



The Snake River Geothermal Consortium

is a research partnership focused on

advancing geothermal energy, hosted

by Idaho National Laboratory.

Data Dissemination and Intellectual Property Plan

April 2016



DISCLAIMER

This information was prepared as an account of work sponsored by an agency of the U.S. Government. Neither the U.S. Government nor any agency thereof, nor any of their employees, makes any warranty, expressed or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness, of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. References herein to any specific commercial product, process, or service by trade name, trade mark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the U.S. Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the U.S. Government or any agency thereof.

Data Dissemination and Intellectual Property Plan

**Jon Weers, National Renewable Energy Laboratory
Robert K. Podgorney, Idaho National Laboratory**

April 2016

**Snake River Geothermal Consortium
Hosted by Idaho National Laboratory
Idaho Falls, Idaho**

www.snakerivergeothermal.org

**Prepared for the
U.S. Department of Energy
Office of Energy Efficiency and Renewable Energy
Under DOE Idaho Operations Office
Contract DE-AC07-05ID14517**

EXECUTIVE SUMMARY

This plan details the methods and approaches that will be used to provide data generated by the Frontier Observatory for Research in Geothermal Energy (FORGE) in a transparent and easy-to-access manner. FORGE is the U.S. Department of Energy’s largest effort to advance the deployment of enhanced geothermal systems (EGS) as a viable part of the nation’s energy portfolio. This project is being conducted by the Snake River Geothermal Consortium (SRGC) at the 110-km² (42.6-mi²) Geothermal Resource Research Area on the Idaho National Laboratory Site.

SRGC, which is conducting the FORGE project, will provide easy access to these data to advance the adoption of EGS technologies, improve the understanding of the Eastern Snake River Plain’s deep geothermal system, fuel innovation in the geothermal sector, engage stakeholders, describe the fundamentals of EGS technology, provide transparency in FORGE operations, and inform local communities of FORGE activities.

In addition to the methods and approaches that will be used provide access to FORGE data, this plan describes the existing platforms, tools, and expertise that will be leveraged to construct a data management platform that is fully compatible with the Geothermal Data Repository (GDR) and National Geothermal Data System (NGDS) in a cost-effective and timely manner.

The SRGC data repository will be housed in the National Renewable Energy Laboratory’s new secure Amazon cloud environment. This modern environment has already undergone a rigorous approval and testing process, has been fully vetted by U.S. Department of Energy cyber security, and has been granted authority to operate with moderate, sensitive data in “the cloud.” This secure platform will be the basis for a repository that will leverage the existing GDR architecture to ensure full compatibility with the GDR and the NGDS, while providing secure access for FORGE partners from anywhere in the world to collaborate on sensitive data before they are released to the public.

Finally, this plan details the SRGC’s intellectual property management strategy, which is based on previous successfully implemented plans for collaborative projects, involving multiple national laboratories, universities, and industry partners (for example, the National Advanced Biofuels Consortium). Both the National Renewable Energy Laboratory and Idaho National Laboratory have vetted the resulting plan, ensuring that FORGE participants will be free to share data and innovate in a supportive and collaborative environment while maintaining their respective interests.

CONTENTS

EXECUTIVE SUMMARY	iii
ACRONYMS	vii
1. OVERVIEW	1
2. TYPES OF DATA TO BE PRODUCED	2
2.1 Continuous/Time-Series Data	4
2.2 Tabulated Data, Spreadsheets, and Databases	4
2.3 Geospatial Data	5
2.4 Three-Dimensional Data	5
2.5 High-Resolution Imagery	6
3. USE OF STANDARDS AND MODELS	6
3.1 Data Models	6
3.2 Metadata Models	7
4. DISSEMINATION	7
4.1 Access and Sharing	8
4.2 Default Licenses	9
4.3 Restricted-Use Data	10
5. ARCHIVING, PRESERVATION, AND PROVENANCE	10
6. CURATION	11
6.1 Definition of Curation	11
6.2 Curation Process	11
6.3 Identifying Curators	12
7. INTELLECTUAL PROPERTY PLAN	12
REFERENCES	14
Appendix A—Sample IP Provisions Agreement	15

FIGURES

Figure 1. Propagation of FORGE data through GDR to external sites like NGDS and Data.gov.	3
Figure 2. An example of an interactive data dashboard showing example data streams from FORGE. SRGC’s vision for the data dashboard includes both public and site management team (SMT) only views.	8
Figure 3. Example of a map-based data interface.	9
Figure 4. Data Backup Strategy.	11

TABLES

Table 1. EGS development functional stages and FORGE phases.	2
Table 2. Examples of continuous, time-series data to be collected.....	4
Table 3. Examples of anticipated tabular data.	5
Table 4. Examples of anticipated geospatial data.	5
Table 5. Examples of 3D data.....	6
Table 6. Examples of high-resolution imagery data.	6

ACRONYMS

3D	three dimensional
API	application program interface
DOE	U.S. Department of Energy
DOI	digital object identifier
EGS	enhanced geothermal systems
FORGE	Frontier Observatory for Research in Geothermal Energy
GDR	Geothermal Data Repository
GSA	General Services Administration
IP	intellectual property
NDA	nondisclosure agreement
NGDS	National Geothermal Data System
NREL	National Renewable Energy Laboratory
OpenEI	Open Energy Information
OSTI	Office of Science and Technical Information
SRGC	Snake River Geothermal Consortium
URI	uniform resource identifier

Data Dissemination and Intellectual Property Plan

1. OVERVIEW

This plan describes the data-collection, management, and dissemination solution that will provide open, transparent access to all data generated by the Frontier Observatory for Research in Geothermal Energy (FORGE). FORGE marks the U.S. Department of Energy's (DOE's) largest effort to advance the deployment of enhanced geothermal systems (EGS). These systems have the potential to tap into a conservatively estimated 100 GW of baseload power-generating capacity by harnessing the earth's heat through engineered geothermal reservoirs. The FORGE project aims to develop methodologies and technologies that will bring this resource into the nation's energy portfolio (Metcalf, 2015). This project is being performed by the Snake River Geothermal Consortium (SRGC) at the 110-km² (42.6-mi²) Geothermal Resource Research Area on the Idaho National Laboratory Site.

The SRGC will develop a comprehensive, state-of-the-art instrumentation, data-collection, management, and dissemination solution that will make all data generated by FORGE available in a transparent and easy-to-access manner. By building transparency and ease of data access into the project from inception, we can enable EGS development through the timely dissemination of the latest scientific and technical information, instill confidence in industry, and build trust with the public, all of which are critical to the success of EGS. At the core of this plan, a robust data repository, built leveraging the code of DOE's Geothermal Data Repository (GDR) and the expertise of the National Renewable Energy Laboratory (NREL) team that manages the GDR, will house and process data generated from the FORGE project, as well as metadata. This SRGC repository will use the same metadata schema developed for the GDR, allowing FORGE data submissions to be automatically added to the GDR and taking full advantage of the data dissemination pathways already built into the GDR. Our data repository will be fully integrated into our website (snakerivergeothermal.org), which will also provide access to relevant external media, upcoming events, frequently asked questions, and community support information. Data generated elsewhere but relevant to FORGE, such as key reports, publications, and external media, will also be disseminated through our website.

