag Journal Editors, data sharing plans for USDA grant proposals.	1	57.2%
4	The key lies in interoperability of data; Ensure an underlying data layer based on existing interoperability standards. Then, different facets/user interfaces can "sit" on top of them, offering different functionalities & serving different purposes.	1	56.8%
5	Need for automated means of annotating data as it is captured, e.g. from field experiments, with standard vocabularies and metadata extraction. Such systems will ensure that data are interoperable when the collector is ready to release them.	1	55.5%
6	Facilitate the training of data management specialists who can work with scientists to ensure that data are properly annotated with metadata and linkages to standardized vocabularies.	2	51.8%
7	Entities that facilitate development of open data exchange standards (e.g. AgGateway) are an opportunity to foster consensus on data needs and to promote technology transfer via public/private partnerships.	1	49.7%
8	Must address the various facets of risk vs incentive, recognition and reward for data sharing for different types of data (e.g., Experimental vs Funder Program) by different types data producers (e.g. Researchers vs Private Sector).	1	49.2%
9	Producing and sharing data needs to be part of the research system. Journals, departments and colleges need to consider this in promotion and tenure systems to reward integration of food production systems.	3	48.2%
10	What is the overall goal for each agricultural/food-production challenge? Is it to prevent a major outbreak or solve a minor issue? Is there a need for both private and public sectors? Then, cross-sector advances in data applications could happen.	1	47.2%
11	Must consider the role of library & information science experts, i.e. librarians, along with the role of engineering and math researchers in computer science and data science to address this issue.	1	47%
12	rebuilding university capacity to solve problems, not just get grants and publish papers that are never used. strengthen national research committees with specific goals for generating, sharing and publishing large data to integrate agriculture	2	43.2%

Rank	Idea	Generation	Score
13	Build open source communities of developers working collaboratively with practioners on managing and applying data in a way that is publicly transparent.	2	43.2%
14	We have the opportunity stop using "Big Data" as a buzzword. We can can name specific large data sets to generate. Example: We need way more on-farm, large-scale field trial data, and can use precision technology to run the trials. Coordinated.	2	42.5%
15	The fed funders MUST work together, not in separate silos: USDA, NIH, NSF, DOE, NIST.	1	39.6%
16	Develop a standard format, structure, and reference for sharing data. Data can be recycled, aggregated, and/or mined after the initial collection, if there they are consistent and have a usable coding sheet	1	35.7%
17	Build open source communities of developers working on managing and applying data collaboratively and publicly.	1	
18	Build open source communities of developers working on managing and applying data collaboratively and publicly.	1	
19	Build open source communities of developers working on managing and applying data collaboratively and publicly.	1	

Relegated Ideas

United States Department of Agriculture | National Institute of Food and Agriculture

What are the most promising opportunities to ensure future generations of data expertise?

2016-11-01

math young program engineering sciences problems **data** sharing information biology types school **science** postdocs professors issues extension agriculture role application grad access building create **computer** courses experiential teachers reward solving k12 management systems high engage learning integrate activities online programs software develop **students** meaningful big show incorporate datasets

44	23	13	764
Contributors	Ideas	Edits	Votes

TOP 3 IDEAS

Rank	Idea	Generation	Score
1	Develop an experiential learning system for public-private partnerships that will match undergraduate students with teams of industry and/or community practitioners and university faculty to find data-driven solutions to real world problems.	3	63.3%
2	Incorporate real-world issues into all computer science and data science programs. All STEM areas develop ways to introduce the role of data in all,fields and societally relevant problems.	1	61%
3	Build an infrastructure in academia and industry that rewards integration of ag systems, that keeps ownership as open as possible, encourages and rewards data sharing and patents and IP, reward solving problems instead of increasing profits	3	59.9%

