

DOE Bioenergy Technologies Office (BETO) 2023 Project Peer Review

6.3.0.2 – Bioenergy Knowledge Discovery Framework (KDF)

April 5, 2023

Data, Modeling, and Analysis Program

Dr. Esther Parish

Oak Ridge National Laboratory (ORNL)

ORNL is managed by UT-Battelle, LLC for the US Department of Energy

Project Overview

- Develop and maintain a “go-to” web-based repository of BETO-funded datasets, reports, and publications.
- Target researchers proposing new projects needed to accelerate decarbonization and development of the US bioeconomy.
- In 2011 and 2016, KDF was fundamental in the distribution of Billion Ton Report datasets and visualizations.
- A new database and dashboard are being prepared for the next Billion Ton report (BT23).

BIOENERGY
KNOWLEDGE DISCOVERY FRAMEWORK
U.S. DEPARTMENT OF ENERGY

Log in

Key Topics ▾ Bioenergy Library Search the KDF Overview Models & Tools

About the Bioenergy KDF

The Bioenergy Knowledge Discovery Framework (KDF) supports the development of a sustainable domestic bioeconomy by providing access to a variety of data sets, publications, and visualization tools that support bioenergy research, analysis, and decision making.

[Overview](#)

DOE Bioenergy Research

This site provides access to reports of current DOE BioEnergy Technology Office (BETO) research and development (R&D) efforts, including state of technology (SOT) reports, technology design pathways analyses, life cycle assessments (LCAs), techno-economic analyses (TEAs), and a supply chain analysis.

[Search](#)

In the KDF, users can search for information, explore key topics in bioenergy research and the bioeconomy, download data, and visualize information in an integrated manner.

Bioenergy Data

- Bioenergy Library
- DOE Funded Research
- Models & Tools
- Techno-Economic Analysis Database

Key Topics

- SOT & Design Pathway Reports
- Data Valorization & FAIR Data Delivery
- High-Octane Fuel Study
- Sustainability

Billion-Ton Study

- Billion-Ton 2016 Vol. 1
- Billion-Ton 2016 Vol. 2
- Billion-Ton Vol. 1 Data Download
- Billion-Ton Vol. 2 Data Download

<https://bioenergykdf.net/>

1 – Approach

- Data Access

- Collect and provide easy access to critical bioenergy datasets and information from BETO-funded projects
- Design website interface to make relevant and often complex data quickly accessible

- Community Engagement

- Interact with BETO and principal investigators (PIs) across the DOE Laboratories to ensure access to bioenergy datasets and reports generated through BETO's investments
- Establish a stakeholder engagement plan and focus groups to guide technical development tasks and priorities (e.g., new Billion Ton report data download interface)

- Success Factors

- Relevant, up-to-date content
- Easy access to priority content and data
- Engagement from user community (e.g., new contributions)
- Downloads and citations



Argonne
NATIONAL
LABORATORY



Idaho National Laboratory



NATIONAL RENEWABLE ENERGY LABORATORY



National Laboratory



Pacific Northwest
NATIONAL LABORATORY

1 – Approach

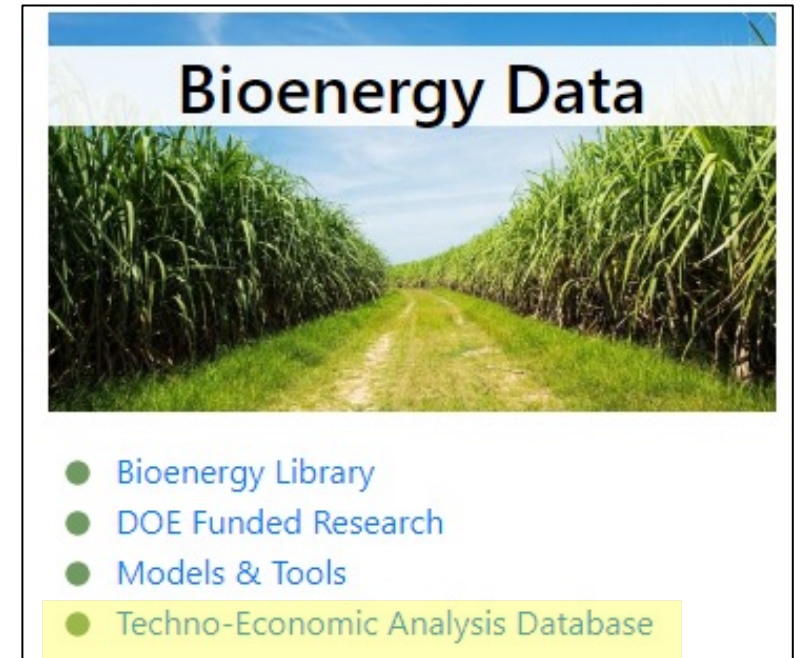
- Build from and customize existing software tools to develop a Government-owned web-based collaboration framework for knowledge management and data visualization:
 - Landing pages
 - Data interface(s)
 - Report downloads
 - Interactive visualizations
 - User-contributed datasets & publications
- Ensure KDF site is fully operational, security and software patches and updates are applied, and all contributed documents and datasets are backed up on a regular basis.