In addition to being a fully functional node on the National Geothermal Data System (NGDS), our repository will automatically supply metadata records for all data generated by the FORGE project to the GDR using application program interfaces (APIs). This will enable all FORGE data stored in our repository to be accessible through our website, through the GDR, and through every other federated site. Federating the metadata catalog to numerous sites, including Data.gov and the DOE Data Explorer, will dramatically increase the outreach and exposure of FORGE data without duplicating data or incurring additional storage costs (Weers and Anderson, 2015). Whenever possible, data generated will be stored in industry-approved content models to further improve the usability of the data.

The repository will also be integrated with DOE's Office of Science and Technical Information (OSTI) DataCite catalog to facilitate the assignment of digital object identifiers (DOIs) to each data set, further enabling the use of FORGE data for scientific research and helping to ensure that the data are properly attributed to their sources. This approach to data provenance, with a focus on transparency, will help to quickly build trust in both the scientific and local communities while making sure that FORGE data are available to a wide audience of researchers, students, decision-makers, and industry professionals.

The repository will enable multiple levels of access appropriate for different groups of users and stages of data release. Publicly available data sets will be automatically federated to DOE's catalog on OpenEI (Open Energy Information), the NGDS, and Data.gov. Live, streaming data will be accessible through APIs on our website, with periodic snapshots automatically assembled and added to the GDR catalog. The website will also feature a data dashboard for quick visual dissemination of key data streams and a map-based interface for live production data, providing full transparency of production operations.

Proprietary data, or any other data not deemed suitable for immediate public release, will be made available to select SRGC members, including DOE, through a tightly controlled, state-of-the-art data access system, for a period of up to 2 years starting from the time the data are generated, after which they will be made available to the public unless other specific, pre-negotiated arrangements have been made. All data generated by the FORGE project will eventually be publicly accessible.

2. TYPES OF DATA TO BE PRODUCED

The FORGE project will generate a wide variety of data, all of which will be available through our website and through the GDR by federating the metadata of SRGC data sets through existing GDR APIs.

Data will be collected from a variety of instrumented devices and activities during all phases of EGS development and FORGE maturity, following generally along the functional phases of “characterize”, “create”, and “operate” discussed by Ziagos et al. (2013). Data will also originate from analyses, models, and simulations of the area, the proposed reservoir, and its geologic properties, as well as research projects done on the FORGE site. Table 1 identifies how the EGS functional stages overlap with the FORGE phases.

Table 1. EGS development functional stages and FORGE phases.

FORGE Phase	EGS Functional Stage (Phillips et al., 2013)	Notes
1	Pre-characterize Characterize	Pre-characterize not identified by Phillips et al. (2013). Do the planning and preparation for FORGE site establishment.
2A	Characterize	—
2B	Characterize Create	Data from research done at the FORGE site and supporting locations. In this instance, “create” refers to creating the site for FORGE, not creating a reservoir.
2C	Characterize Create	Data from research done at the FORGE site and supporting locations. In this instance, “create” refers to creating the site for FORGE, not creating a reservoir.
3	Characterize Create Operate	Data from research done at the FORGE site and supporting locations. At this point, the site has been “created,” and now we focus on reservoir creation.

The collecting and sharing of FORGE data are essential to advancing EGS research, accelerating discovery, and providing full transparency of the EGS creation process. After all, “DOE’s success should be measured not when a project is completed or an experiment concluded, but when scientific and technical information is disseminated” (DOE, 2011).

The FORGE project will generate high-quality, curated data. Curators working for the SRGC will work closely with the GDR curation and development teams to ensure that FORGE metadata meet all of the requirements for GDR integration, including the metadata requirements necessary to propagate successfully to the NGDS, Data.gov, and OSTI, without adversely impacting the scope of the GDR project. This will be accomplished through the use of a common, shared code base and by leveraging the shared expertise of the SRGC and GDR teams. The SRGC repository will also employ the use of NGDS content models for all relevant data types. Additionally, the SRGC curator(s) will work closely with NREL’s geographic information system (GIS) team to ensure that any geospatial data generated meet the minimum requirements necessary to appear in NGDS map searches and to be loaded as layers in the Geothermal Prospectors.

Though all data generated by the FORGE project will eventually become publicly accessible, some data may initially be proprietary. Proprietary data will be held in our repository until a predetermined release date not more than 2 years from the date of submission, at which time such data will become publicly accessible. During this moratorium period, proprietary data will be accessible by users explicitly granted access for the purposes of research, curation, and clarification. The repository will enable multiple levels of access appropriate for different groups of users and stages of data release. For example, select individuals from collaborating universities or national laboratories, as well as SRGC curators and DOE personnel, could access proprietary data before they are available to the public.

Our repository will be home to data of all types, including site characterization and monitoring data, research and development project data, analytical models, and real-time instrumentation data. To affordably house such a variety of data, the repository will employ a combination of storage techniques, including the ability to directly store conventional data, link to externally hosted data, provide access to real-time streaming data, and integrate with high-performance compute solutions for the storage and retrieval of “big data.” The proper metadata for all of these storage solutions will be provided to the GDR for federation to the NGDS, Data.gov, OSTI’s DataCite index, and more (Figure 1), so that users of all of these websites will have direct access to FORGE data regardless of the storage solution implemented. The primary metadata repository, along with the SRGC website and the majority of data submission, will be housed on the Amazon cloud. This innovative and unique environment, built in collaboration between NREL and Amazon, has been approved by DOE’s cyber security team to house moderate-level data and has been granted official authorization to operate. This cloud-based environment will enable the SRGC repository to scale rapidly in any direction to respond to the ever-changing needs of “big data,” unanticipated spikes in user traffic, and special online events, while still providing affordable, long-term data storage and dissemination.

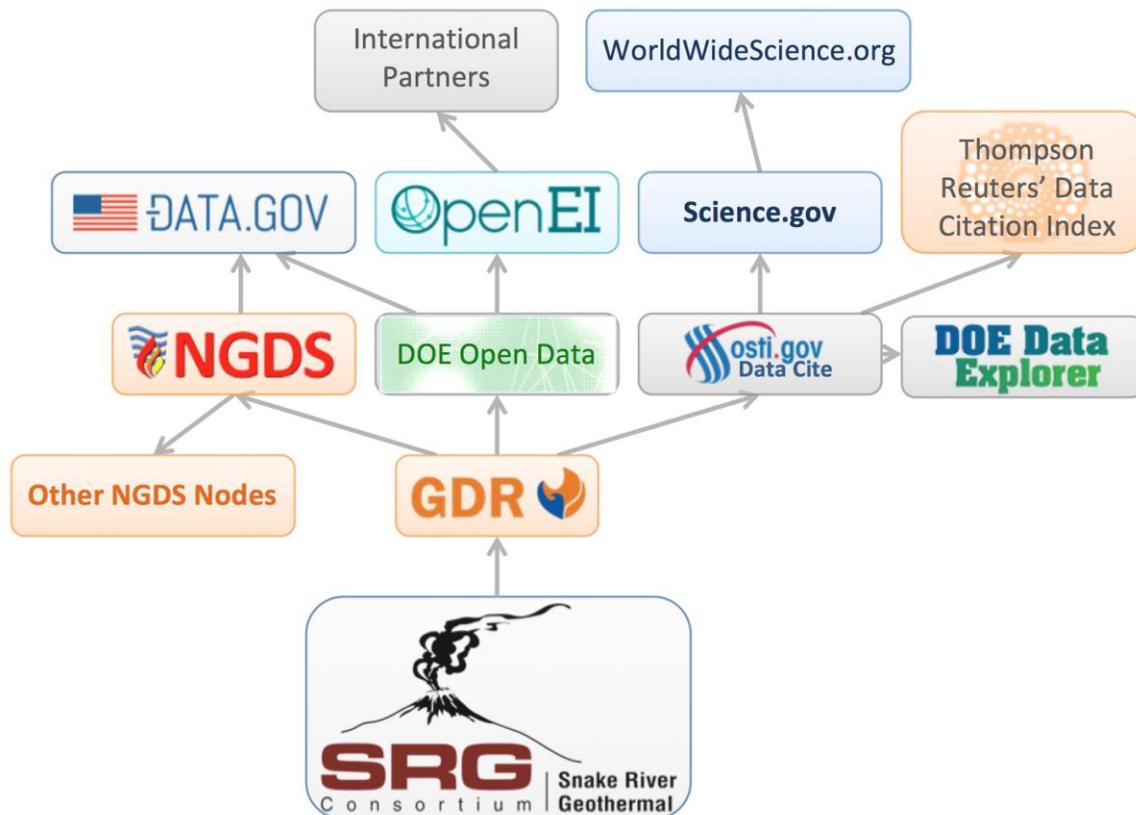


Figure 1. Propagation of FORGE data through GDR to external websites like NGDS and Data.gov.