ALL 23 IDEAS

Rank	Idea	Generation	Score
1	Develop an experiential learning system for public-private partnerships that will match undergraduate students with teams of industry and/or community practitioners and university faculty to find data-driven solutions to real world problems.	3	63.3%
2	Incorporate real-world issues into all computer science and data science programs. All STEM areas develop ways to introduce the role of data in all,fields and societally relevant problems.	1	61%
3	Build an infrastructure in academia and industry that rewards integration of ag systems, that keeps ownership as open as possible, encourages and rewards data sharing and patents and IP, reward solving problems instead of increasing profits	3	59.9%
4	A new type of AgSchool graduate / program is needed, one that combines agro-ecosystem knowledge with analytical skills in areas such as: machine learning and data mining; statistical and quantitative Analysis; data visualization and problem solving.	1	59.5%
5	Create workshops that bring together data scientists (e.g. Math, CS, and Stats grad students, postdocs, and professors) and biologists (e.g. grad students, postdocs, and professors from ag-related departments) with real-problem datasets to analyze.	2	57.3%
6	Request all proposals in the field of big data development and application to have a significant extension and outreach component. This increases public awareness and promotes grass-root activities toward incorporating big data in teaching curricula.	2	56.9%
7	Engage biology high school teachers & students in data sciences using free, online resources (e.g., NCBI, UniProt) to introduce fundamental principles of data science at an early stage.	1	55.6%
8	Create an endless supply of experiential data learning activities that are personalized for agriculture workers, e.g. using computational creativity, and delivered online, in the classroom, or through extension.	1	54.4%
9	Adjust the metrics for university tenure and promotion to expand beyond just grant funding and publishing, as in business, measure the contribution to a team, product, impact, outcome, whether extension or research	1	53.1%
10	Develop an experiential learning model that utilizes professional societies that represent a diversity of fields as cross-sector conduits, which first engages newer members and then partners with science-based organizations (COEs) to train them.	1	52.6%
11	Develop materials for agriculture extension that are focused on data analytics, in consultation with electrical and computer engineering programs, that have foots in both hardcore engineering and hardcore data science.	1	51.4%

Rank	Idea	Generation	Score
12	Diversify opportunites for young people especially those most likely to be under-engaged because of their circumstances to generate/use data to solve local problems that are meaningful to them while at same time connecting to global issues.	1	50.7%
13	Engage biology undergraduates in data management courses & projects, in addition to more traditional bio-informatics & computer science courses.	2	50.6%
14	Better integrate data management best practices as well as ethics and responsible research training into computer and data science programs. Case studies using real data would be perfect.	1	50.2%
15	Show primary and secondary school students how data matters to them, their families and communities; demo job availability in data systems building and application, and ensure they understand and access the pathways into them.	1	48.8%
16	show students in middle school and high school that data acquisition, sharing and management is exciting and useful and that there will be jobs available in systems building and application, then make it happen at universities and companies	2	47%
17	Must consider the role of library & information sciences schools along with role of engineering and math programs in computer science and data, and engage the librarians, they've been experts in information for a very long time!	1	46.8%
18	Develop and share materials (e.g. software and datasets) that can help instructors incorporate computation into life-science courses.	2	45.5%
19	Create systems that can be navigated and used by a wide audience (i.e. UX design). This will help K-12 school teachers integrate open data sources into their curriculum.	1	45.3%
20	Provide easier access to computer and data science courses in K-12. We need a populace at every level that is literate in methods of data use and re-use.	2	45.2%
21	Must address the various facets of risk vs incentive, recognition and reward for data sharing for different types of data (e.g., Experimental vs Funder Program) by different types data producers (e.g. Researchers vs Private Sector).	1	43.2%
22	Get young people---especially our most under---engaged youth- involved in generating and analying data to solve problems that are meaningful to them.	3	40.9%
23	Imagine a farmerâ€™s young son. Heâ€™s a computer whiz and loves to tinker. Dadâ€™s $300,000 combine keeps breaking down. Can Jr try to fix it without violating a software license that neither of them has read? Don't let IP law stifle Jr's ingenuity.	2	29%

United States Department of Agriculture | National Institute of Food and Agriculture

What are the most promising opportunities for big data in communications, property rights, and communities?