2 – Progress and Outcomes

2021 Highlights

- Added NREL's Biofuels Techno-Economic Analyses (TEA) Database
- Upgraded system to Drupal 9 and continued bug fixes and system maintenance for the latest release
- Made a Go/No-Go Decision in February 2021:
 - Continue the KDF project but with a more concentrated focus on the needs of bioenergy researchers
 - Work more closely with BETO and other bioenergy researchers to identify new data, tools, and models that can be disseminated via KDF



2 – Progress and Outcomes

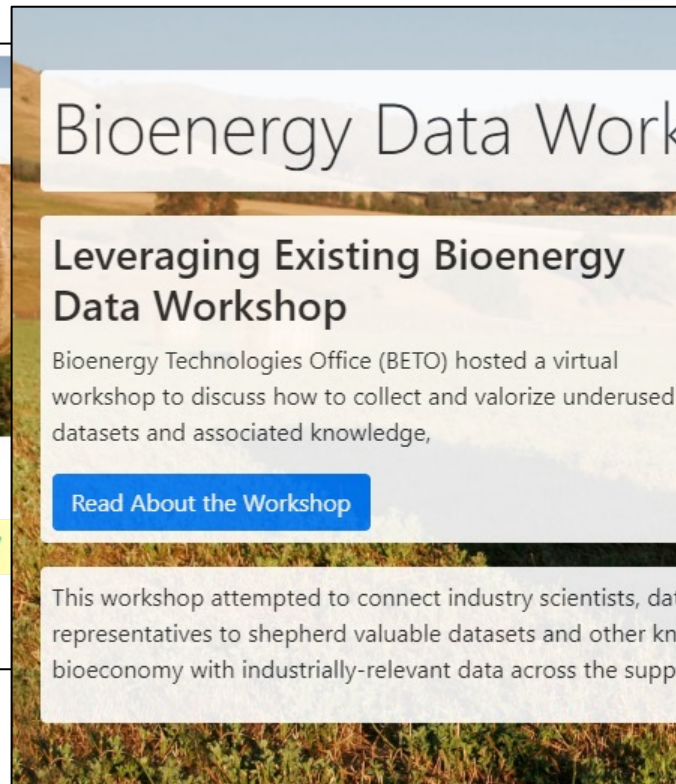
2022 Highlights

- Project Lead changed from a Geospatial web application developer to a Bioenergy researcher
- Created landing page to showcase BETO's "Data Valorization and FAIR Data Delivery" project
 - FAIR datasets meet principles of Findability, Accessibility, Interoperability, and Reusability
 - Previously unpublished biorefinery datasets can provide lessons learned for new ventures



Key Topics

- SOT & Design Pathway Reports
- **Data Valorization & FAIR Data Delivery**
- High-Octane Fuel Study
- Sustainability