2.1 Continuous/Time-Series Data

Instrumented devices throughout the FORGE site will provide live, streaming data from multiple aspects of development, documenting the evolution of the FORGE project and providing a transparent view of FORGE operations. These devices will provide live data on production, monitoring, seismic activity, pumps, and wells. The data will be available through a data dashboard and interactive map on our website and through APIs on the FORGE site. Snapshots of the streaming data will be taken at regular intervals and provided as data submissions to the GDR along with links to the live assets. Table 2 provides examples of the continuous, time-series data that will be collected.

Table 2. Examples of continuous, time-series data to be collected.

Instrumented Data Source(s)	Data	Phase(s)	Notes
Injection Well Production Well Water supply Well	Pressure Flow Rate Temperature	Create Operate	All wells will be instrumented and monitored in real time.
Seismometers	Seismic	Characterize Create Operate	Continuous monitoring of seismic events will be available through the data dashboard.
Local Weather	Temperature Humidity Cloud Cover Precipitation Web Cam(s) Imagery	Characterize Create Operate	Weather will be tracked as soon as the FORGE site has been instrumented with weather sensors, including a web cam, to provide a live view of the site.
Generators (Phase 3, Year 5)	Generation Conversion Efficiency	Operate	All aspects of power generation will be monitored and calculated in real time.

2.2 Tabulated Data, Spreadsheets, and Databases

The FORGE project will generate a large number of tabulated data sets from a variety of sources, including site characterization, well data, engineering data, environmental data, drilling data, analytical models, and FORGE research and development project data (Table 3). Whenever possible, the appropriate NGDS content model will be used to provide these data. All tabulated FORGE data will be automatically included in the GDR catalog, enabling proper federation to the NGDS and other sites.

Table 3. Examples of anticipated tabular data.

Data Type	Phase(s)	Notes
Water Samples	Characterize Create Operate	Water samples will be taken at routine intervals.
Core Analysis	Characterize	All information will be archived.
Research Project Data	Characterize Create Operate	Data will be collected from research done at the FORGE facility.
Analytical Models	Characterize Create Operate	Models will be used to periodically update understanding of subsurface and reservoir development.
Quarterly Reports	Characterize Create Operate	Analytics and progress data will be collected during all phases of FORGE development.

2.3 Geospatial Data

We will collect and generate numerous geospatial data sets throughout the FORGE project life cycle, including seismic data, gravity survey data, resistivity, magnetotelluric data, and other map-based data. Whenever possible, geospatial data will be provided in the appropriate NGDS content model, enabling the data to be included in NGDS map-based searches and integrated into Geothermal Prospector (Table 4).

Table 4. Examples of anticipated geospatial data.

Data	Phase(s)	Notes
Seismic Activity Maps	Characterize Create Operate	Geospatially tagged seismic hypocenters will be calculated throughout the FORGE project.
Gravity Survey Data	Characterize Potentially Operate	—
Resistivity Data	Characterize Potentially Operate	—
Magnetotelluric Data	Characterize Potentially Operate	—
Two-Dimensional Models	Characterize Create Operate	Models of the FORGE site will be continuously updated with the latest information throughout the various project phases.

2.4 Three-Dimensional Data

We will incorporate the use of three-dimensional (3D) models throughout the project. These detailed data sets will provide FORGE users and the geothermal community with the data needed for additional simulation, well design, and other comprehensive analysis. Table 5 provides examples of the 3D data we will gather during the FORGE project.

Table 5. Examples of 3D data.

Data	Phase(s)	Notes
Geologic Models	Characterize Create Operate	A geologic model will be updated throughout the development of the FORGE site.
EGS Models	Characterize Create Operate	Accurate models of EGS will be kept up to date with the latest information.
Simulations	Characterize Create Operate	Fluid flow, thermal transport, and geomechanical evolution will be simulated.

2.5 High-Resolution Imagery

Additional information will be provided as high-resolution imagery in the form of detailed site maps, digitized core samples, and video from recorded presentations, downhole cameras, and a live web cam of the FORGE site (Table 6).

Table 6. Examples of high-resolution imagery data.

Data	Phase(s)	Notes
Site Maps	Characterize Create Operate	Detailed site maps will be assembled from publicly available satellite imagery and planned infrastructure.
Digitized Core Samples	Characterize	—
Downhole camera recordings	Characterize Create Operate	Video recordings of downhole cameras will be made available as downloadable data.
Recorded Presentations	Characterize Create Operate	Presentations from all phases of the FORGE project will be available for viewing on the SRGC website.
Live Web Cam Feeds	Characterize Create Operate	One or more live web cams of the FORGE site will provide real-time visibility into site activities.

3. USE OF STANDARDS AND MODELS

Our data repository will take full advantage of the work done by the U.S. Geoscience Information Network (USGIN), OSTI, and the General Services Administration’s (GSA’s) Project Open Data in creating comprehensive data through the use of data and metadata models to increase the utility and exposure of FORGE data.

3.1 Data Models

The use of data models, structured data conforming to one or more widely used standard formats, including NGDS content models, will be encouraged for all applicable FORGE data. Our repository will support the submission of data conforming to these models in a manner fully compatible with the NGDS. Built on the same code base at the GDR, the repository will be able to automatically detect the use of an

NGDS content model and modify the submission metadata appropriately to ensure full interoperability with the NGDS site. Data provided in content models with a geospatial component will be automatically loaded into the web feature service at geothermaldata.org for inclusion in map-based searches and integration into Geothermal Prospector.

3.2 Metadata Models

The metadata collected with each data set will conform to the metadata standards currently employed by OSTI, the NGDS, and GSA's Project Open Data. Using a common metadata core, our repository will have active translators for each of the required metadata standards, enabling real-time integration with high-profile U.S. Government sites, including Data.gov, OSTI's DataCite index, the DOE Data Explorer, Thompson Reuters' index, the NGDS, and OpenEI (Figure 1). Integration with these sites will make FORGE data discoverable by a broader audience.

The metadata used by the repository will be fully compatible with the GDR and the NGDS and will be exposed through the proper mechanisms to enable inclusion of FORGE data in NGDS search results.

Metadata about each data set added to the repository will also be submitted to OSTI for assignment of a DOI, increasing the utility of data within the repository and making it easier for users to cite FORGE data in their research.