2016-11-01

40	21	10	494
Contributors	Ideas	Edits	Votes

TOP 3 IDEAS

Rank	Idea	Generation	Score
1	Support training and online educational opportunities in data mining and data visualization for scientists and researchers	1	68%
2	There is a need to work on inconsistencies regarding standards, data quality, security, policy, data privacy, confidentiality, and ethics regarding the retaining and sharing/transfer of data	1	64.2%
3	Big data is only useful if it is accessible. It needs to be formatted in a readily usable form. Simple, straightforward interfaces are needed with GIS and informatics analyses. Public access should be through apps and portals with county level data.	2	58%

ALL 20 IDEAS

Rank	Idea	Generation	Score
1	Support training and online educational opportunities in data mining and data visualization for scientists and researchers	1	68%
2	There is a need to work on inconsistencies regarding standards, data quality, security, policy, data privacy, confidentiality, and ethics regarding the retaining and sharing/transfer of data	1	64.2%
3	Big data is only useful if it is accessible. It needs to be formatted in a readily usable form. Simple, straightforward interfaces are needed with GIS and informatics analyses. Public access should be through apps and portals with county level data.	2	58%
4	Encourage interaction with 'big data' by presenting it in visually appealing, interactive, formats addressing questions of general interest and/or questions that might be included in training. Ensure educators know of the resource.	2	57.1%
5	Access to local data and information through public systems can increase individual stakeholder awareness of land use, climate patterns, water use, or other common topics. Access can help to break down knowledge barriers and serve communities.	2	56.6%
6	Working with communities in developing applications, information and planning from data will show the benefits to individuals and communities while building trust and confidence in the use of data while respecting property rights.	2	53.2%
7	Precision agriculture can increase efficiency to minimize environmental impacts and maximize farm profits. Integration of GIS, climate data, and biological rates (photosynthesis, water use, etc) are needed on a single, interoperable platform.	1	51.6%
8	We talk about big data in relation to scientist, consumers, extension, funding and other aspects. However, there needs to be some clarity in understanding what is this big data for productive ideas to come forward and develop directions on its use.	1	51.3%
9	Big Data can provide real-time insights on community needs and issues allowing timely response by leaders and individuals. These big data sets can also offer confidence to projections towards marketing, performance and many more future activities.	1	51%
10	Support symposia, workshops, meetings etc. to facilitate transdisciplinary collaborations among researchers, data scientists, computer engineers etc.	1	50.3%
11	Need for support from Funding agencies to establish Bigdata infrastructure (e.g. NIH BD2K for data submissions, data sharing repositories) which can facilitate a well-defined computational architecture to facilitate large scale data analysis	1	49.6%
12	Data is a key to community outreach and communication. Communication requires effort and time. Resources should designated for extension and teaching for authentic communication about the positive contributions agriculture provides to society.	5	48.6%

Rank	Idea	Generation	Score
13	Technologies already exist to provide layered access to GIS information that anyone can access. This is low-hanging fruit for communities to connect people with data and could include agricultural production data and environmental constraints.	1	47%
14	Must address the various facets of risk vs incentive, recognition and reward for data sharing for different types of data (e.g., Experimental vs Funder Program) by different types data producers (e.g. Researchers vs Private Sector).	1	46.8%
15	Same thread of communication to achieve transparency. We can use 'big data' to back up conversations to society on agriculture based on evidence, not emotion or misguided philosophy. This should emanate from the scientific community.	2	46.2%
16	Big data can provide information that helps communities, but only after assessment. This information needs to be communicated through university extension and government research working together. Knowledge gained from big data may protect property.	2	44.8%
17	If biological challenges are formulated into mathematical models and competitions with prizes, then mathematicians will be attracted to contribute efficient algorithms for solving the problems without having to learn domain knowledge.	1	41.9%
18	Analysis and assessment of big data requires specialty computation knowledge, expensive computational resources and storage (TBs), and scientific understanding. This is an integral to delivering knowledge from big data to answer community questions.	1	40.9%
19	We need to better understand the importance of diversity. Ag/food systems are inherently diverse, yet their management is mostly single-focused. Technology is helping to capture diversity, which is critical for sustaining communities at all levels.	1	36.6%
20	It is better to adopt computing infrastructure similar to big corporations (such as Microsoft, Apple, Oracle, Google, Amazon, Facebook and twitter) for data mining and machine learning applications to inform decision-support	1	36%