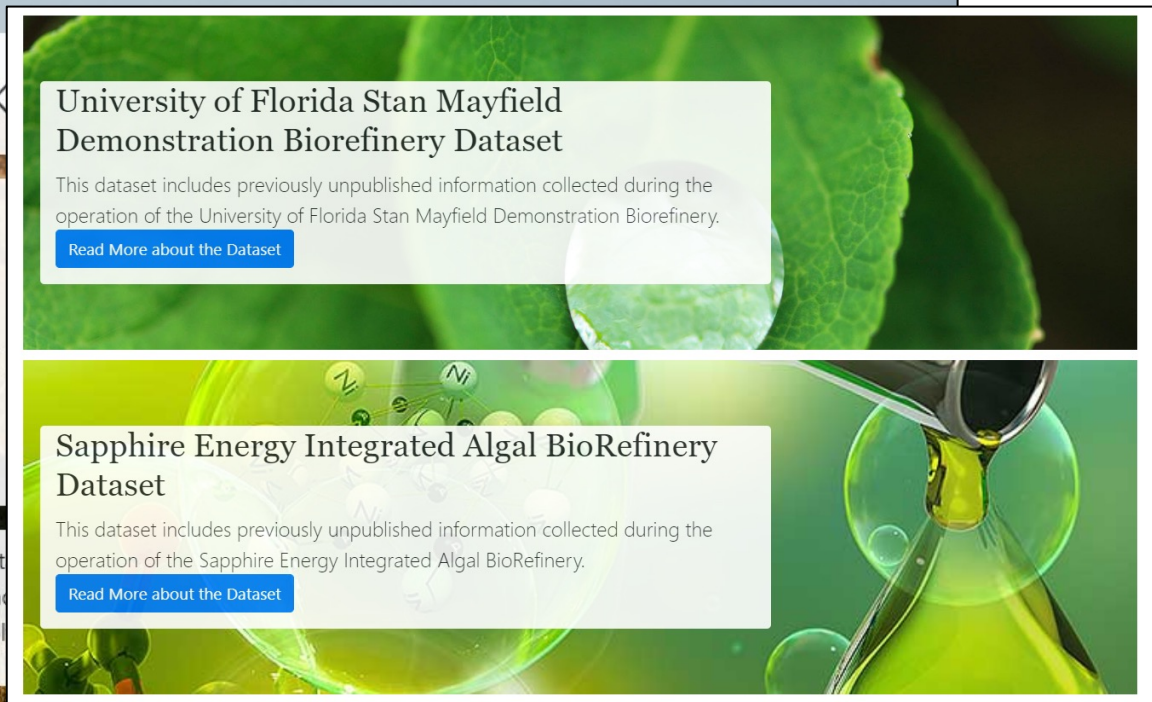
Bioenergy Data Workshop

Leveraging Existing Bioenergy Data Workshop

Bioenergy Technologies Office (BETO) hosted a virtual workshop to discuss how to collect and valorize underused datasets and associated knowledge,

[Read About the Workshop](#)

This workshop attempted to connect industry scientists, data representatives to shepherd valuable datasets and other knowledge into the bioeconomy with industrially-relevant data across the supply chain.



University of Florida Stan Mayfield Demonstration Biorefinery Dataset

This dataset includes previously unpublished information collected during the operation of the University of Florida Stan Mayfield Demonstration Biorefinery.

[Read More about the Dataset](#)

Sapphire Energy Integrated Algal BioRefinery Dataset

This dataset includes previously unpublished information collected during the operation of the Sapphire Energy Integrated Algal BioRefinery.

[Read More about the Dataset](#)

2 – Progress and Outcomes

2022 Highlights

- Created a data download interface for the Billion Ton 2016 Volume 2 Sustainability datasets available (e.g., changes in soil carbon, water quality, biodiversity, associated with BT16 biomass production scenarios)



Chapter Title	
Climate Sensitivity of Agricultural Energy Crop Productivity	1000463
Air Emissions from Agricultural and Forest Biomass Production	826994
Water Quality Effects on Forest Lands	236284
Water Consumption Footprint on Agricultural and Forest Lands	134637
Fossil Energy Consumption, Greenhouse Gas Emission, and Soil Carbon Effects	55962
Water Yield Effects on Forestlands	35466
Effects on Avian Biodiversity	18666

<https://bioenergykdf.net/bt16-vol2>

BT16 Vol2 Download Tool

[Go back to Billion-Ton 2016 Interactive Report](#)

Vol. 2 Download (Data for Chapters 4, 6, 7, 8, 9, 10, & 13)

Vol. 2 Download (only for Chapter5 - Water Quality Agriculture)

Billion-Ton 2016 Vol2 Download Tool (Chapter Data)

Rows of data to be downloaded

18666 Rows

Estimated time for CSV dataset generation

11.2 Seconds (Estimated)

[Generate Downloadable filtered dataset as CSV](#)

(After dataset generation, a download dialog will show, depending on size of dataset it may take some time.)