Our repository will also transmit metadata conforming to GSA's Project Open Data standards to GSA and OpenEI, ensuring that FORGE data appear on Data.gov, are properly attributed to DOE, and are within the DOE open data catalog on OpenEI. The inclusion of FORGE data on Data.gov will greatly increase the exposure of FORGE data by reaching a larger community of users and by taking full advantage of the many efforts GSA has already built to improve data dissemination.

4. DISSEMINATION

Publicly accessible data will be openly available, without restriction, via the repository interface for download and through machine-readable web services (APIs) for application and analysis support. Data streaming from instrumented devices will be available through a data dashboard (Figure 2). An interactive map display (Figure 3) will feature live site data, provide full transparency of FORGE operations, and provide users with another means of interacting with FORGE data. All of these data streams will also be accessible through APIs. Links to these live data streams will be uploaded, along with regularly timed snapshots of streaming data, to our data repository, and subsequently the GDR. Our data dashboard, the interactive map, and the data repository will all be accessible through our website (snakerivergeothermal.org), which also features information on EGS, FORGE-related events, the SRGC, and the Snake River Plain.

Metadata for all data added to our repository will be made available through the repository interface, exposed through machine-readable APIs, and federated to several high-profile government sites, including the GDR, the NGDS, OpenEI, OSTI's DataCite catalog, Data.gov, and more. The additional exposure will benefit the FORGE project by reaching a wider audience and increasing the usability of FORGE data. As a result, FORGE data will be openly available, without restriction, through the search interfaces of the many sites shown in Figure 1.

Metadata for all data within our repository will also be made available to search engines by using the latest in search engine optimization strategy, with the goal of achieving top search rankings, increasing the discoverability of FORGE data.

Our repository will use an "API-first" development process, meaning it will use its own API for almost all base functions, ensuring that these functions are available for use by both the development and analytical communities. This development methodology will ensure the availability of FORGE data through machine-readable standards by prioritizing the development on public-facing APIs.

Our website will further disseminate FORGE data by showcasing select data sets through various social media channels, including blogs, YouTube, LinkedIn, and Twitter, in accordance with the targeted social media campaign outlined in the FORGE *Communications and Outreach Plan* (Ulrich and Podgorney, 2016).

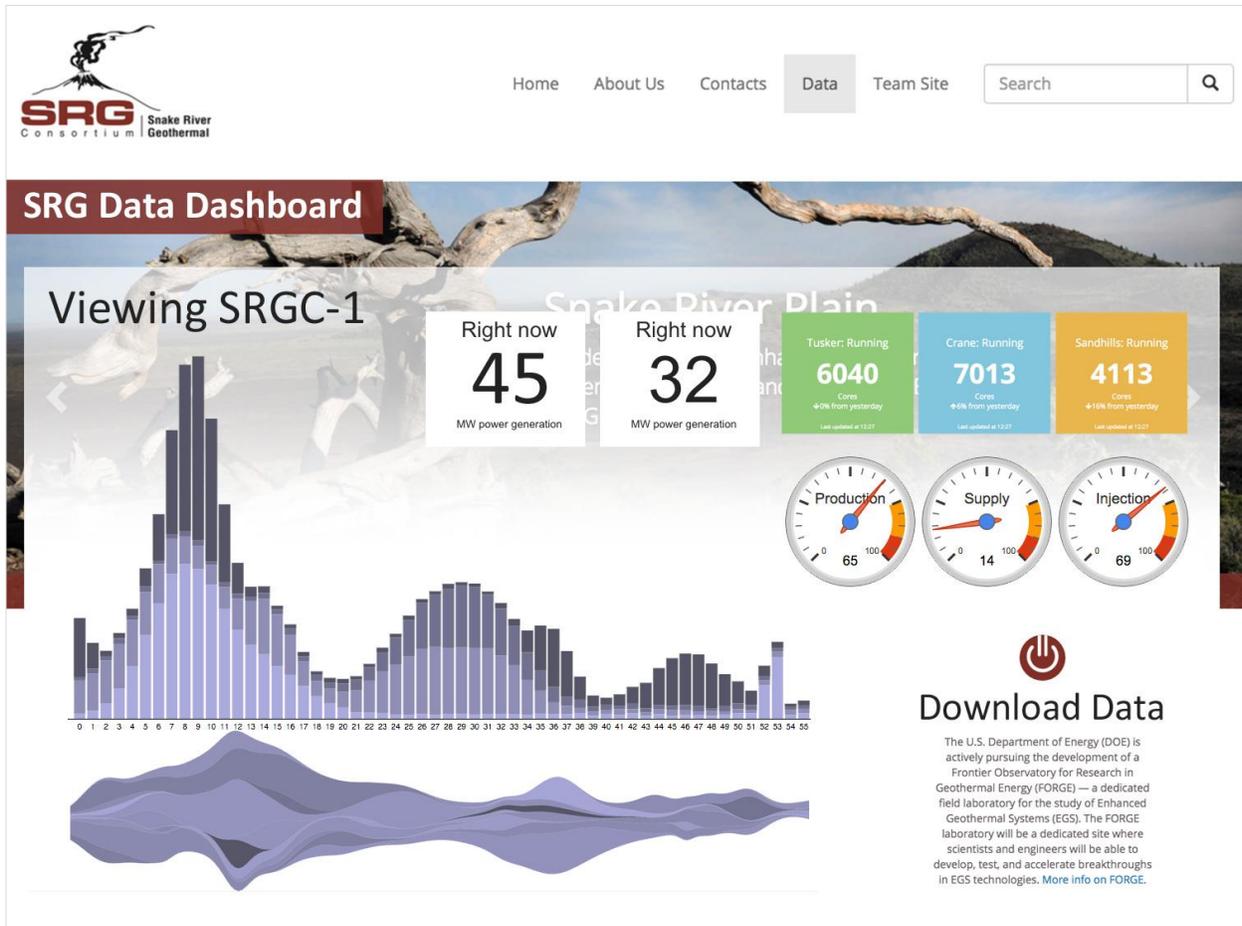


Figure 2. An example of an interactive data dashboard showing example data streams from FORGE. SRGC’s vision for the data dashboard includes both public-access and site-management-team-only views.

4.1 Access and Sharing

The majority of FORGE data will be made available to the public in near-real time. Publicly accessible data will be openly available, without restriction, for download via the application interface, through machine-readable web services, the data dashboard (Figure 2), the interactive map (Figure 3), or any of the federated partner sites (Figure 1). All users will be free to share publicly accessible data as they see fit. Relevant and timely data may also be linked to events on the SRGC website, shared through the SRGC blog (i.e., “the Director’s Message”), or social media channels in accordance with the FORGE *Communications and Outreach Plan* (Ulrich and Podgorney, 2016).

A data dashboard on our website will continuously transmit and monitor the status and health of FORGE. The dashboard will allow users to view data streams for specific instrumented devices in real time or to download aggregated data, as illustrated in Figure 2.

