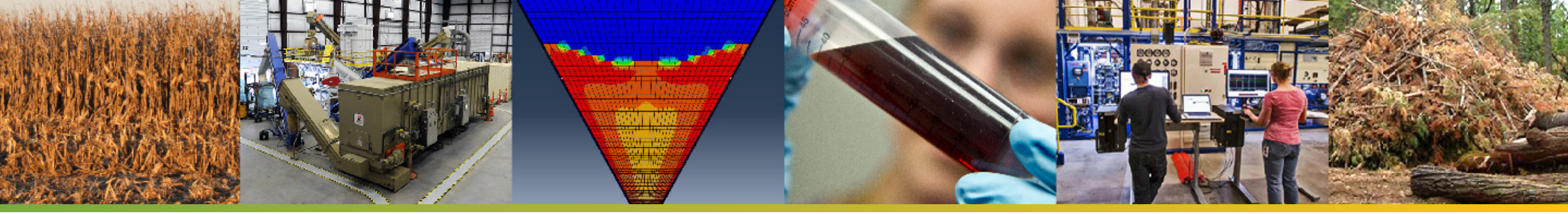


DOE Bioenergy Technologies Office (BETO) 2021 Project Peer Review

3/16/2021

Feedstock-Conversion Interface Consortium

Task 4 – Data Integration and Management

Jim Collett (PNNL) and Rachel Emerson (INL)

fcic
FEEDSTOCK-CONVERSION INTERFACE CONSORTIUM

PNNL-SA-159498

FCIC Task Organization

Feedstock → Preprocessing → Conversion

Task 2: Feedstock Variability

Task 5: Preprocessing

Task 6: Conversion High-Temp

Task 1: Materials of Construction

Task 7: Conversion Low-Temp

Task 3: Materials Handling

Enabling Tasks

Task X: Project Management

Task 4: Data Integration

Task 8: TEA/LCA

Task X: Project Management: Provide scientific leadership and organizational project management

Task 1: Materials of Construction: Specify materials that do not corrode, wear, or break at unacceptable rates

Task 2: Feedstock Variability: Quantify & understand the sources of biomass resource and feedstock variability

Task 3: Materials Handling: Develop tools that enable continuous, steady, trouble free feed into reactors

Task 4: Data Integration: Ensure the data generated in the FCIC are curated and stored – FAIR guidelines

Task 5: Preprocessing: Enable well-defined and homogeneous feedstock from variable biomass resources

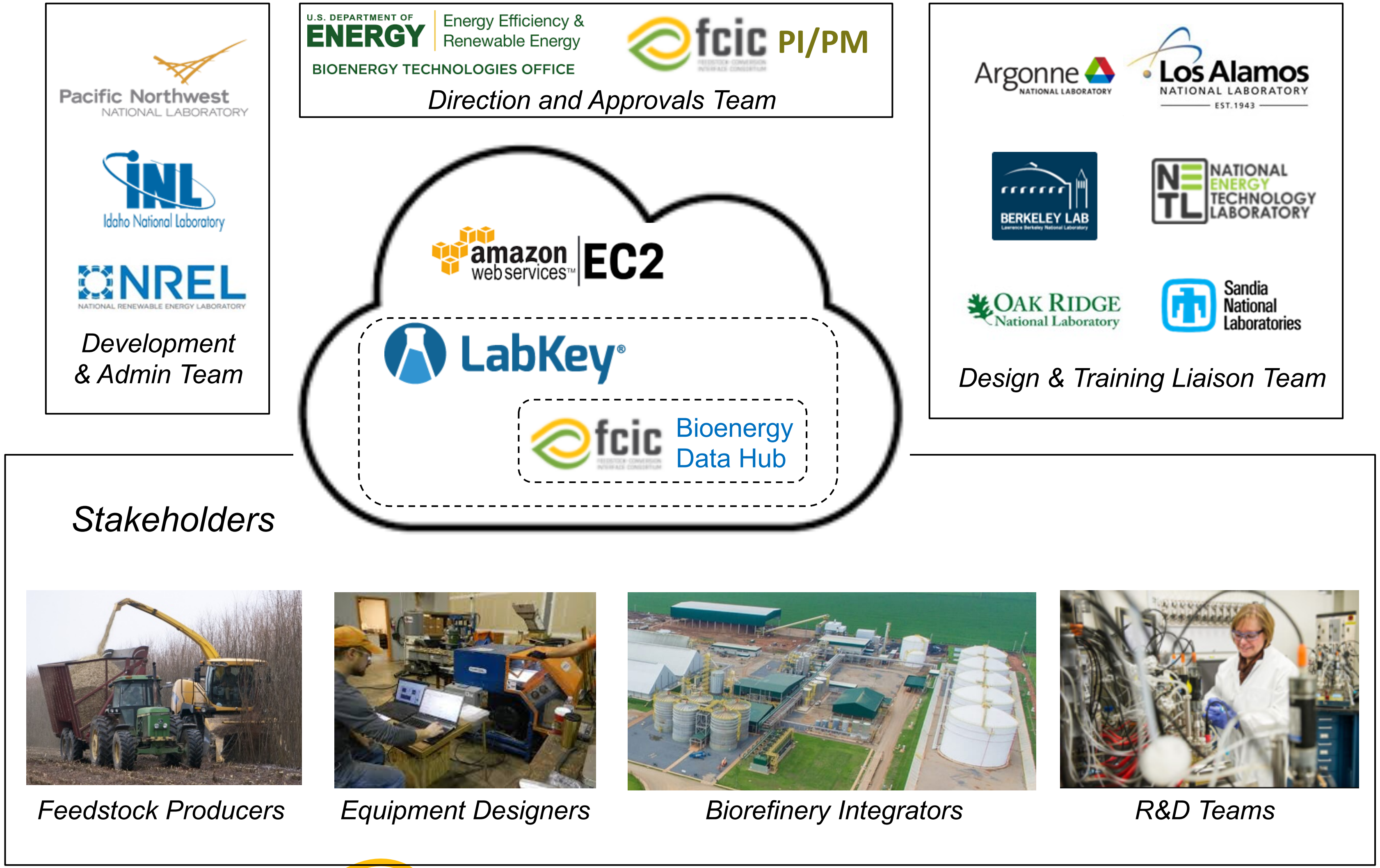
Task 6 & 7: Conversion (High- & Low-Temp Pathways): Produce homogeneous intermediates to convert into market-ready products