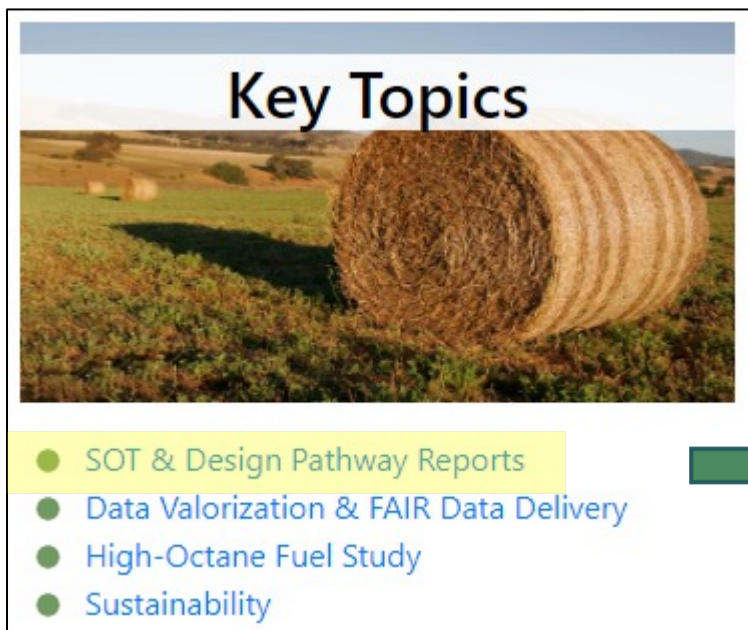
Displaying 1 - 80 of 18666

Year	Scenario	County FIPS Code	Feedstock	Indicator	Source	Value	Units	Chapter Title
2040	BC1	01001	Agriculture	birdfrs_rch	Prod	11.15312134	spc/km2	Effects on Avian Biodiversity
2040	BC1	01001	Agriculture	birdgen_rch	Prod	18.72627139	spc/km2	Effects on Avian Biodiversity
2040	BC1	01001	Agriculture	birdgls_rch	Prod	6.812632939	spc/km2	Effects on Avian Biodiversity
2040	BC1	01003	Agriculture	birdfrs_rch	Prod	10.92425027	spc/km2	Effects on Avian Biodiversity

2 – Progress and Outcomes

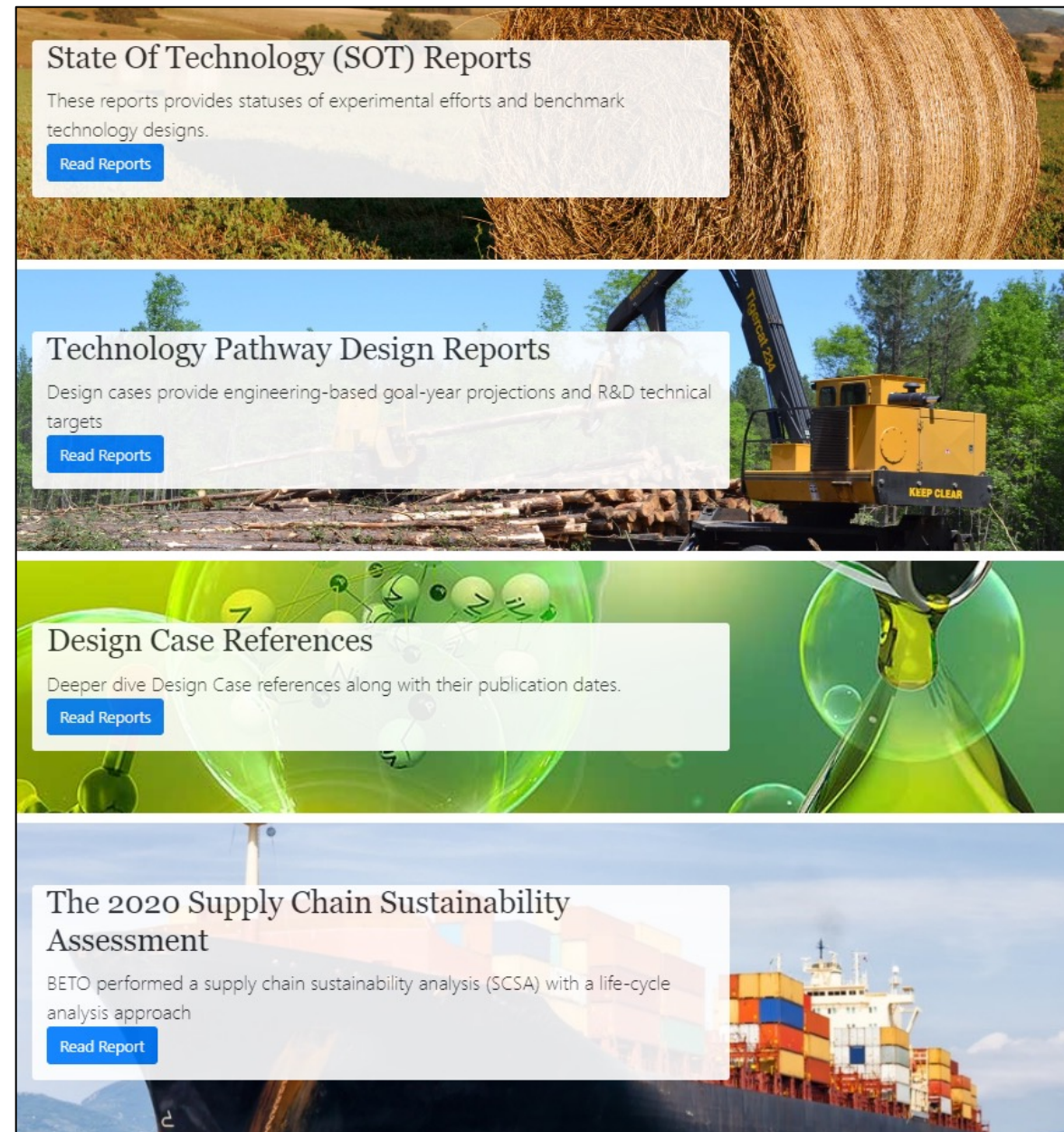
2022 Highlights

- Created a new landing page to consolidate access to BETO's State of Technology (SOT) and Design Pathway Reports produced across several different National Laboratories