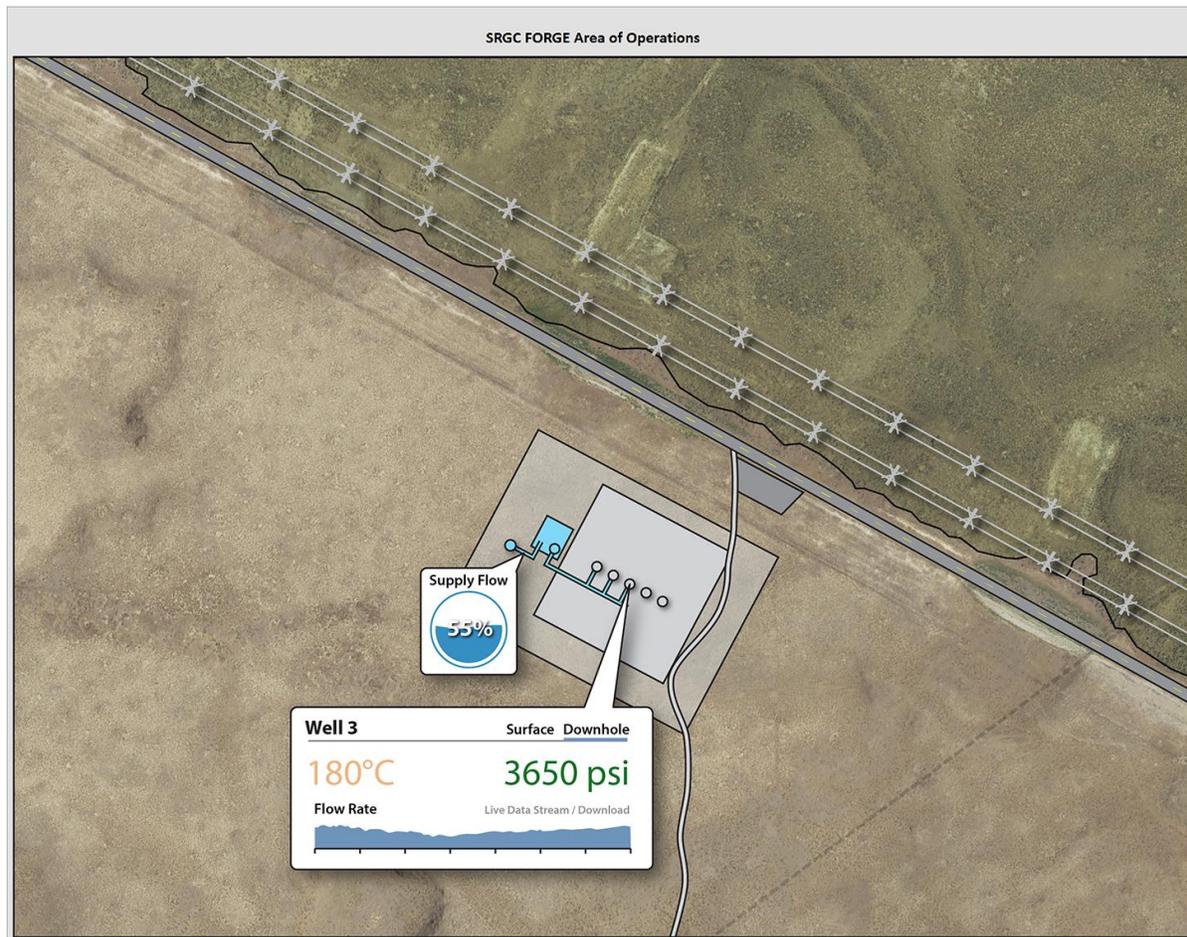


Figure 3. Example of a map-based data interface.

An interactive, map-based interface will provide another means of accessing FORGE data. Figure 3 shows a conceptual interactive map displayed as an example of what will be developed. Interactive elements in the image will allow users to simply click on a well to see data associated with it, such as wellhead or downhole pressures, temperatures, and flow rates. Data from other aspects of production that have been instrumented will also be accessible, such as the flow rate of water from the supply well, to provide a fully interactive, transparent picture of FORGE operations.

Data streaming from instrumented devices will also be aggregated into tabular data sets at regular intervals for easy download and inclusion in the GDR.

High-resolution images of core samples, developed in accordance with the FORGE *Sample and Core Curation Plan* (Snyder et al., 2016), will also be available for download through the SRGC data repository and included in the GDR.

4.2 Default Licenses

By default, FORGE data will be licensed Creative Commons Attribution (CC-BY: <http://creativecommons.org/licenses/by/4.0>), which allows unrestricted use and reuse with proper attribution to the original creating party. License and use information for all data, both public and proprietary, will be clearly visible on data set landing pages to inform users of permissible use and proper

citation. The landing pages for each data set will automatically generate their own proper citation for users to simply copy and paste into their derivative works.

4.3 Restricted-Use Data

Our repository will be able to accommodate restricted-use data, such as data that are part of a collaborative research agreement, governed by proprietary intellectual property (IP) provisions, or otherwise subject to a contractual moratorium on public release.

Seamless, cost-effective, and secure access to restricted use data will be provided through NREL's new, cloud-based secure platform, which resides on a FedRAMP-approved architecture and has already been vetted by DOE cyber security. The repository will leverage the existing GDR architecture, which has already been approved by DOE cyber security, granted an official authorization to operate, and approved for housing sensitive (moderate-level) data.

Restricted-use data will be available only to users with the explicit permission needed to access the data. This will allow SRGC members, including multiple national laboratories, universities, and DOE, to securely access restricted-use data sets before they are publicly available, enabling easy collaboration on proprietary or sensitive projects.

Behind the scenes, the repository will be split into two systems, one housing publicly accessible data and one securely housing restricted-use data. The public-facing system will act as a gateway to the secure system, providing access only to users with explicit permissions. From the user's perspective, the delineation will be invisible. Our repository will appear to be a single, seamless site with access to both open and restricted-use data through the same interface.

Metadata for restricted-use data will still be made available to the public and will include contact information as well as the date on which the data will become publicly accessible. Though unable to access restricted-use data, the general public will have the basic information necessary to make inquiries about the data and their eventual release, reducing duplication of efforts and promoting further collaboration.

5. ARCHIVING, PRESERVATION, AND PROVENANCE

The repository will store a complete revision history for all data sets. Individual collections of data will exist at a fixed uniform resource identifier (URI) to support proper data provenance, and these collections will be assigned a permanent unique identifier to enable proper citation.

The repository will be capable of storing data indefinitely and will be designed with proper data provenance in mind. Publicly available FORGE data will remain unchanged and available to the public at a fixed Internet location (URI). The assignment of DOI numbers through integration with OSTI will provide another permanent means of identifying individual data resources. These steps are essential for building trust in the scientific community and ensuring that any data cited in derivative works continue to be discoverable, increasing the validity of those works.

On a system level, the repository itself will be archived in an economical but appropriately redundant manner. At a minimum, daily archives of the system will be kept for up to a month, monthly archives will be kept for a quarter, quarterly archives for a year, and annual archives indefinitely (Figure 4).

The repository will support proper data provenance. The underlying cloud architecture, combined with the metadata schema employed, the submission process, and curation workflow will all be designed to maintain an indefinite, uncorrupted digital resource for the scientific community. All data sets submitted to our repository will be assigned unique identifiers, receive DOIs when appropriate, be properly versioned, and reside at a fixed location online (URI). This will allow them to be appropriately referenced in research papers and provide reliable and credible pointers to data residing at a specific reference.

This standard approach to archiving is considerably more economical than storing daily archives indefinitely, and it serves the vast majority of system restoration needs.

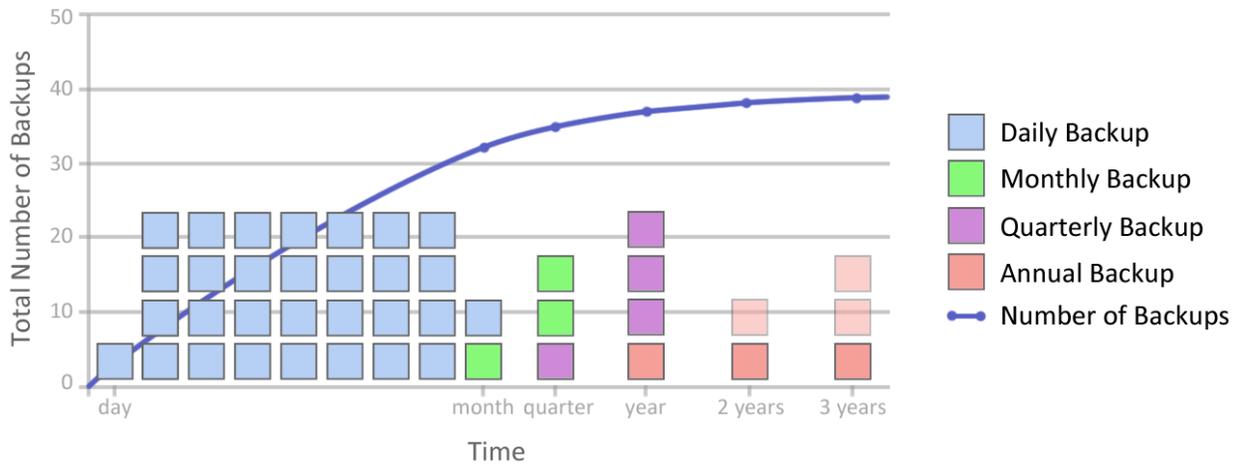


Figure 4. Data backup strategy.

6. CURATION

A data repository is only as valuable as the quality of its contents. As a result, curation of data introduced to the SRGC repository will be critical to the project. Additionally, the repository will be built in such a way that it can easily support external curators and enable the sharing of curation responsibilities among the SRGC, DOE, and potentially other parties.

6.1 Definition of Curation

Curation is commonly defined as a process or activity performed to help preserve the discoverability, relevance, and usefulness of data. Curation of our repository will include the review of data for completeness and validity. Curation will also include the review of the metadata provided with each submission for accuracy, completeness, and relevance, as well as adherence to the metadata standards of the external sites identified in Figure 1. If necessary, curators may contact data originators to resolve any discrepancies or omissions detected during this review. Curators will also be able to change the status of existing data collections, moving them toward public accessibility or removing them from public availability to address a concern with the data. Curation may, at times, include the complete removal of inappropriate or outdated data, the development of new data models, or the reclassification of existing data.

6.2 Curation Process

The curation process begins with the review of metadata provided for each new data set. The curator will be checking metadata for:

- **Context** to make sure the metadata provided accurately describe the associated data in a manner easily understood by someone not familiar with the project
- **Relevance** to ensure the metadata are descriptive of the data included
- **Appropriateness** to make sure the data are suitable for public release and do not inadvertently contain personally identifiable information or other obviously inappropriate content
- **Completeness** to ensure no critical components are missing and that the metadata do not reference any resources not present within the data set.

Additionally, the curator will examine each data resource for quality, completeness, and accuracy. Any issues discovered will be resolved with the help of the original data submitter before making the data publicly accessible. Once satisfied with the data set, the curator will mark curation of the data set complete, triggering the repository to automatically advance the data set to the next step in the workflow.

The curation process will vary slightly for different data types. Reports or analyses, for example, may require more individual scrutiny than single-source streaming data.

6.3 Identifying Curators

The SRGC governing members will select the curators necessary to interface and develop a professional relationship with data providers and who have the technical skills needed to ensure metadata compatibility with external sites. Curators will be asked to commit to at least 1 year of fulfilling their curation duties.

Curators will be able to curate remotely and can include personnel from any of the SRGC members, national laboratories, universities, or DOE-appointed individuals. To facilitate remote curation, our data repository will allow for secure access to restricted-use data sets. This same functionality could be used to grant individuals from DOE access to data before they are made publicly available.

7. INTELLECTUAL PROPERTY PLAN

IP may be developed during FORGE research by members of the SRGC or by awardees from the solicitation process, described in the *Research and Development Implementation Plan* (Podgorney et al., 2016). The general purpose of this IP plan is to address the protection and disposition of IP developed under the FORGE project, within the framework of federal IP laws, regulations, and policies. It defines the ownership, disposition, and sharing of IP. The objectives include:

- Promoting the patenting, licensing, and rapid commercialization of inventions developed under the DOE FORGE award
- Promoting the rapid dissemination of scientific data for the public good.

Although management of proprietary information and IP is similar for SRGC members and solicitation awardees, the process is divided into different agreements. An IP Management Plan (Appendix A) is the controlling document for the SRGC; solicitation awardees' IP will be managed through the procurement contracting documents.

7.1 SRGC Agreements

There are two primary documents that define the charter and operation of the SRGC. These are a multiparty nondisclosure agreement (NDA) and an IP management plan (Appendix A). Lead SRGC members are required to enter into the NDA and IP management plan prior to submission of the FORGE Phase 2 renewal application to DOE.

The NDA and IP management plan are signed by each SRGC member institution. The SRGC NDA has a nondisclosure period of 5 years after the end of Phase 3, so it remains active and in force until approximately April 2029. The IP management plan establishes a procedure for rapid dissemination of information, provides effective and coordinated efforts for commercialization, and allows for broad access to technology developments. Lead members had input to the language of the NDA and IP management plan, and it is mandatory that they sign the agreements. Members who joined the SRGC later, or who serve in more of a support role, are required to sign the NDA and IP management plan as executed by the lead members without modification. Early development of these two documents provides a basic framework inside the SRGC environment and is essential to fully developed management and IP plans.

7.2 Solicitation Awardee Agreements

IP is owned by the respective institution(s) that invented the IP, as prescribed by U.S. patent law. A nonproprietary description of the IP will be shared with DOE and SRGC members, and will be available to the general public. Solicitation awardees will enter into a contractual agreement through the INL procurement organization, which includes provisions for IP management. Provisions presented in the attached plan for SRGC members that are relevant for solicitation awardees will be included in the contract. Specifically, Sections 3, 4, and 5 will be applied in the solicitation awardee's contract. Background IP will be identified by the solicitation awardee and protected as described in Section 3.a. SRGC members will have the first option to license IP from the inventor during a 9-month option period as described in Section 3.d and 5.c. DOE maintains its statutory rights as defined in Section 3.f. A non-confidential version of IP will be disseminated as presented in Section 4 and publication of results by the inventor is expected. Data collected by the solicitation awardee will be governed as described in earlier sections of this document; nonproprietary data will be distributed to the public in real time. Proprietary data will be protected for a period of up to two years, at which time it will also be distributed.

To the extent this plan may conflict in any way with the provisions of any phase of the FORGE project award, including any IP provisions, the FORGE award provisions will take precedence. SRGC and DOE rights to IP and/or project technical data not otherwise specified in solicitation awardee agreements will be governed by the terms of the prime contract.

REFERENCES

- DOE, 2011, Strategic Plan, U.S. Department of Energy, DOE/CF-0067, p. 43.
- Metcalfe, E., 2015, Road Tripping through the Geothermal Frontier: <http://energy.gov/eere/articles/road-tripping-through-geothermal-frontier> (accessed March 2016).
- Ziagos, J., Phillips, B., Boyd, L., Jelacic, A., Stillman, G., and Hass, E., 2013, A Technology Roadmap for Strategic Development of Enhanced Geothermal Systems, Proceedings, Thirty-Eighth Workshop on Geothermal Reservoir Engineering, Stanford, California, February 11–13, 2013.
- Snyder, S., McLing, T., and Bartholomay, R., 2016, Sample and Core Curation Plan: Snake River Geothermal Consortium, INL/LTD-16-38122.
- Ulrich, J., and Podgorney, R.K., 2016, Communications and Outreach Plan: Snake River Geothermal Consortium, INL/LTD-16-38119.
- Weers, J., and Anderson, A., 2015, DOE Geothermal Data Repository: Getting More Mileage Out of Your Data, in Proceedings, 40th Workshop on Geothermal Reservoir Engineering, Stanford University, Stanford, California.