Key Topics

- SOT & Design Pathway Reports
- Data Valorization & FAIR Data Delivery
- High-Octane Fuel Study
- Sustainability



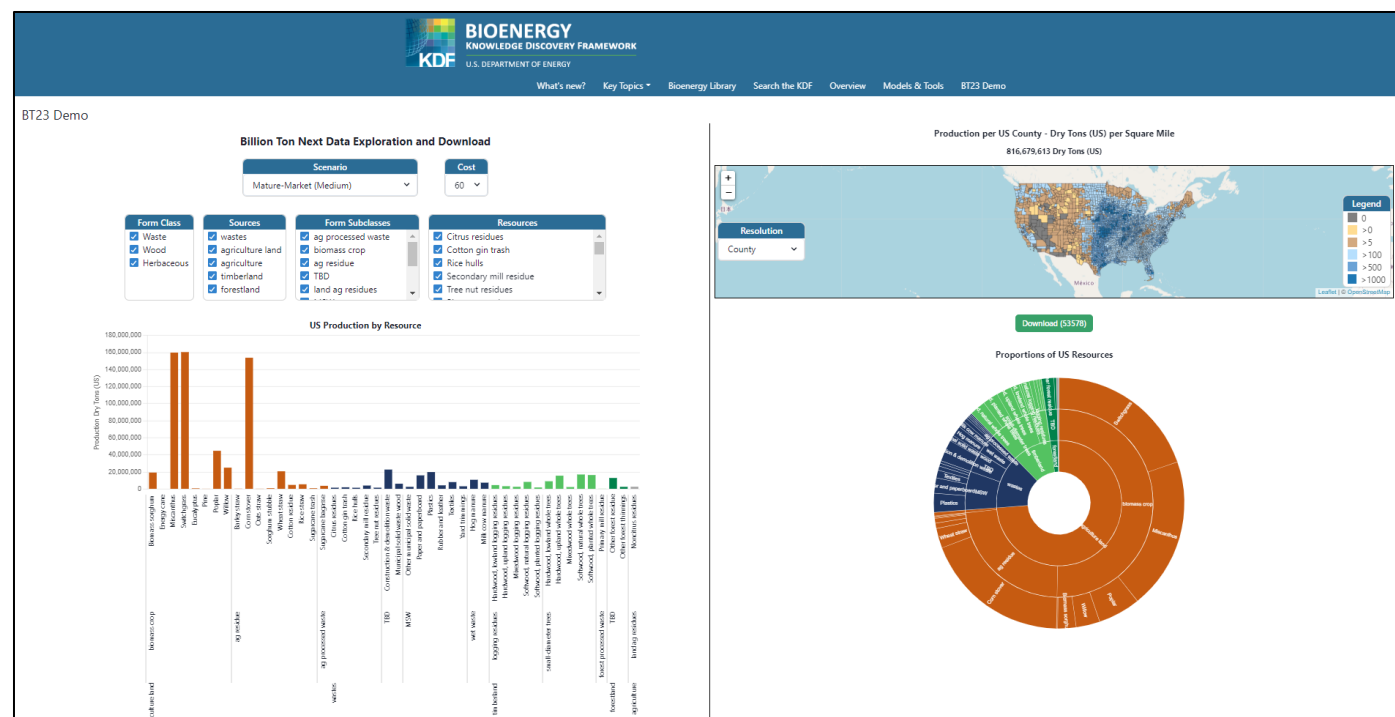
- State Of Technology (SOT) Reports**
These reports provides statuses of experimental efforts and benchmark technology designs.
[Read Reports](#)
- Technology Pathway Design Reports**
Design cases provide engineering-based goal-year projections and R&D technical targets
[Read Reports](#)
- Design Case References**
Deeper dive Design Case references along with their publication dates.
[Read Reports](#)
- The 2020 Supply Chain Sustainability Assessment**
BETO performed a supply chain sustainability analysis (SCSA) with a life-cycle analysis approach
[Read Report](#)