Appendix A

Snake River Geothermal Consortium Intellectual Property Management Plan

Appendix A

Snake River Geothermal Consortium Intellectual Property Management Plan

The principal goals of this Intellectual Property Management Plan (IP Management Plan) are to (i) establish procedures for the rapid dissemination of information among Snake River Geothermal Consortium (SRGC) Members, (ii) provide effective and coordinated efforts to commercialize technologies developed in SRGC and (iii) supply broad access to relevant technologies to researchers in the Core Research Area. This IP Management Plan relates solely to the SRGC.

1. Background and Award

- a. The United States Department of Energy (DOE) awarded a grant to fund Phase 1 of the SRGC Frontier Observatory for Research in Geothermal Energy (FORGE). The purpose of the SRGC is to establish and manage a dedicated Enhanced Geothermal System (EGS) field laboratory to focus on science and technology research and development in an ideal EGS environment. The resultant goal is a rigorous and reproducible methodology that will enable development of large utility scale, cost-competitive EGS power. The Phase 1 objective is to complete mission-critical technical and logistical tasks that demonstrate site viability and commitment and capability to meet FORGE goals. The Phase 2 objective is to fully instrument the site and bring FORGE to full readiness for the testing of new technologies. The Phase 3 objective is full implementation of FORGE and tasks specific to solicitation, selection, testing and evaluation of new and innovative EGS tools, techniques, and supporting science. Phases 2 and 3 have not yet been awarded, but if awarded by DOE to SRGC, this IP Management Plan will also apply during the performance of those phases.
- b. The SRGC is composed of the Members specified herein, and led by the Battelle Energy Alliance, LLC as the Management and Operating contractor of the Idaho National Laboratory (the “SRGC Lead”).
- c. Members shall work with the SRGC Lead to prepare and submit statements of work (each, a “Statement of Work”) to the SRGC Lead that set forth a proposed scope of research and development work to be conducted by such Member(s) with SRGC Program funding. If a Statement of Work is accepted and approved for a sub-grant by the Member(s) to be participating in such Statement of Work and the SRGC Lead, the respective Member(s) and the SRGC Lead shall then enter into a detailed agreement (each, a “Sub-Agreement”) setting forth the specific terms and conditions for that sub-grant. Members intend for all Statements of Work and Sub-Agreements to be prepared in accordance with the principles set forth in this IP Management Plan.
- d. This IP Management Plan is intended to further the SRGC Purpose by promoting a consistent, uniform approach to the disposition of intellectual property (IP) rights and the handling of proprietary and protected data across all Members of the SRGC to accelerate research via an open exchange of information.

2. Definitions

- a. The Consortium is comprised of the following members (each a “Member” or collectively “Members”);
- b. “SRGC Funding” means (i) the funding that was awarded by the Department of Energy (DOE) to SRGC, and/or the SRGC Lead and/or other Members (ii) the funding that was awarded by DOE to the other Federal Laboratory Members under their Prime Contracts specifically for the SRGC Program, and (iii) any funding provided by the Members as part of their cost share.
- c. “SRGC Program” means the research program conducted by and amongst the Members using SRGC funding.
- d. “Core Research Area” means enhanced geothermal systems as described in the Statement of Work approved by the DOE in granting the SRGC Funding.
- e. “Intellectual Property” means all intellectual property rights, discoveries, innovations, know-how, works of authorship and inventions, along with derivative works, whether patentable or not, including computer software and code, patents and patent applications, trade secrets, mask works, copyrights and copyrightable materials.
- f. “Background Intellectual Property” or “Background IP” means any IP that meets each of the following conditions (i) was developed before or outside the SRGC Program, (ii) is necessary for use in the Core Research Area, and (iii) is attached hereto by the express written agreement of the IP owner.
- g. “Program Intellectual Property” or “Program IP” means any IP conceived or first actually reduced to practice during and as part of the SRGC Program.
- h. “Affiliate” means, as to any person or entity, any other person or entity that directly or indirectly controls, is controlled by, or is under common control with such person or entity, and is identified in writing to the Members. For purposes of the preceding sentence, “control” means the right to control, or actual control of, the management of such other entity, whether by ownership of securities, by voting rights, by agreement or otherwise.
- i. “Inventing Member” means a Member or Members who first conceive or reduce to practice Program IP.
- j. “Federal Laboratory Members” means Battelle Energy Alliance, LLC and Alliance for Sustainable Energy, LLC.

3. Intellectual Property Ownership

- a. Operating Principles. The Inventing Members, acting reasonably and in good faith, shall have the sole discretion to seek patent or copyright protection on Program IP. Upon mutual agreement, the Inventing Members may appoint a single Member to manage and facilitate the filing, prosecution, and maintenance of patents and copyrights on Program IP. Each Member

- electing and filing any such patent application is responsible for patent costs incurred by such Member in connection with the filing and prosecution thereof, unless otherwise agreed.
- b. Background IP. Ownership of Background IP shall remain unaffected by this IP Management Plan. Members shall have the sole discretion as to whether or not to include Background IP in the SRGC Program. Members shall act reasonably and in good faith to identify Background IP that is available and necessary for conducting research under the SRGC Program. All included Background IP shall be attached, which may be amended from time-to-time. Any such amendment shall be by written notice of the contributing Member to all Members of the SRGC. Access to included Background IP by any Member shall be subject to prior written notice to, and expressed written consent of, the contributing Member. Members who have no need to access Background IP for work in the Core Research Area will not be required to receive Background IP.
 - c. Program IP. Program IP solely developed by a Member shall be owned by that Member. Program IP developed by two or more Members shall be jointly owned by the Inventing Members. Individual Members creating joint inventions may enter into a separate agreement to facilitate the filing of patent applications and licensing.
 - d. Protecting Program IP. Non-Federal Laboratory Inventing Members shall have 90 days from receipt of a complete invention disclosure by their respective Intellectual Property department (or equivalent) to file the necessary applications to protect the Program IP if the Inventing Member elects to do so. Federal Laboratory Inventing Members shall have the time specified in their respective Prime Contracts to file the necessary applications to protect the Program IP if the Federal Laboratory elects to do so. Upon such filing, the Inventing Member(s) shall send a copy of the patent applications or other full description of the Program IP to the SRGC Lead. Within 10 business days, the SRGC Lead will disseminate copies of such patent applications or other full descriptions covering such Program IP to all Members who have requested such copies, initiating the nine (9) month Option Period as described in Section 5(c), below. Such dissemination will be subject to the SRGC NDA unless the disclosing Member expressly provides otherwise by written documentation.
 - e. Governing Law. Ownership of any Program IP shall be determined in accordance with applicable state and federal law and as required by Federal Laboratory Members' prime contracts with U.S. Department of Energy (DOE) for the management and operation of the Federal Laboratory and as prescribed in any subcontract between the SRGC Lead and non-Federal Laboratory Members. Any assignment of rights between an inventor, author or other individual involved in the creation of Program IP and the Member who sponsors or employs him or her shall be governed by the law and policies applicable to the sponsoring Member. The Bayh-Dole Act, 35 U.S.C. 200 et seq., requires that universities, non-profit institutions and small businesses who are participating under a funding agreement (as defined in the Bayh-Dole Act) will have the option to retain title to their own employees' inventions. Inventions made by employees of Battelle Energy Alliance, LLC and Alliance for Sustainable Energy, LLC will be subject to their respective Management and Operating (M&O) contract terms and conditions with respect to ownership of inventions made by their employees. The M&O contract generally provides that the laboratory contractor has the right to elect to retain title to inventions made by their lab employees. Federal Laboratory Members' option and intellectual property provisions will be in accordance with their DOE Prime Contract requirements and DOE approved agreements to conduct work for non-federal

sponsors such as Cooperative Research and Development Agreements (CRADAs) or Work for Others (WFO).

- f. Reserved Federal Rights. The Members acknowledge that the U.S. Government will have certain statutory rights in any invention that is conceived or first actually reduced to practice under a DOE award. In accordance with 42 U.S.C. 5908, title to such inventions vest in the United States, except where 35 U.S.C. 202 provides otherwise for nonprofit organizations or small business firms. However, the Secretary of Energy may waive all or any part of the rights of the United States subject to certain conditions. Any and all rights afforded to the U.S. Government shall be consistent with the terms and conditions as provided by the DOE for the SRGC Funding.

4. Dissemination and Publication

- a. Dissemination of Research Findings. Inventing Members shall provide a non-confidential title and abstract of any newly created Program IP to the SRGC Lead within 30 days of receipt of a complete invention disclosure by their respective Intellectual Property Offices. The SRGC Lead shall notify all Members of the existence of new Program IP by disseminating the non-confidential title and abstract of Program IP to all Members. Members wishing to receive copies of such Program IP, when available, will formally submit a written request to the SRGC Lead.
- b. Publication. Members shall publish the findings and results of research and development activities under the SRGC Program, provided that they have a reasonable opportunity to seek any patent or other protection for its Program IP. Members agree to share their research findings resulting from the SRGC Program within the general research community through annual meetings, professional conferences, and publications.

5. Intellectual Property Licensing

- a. Background IP. To the extent that a Member has the legal right, title, and interest to do so, and subject to third party rights, and solely if such Member elects in its discretion to include its Background IP in the SRGC Program, such Member shall grant to any requesting Member a fully paid up, royalty-free, nonexclusive, nontransferable license under its Background IP, for a defined period of time, solely to use said Background IP for internal, noncommercial purposes in the Core Research Area for the purpose of the requesting Member to conduct activities pursuant to such requesting Member's Statement of Work under the SRGC Program, so long as the licensing Member is a Member of the SRGC. Except for such limited license, nothing in this IP Management Plan shall constitute a grant, option, or license to commercialize or otherwise use another Member's Background IP for any other purpose. Subject to such license, Members shall retain the right, in their sole discretion, to license their Background IP to any party on mutually agreed terms. Member Background IP can be withdrawn from this IP Management Plan at the discretion of the Background IP owner, subject to any licenses granted to Members and their Affiliates, as applicable.
- b. Program IP—Internal Research Use. Each Inventing Member of Program IP shall grant a royalty-free, nonexclusive, nontransferable license under its Program IP to other Members solely for internal, noncommercial research use as required for such Member to conduct its activities pursuant to any Statement of Work involving such Member(s) under SRGC Program so long as the Inventing Member is a Member of the SRGC.

- c. Program IP—Commercial Use. For a nine-month period following the date that each Program IP is disseminated to all Members by the SRGC Lead as provided in Section 3(d) above (the “Option Period”), each non-Inventing Members shall have the right to negotiate with the Inventing Member, in good faith under reasonable terms and conditions, a nonexclusive commercial license to such Program IP, and such license(s) must be fully negotiated and executed within this Option Period. Any such commercial use license which may include the right to sublicense, will be at the discretion of the Inventing Member, for commercial use of such Program IP. Inventing Members may license their Program IP to non-Member third parties for commercial purposes after the expiration of the Option Period and subject to any licenses granted to Members and their Affiliates, as applicable.
- d. Program IP—Research Community. After the Option Period, each Inventing Member is encouraged to implement and market in good faith a nonexclusive licensing program for the general research community with respect to the Program IP. Notwithstanding the foregoing, the decision to license an Inventing Member’s Program IP to the general research community, as well as the specific terms and conditions under which it shall do so, is within the sole discretion of the Inventing Member. To the extent that a Member has the legal right, title and interest to do so, and subject to third party rights, such Program IP will be made available to academic researchers for research and noncommercial purposes in the U.S. for free.

6. Addition of New Members

The Members have separately agreed to adhere to a SRGC Non-Disclosure Agreement (the “NDA”). New Members will sign the NDA when they join the SRGC and sign this IP Management Plan. Requiring every existing Member to re-sign the NDA and IP Management Plan every time a new Member is added to the SRGC is cumbersome. Therefore, the SRGC Lead will facilitate the addition of new Members as follows: When a new party is proposed to be added as a Member, the SRGC Lead will transmit an abstract via fax or electronic mail regarding the proposed new party, including their proposed Statement of Work, to all current Members for a comment and consent period not to exceed thirty (30) days from the date the abstract is transmitted. The SRGC Lead will collate and disseminate to all Members all comments regarding the addition. All Members will make a good-faith effort to resolve any issues regarding the addition, and will not unreasonably withhold their consent to the addition. If the Members are unable to reach unanimous consent, the SRGC Lead will send the abstract together with all comments regarding the proposed addition to the cognizant DOE program official, who will make a final decision regarding the proposed new party’s addition. Upon unanimous consent of the Members or final decision by DOE, the SRGC Lead will provide the NDA and IP Management Plan for the new party’s signature, after which the SRGC Lead will send an amended NDA and IP Management Plan reflecting the addition to all Members, and the amended NDA and IP Management Plan will be binding on all the Members.

7. Miscellaneous

- a. Nothing herein contained shall constitute a partnership between or joint venture by the Members hereto or constitute any party the agent of the others. No Member shall hold itself out contrary to the terms of this Section and no Member shall become liable by any representation, act or omission of the other contrary to the provisions hereof. This IP Management Plan is not for the benefit of any third party and shall not be deemed to give any right or remedy to any such party whether referred to herein or not.
- b. This IP Management Plan may not be modified except in writing signed by all Members.

- c. If any term of this IP Management Plan is found by a court of competent jurisdiction to be unenforceable, the remaining terms of this IP Management Plan will remain enforceable.
- d. This IP Management Plan expresses the entire understanding of the Members with respect to the subject matter it addresses and supersedes any prior written or oral agreement or discussions. The Members acknowledge that they are not relying on any representation, promise, or other statement, whether written or oral, that is not expressly contained in this IP Management Plan. No Member is obligated to disclose any Proprietary Information. This IP Management Plan does not obligate any Member to enter into any further agreement with the other Member.
- e. All notices, requests or other communications to any Member hereunder shall be in writing (including telecopy, electronic mail or similar writing) and shall be given to each Member at its address set forth below its signature or such other address, facsimile number or electronic mail address as such Member may hereafter specify by notice to the other Members. Each such notice, request or other communication shall be effective (a) if given by facsimile, when such facsimile is transmitted to the facsimile number specified below and the appropriate facsimile confirmation is received or (b) if given by any other means, when delivered at the address specified below.

[*NOTE: Members are not listed nor are the signature pages not included in Data Dissemination and Intellectual Property Plan attachment.]