2 – Progress and Outcomes

2023 Highlights

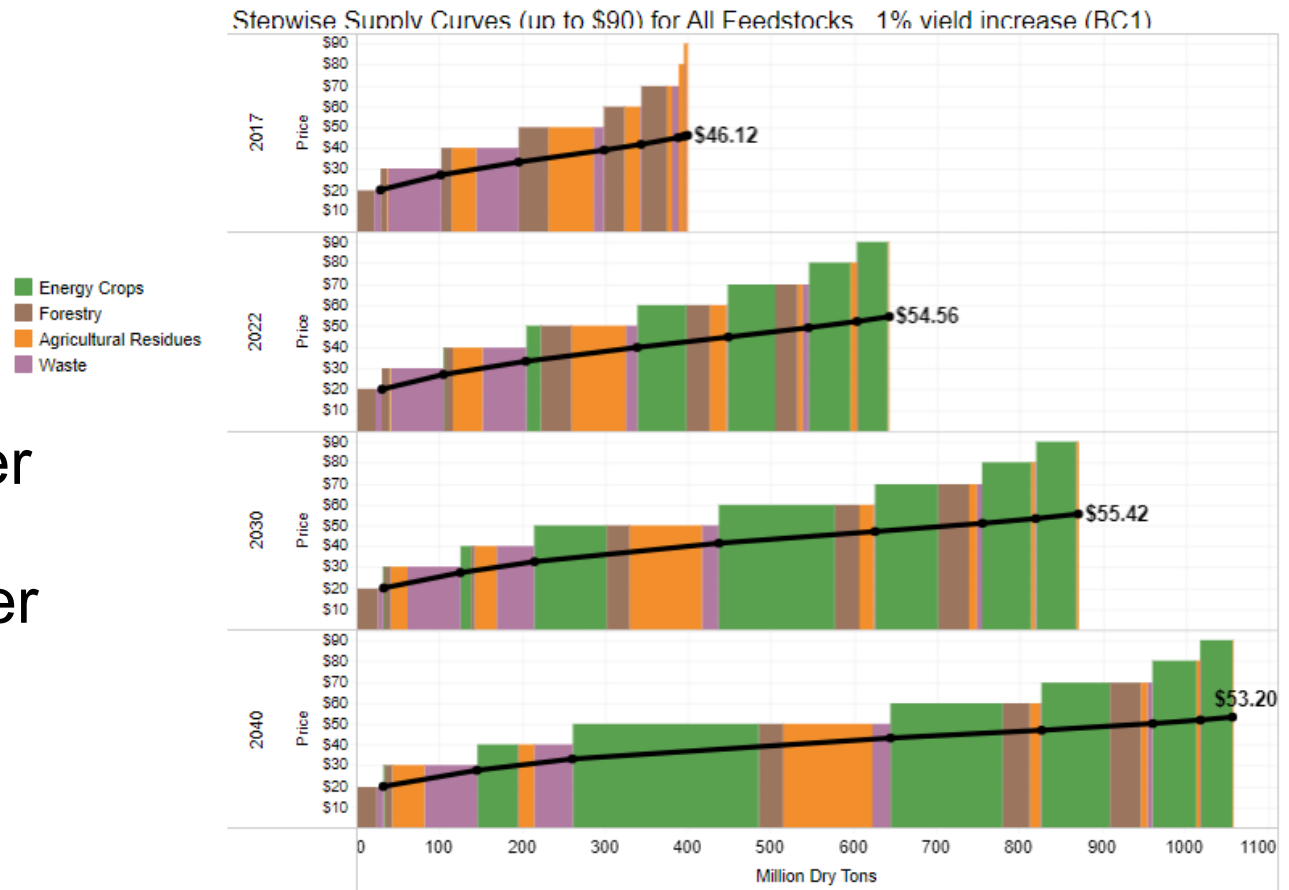
- Developed new code and workflow module to automatically generate DOE Office of Scientific and Technical Information (OSTI) digital object identifiers (DOIs) for user-contributed datasets and reports
 - DOIs all start with a dedicated KDF prefix (i.e., 10.23720/XXXXXX)
 - Users can check a box to add an inset value (e.g., BT23, Sustainability) to allow easier aggregation of like datasets (e.g., 10.23720/BT23/XXXXXX)
 - New DOI module promotes FAIR data principles

- Currently working with the Idaho National Lab's Biomass Feedstock Library (BFL) team to collect & present datasets from three BETO Affordable and Sustainable Energy Crop (ASEC) projects
- Currently developing a new database and interactive data dashboard for the next Billion Ton report based on feedback from EPA, USDA, industry



3 – Impact

- The fundamental objective of the KDF is to disseminate datasets, tools, and information that can help further research needed to accelerate decarbonization and development of the US bioeconomy
- KDF brings together BETO datasets and reports from
 - across the national labs
 - across various bioenergy supply chains
- Lessons learned from stakeholder experiences with the KDF web interface can be extended to other research domains within DOE



Summary

- **Approach**

- Design & develop a robust, collaborative informatics framework

- **Technical Accomplishments**

- Increased collaboration with BETO projects across multiple Labs due to change in project lead
- Consolidation of BETO's State of Technology & Design Pathway Reports produced (SOTs, LCAs, TEAs)
- Increased emphasis on meeting FAIR data principles (Findability, Accessibility, Interoperability, and Reusability)

- **Impact**

- Providing access to most up-to-date Bioenergy Data to accelerate decarbonization and development of the US Bioeconomy



Bioenergy Knowledge Discovery Framework (KDF)/6.3.0.2

Esther Parish/ORNL

WBS # 6.3.0.2-Agreement # 22893

Project Objectives

- Develop a “go-to” web-based repository of BETO-funded datasets, reports, and publications for researchers proposing new projects needed to accelerate decarbonization and development of the US bioeconomy.

Technical Approach

- Focus on three tasks: (1) KDF maintenance, (2) KDF connections with other BETO projects, and (3) KDF enhancements. A primary focus for FY23 will be the development of a new database and data download interface for the Billion Ton 2023 Update (BT23) report.

Project Milestones and Outcomes

- The project’s Go/No-Go Decision was completed in February 2021.
- By 12/31/22: Assign DOI tags to uploaded datasets
- By 06/30/22: Collect and present datasets from first completed ASEC FOA project
- By 09/30/2023: Release new BT23 report and data interface

Decarbonization Pillars and EERE Emphasis Areas

The KDF’s existing structure has been designed to address the five critical elements of the bioenergy infrastructure: (a) Production; (b) Reliability; (c) Security; (d) Agility; and (e) Sustainability. This cross-cutting project will continue to build a repository of datasets, reports and publications related to all three decarbonization pillars (Transportation, Industry & Agriculture) and multiple BETO focus areas and targets.

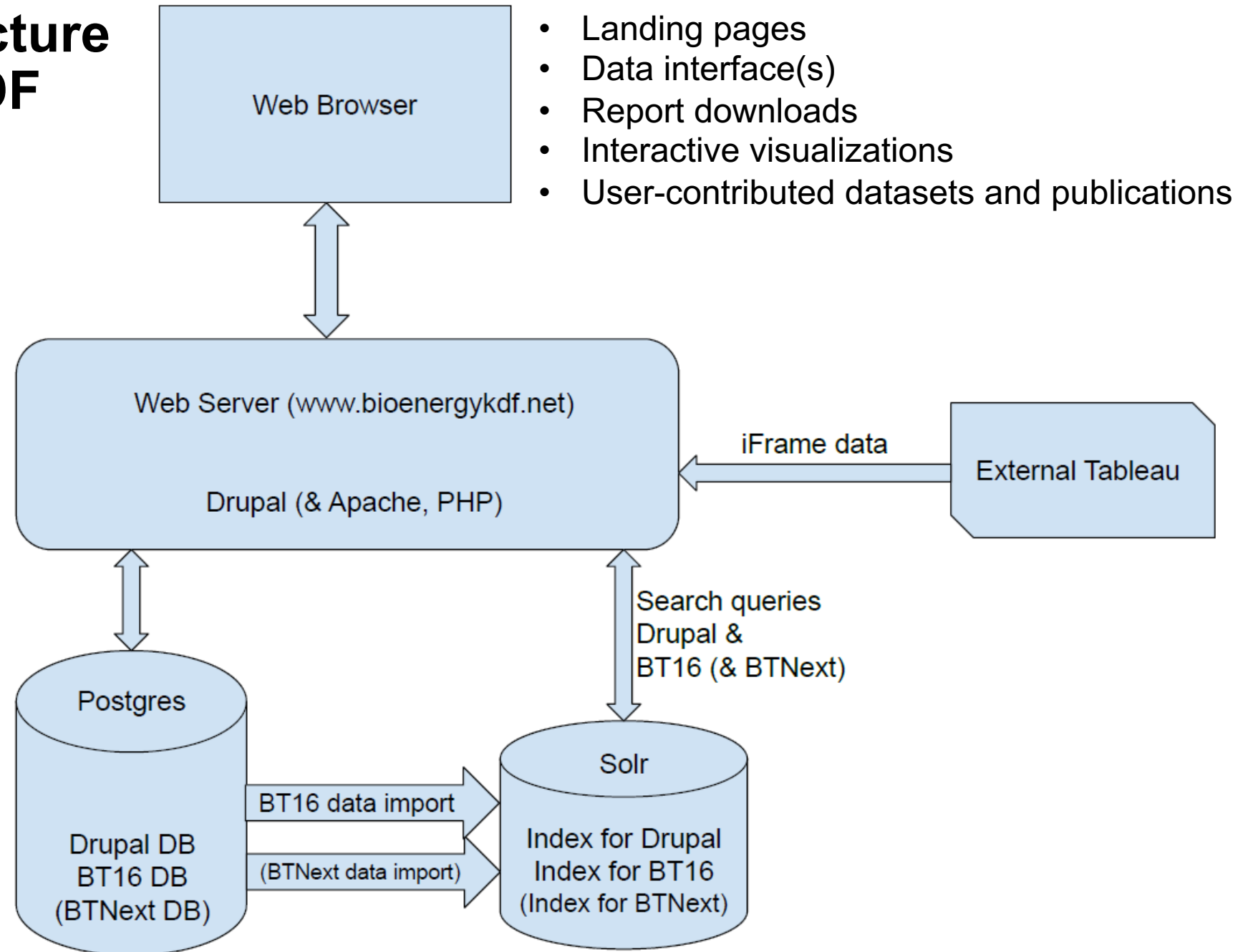
Project Attributes

All BETO Projects

DOE TM Lead

Additional Slides

Current Architecture of Bioenergy KDF



Responses to Previous Reviewers' Comments

- Comment: It is good to see the strategic pivots this project has taken. Given that the KDF is a key conduit to disseminate information, it is good to see the website reflect office priorities on the landing page, such as the marine and aviation fuel pages. The more the project can define its audience and hone interests/needs and barriers for its audience, the more likely that DOE-funded research will be disseminated to the right stakeholders.
- Response: The project is now gaining traction by defining KDF's primary audience as the community of researchers seeking to propose new projects needed to accelerate decarbonization and development of the US bioeconomy. These researchers need easier access to the collection of previously funded BETO reports and datasets so that they can absorb lessons learned and move forward with new ideas that have increased viability. These datasets are currently not as readily as accessible as they could be.

Responses to Previous Reviewers' Comments

- Comment: The apparent shift from a user-updated model to curated content model is a worthwhile strategy because it can ensure that data sets are up to date, maintained, and high quality. In particular, this tool can serve a useful role as a geospatial data repository because a great deal of overhead in both compute resources and data management is involved in hosting geospatial data, making it difficult for smaller entities. It is not immediately clear how much interest or web traffic the KDF is attracting beyond the Billion Ton sites.
- Response: KDF has the capacity to store large numbers of user-contributed datasets. The challenge has been getting users to add their datasets (e.g., geospatial data) to the BETO repository. The project team is now making a more concerted effort to reach out to project PI's with datasets that can be useful to a broader group of researchers. The fact that several recent BETO FOAs (e.g., ASEC, BioRESTORE) have included stipulations that the projects will need to contribute final datasets to the KDF and the INL Biomass Feedstock Library is giving KDF new leverage to ingest and provide access to more datasets.

Responses to Previous Reviewers' Comments

- Comment: Relative site analytics (compared to previous years or peer review cycles) would be helpful to see if usage is growing over time.
- Response: Setting the proper criteria to understand impacts in a web application like the KDF can be difficult. Just looking at number of page views or user accounts doesn't tell the full story. The value of the KDF is easing the access to high value data and exploration tools that were previously difficult to obtain. An example used with the Billion Ton Update was that, in a given year, having the data accessible on the KDF saved DOE around \$1 million because researchers were no longer having to field requests for data access or to subset the data for their specific needs this was done for them on the KDF. That said, we are now working to get KDF registered with ORNL-approved "Site Improve" software so that we can begin to collect metrics related to site usage, downloads, etc. (Google Analytics raised security concerns.)

Responses to Previous Reviewers' Comments

- Comment: It was not clear how this project is different from or related to the Findability, Accessibility, Interoperability, and Reusability (FAIR) data project or the Alternative Fuels Data Center.

- Response:

KDF is making an intentional effort to incorporate FAIR principles into its framework, as is shown by example of the recently developed capability to automatically assign OSTI DOIs to user-contributed datasets. KDF has also developed a landing page to showcase and provide access to previously unpublished biorefinery datasets retrieved through BETO's FAIR project.

The Alternative Fuels Data Center (AFDC) provides information, data, and tools specifically targeted to fleets and other transportation decision makers to find ways to reach their energy and economic goals. By contrast, KDF is targeted to researchers needing to access datasets and reports related to multiple bioenergy pathways so that they can plan future R&D activities for a decarbonized future.