Task 8:Crosscutting Analyses TEA/LCA: Valuation of intermediate streams & quantify variability impact

Project Overview

Objectives:

- Provide a web-enabled database for integrating, standardizing, and archiving FCIC data intended for public release, and for channeling integrated datasets to Task 8 Cross-cutting Analysis.
- Establish controlled vocabularies, data standards, and critical property specifications for biorefinery unit operations and process streams within a Quality by Design (QbD) framework.
- Deploy a public web portal to disseminate FCIC datasets, specifications, analyses, methods, and software to industry stakeholders who will grow the new bioeconomy.



Project Overview

Current limitations:

- A lack of data standards and public, commercial-quality data servers holds back the bioeconomy by limiting dissemination of BETO-funded datasets and software tools for use by equipment designers and biorefineries.
- FCIC R&D teams do not have a modern digital platform to collaboratively assemble and publicly share the massive amount of data they generate, and still rely heavily on emailing spreadsheets or share drives to exchange data.

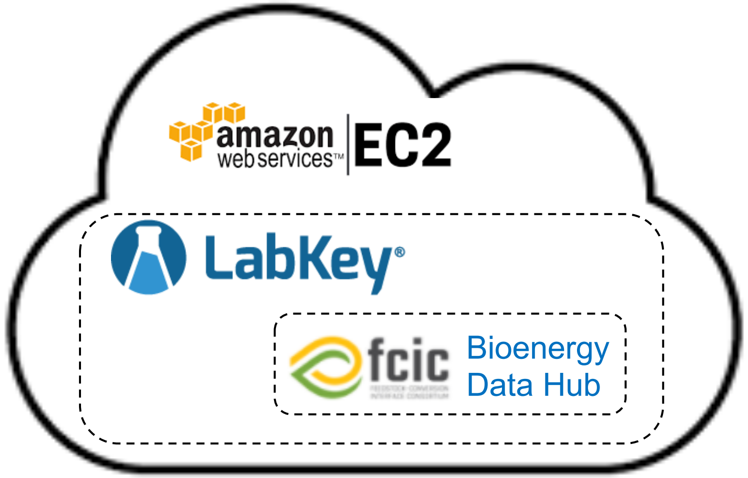
Relevance:

- This enabling task seeks to make every researcher in the FCIC more efficient and effective by providing a modern, commercial-quality platform for collaboration and data sharing across 8 National Labs.
- BETO-funded datasets, models, and QbD tools are now ***F**indable, **A**ccessible, **I**nteroperable, and **R**eusable* (**FAIR**) in accordance with DOE mandates for data modernization.
- This task directly supports 3 (of 5) major recommendations from the FCIC 2019 Peer Review report (p. 67):
  - Develop innovative ways to immerse the national laboratories with industry.
  - Establish quality specifications on feedstocks
  - Aggressively promote FCIC results and visibility.

Risks:

- Sustained commitment to data standards and timely data uploading, integration, and dissemination are essential for encouraging stakeholder confidence in the FCIC.
- Meeting User Experience expectations and maintaining near 100% uptime are necessary to turn new Data Hub users into devoted fans of FCIC knowledge products and tools.

1 – Management



Subtask	Lead(s)	Major Responsibilities
4.2	Jim Collett (PNNL) and Dave Sievers (NREL)	FCIC Data Hub Infrastructure, Tools, and Training
4.4	Rachel Emerson (INL) and Dave Sievers (NREL)	Facilitating Task 8 TEA

Risk Mitigation:

- LabKey Premium Edition Server Software on the FCIC Data Hub provides:
  - enterprise-grade, commercial-quality web application software deployed on the Amazon Web Services cloud for excellent uptime and availability.
  - Same-day technical support and weekly meetings with LabKey developers.
- FCIC Data Hub Content Management
  - Separate "production" and "development" servers are maintained by a professional IT team at PNNL to minimize risks of code deployment and software upgrades
  - Datasets and code are compatible with the free, open-source Community Edition of LabKey Server to allow legacy access to data if funding for the consortium is discontinued.

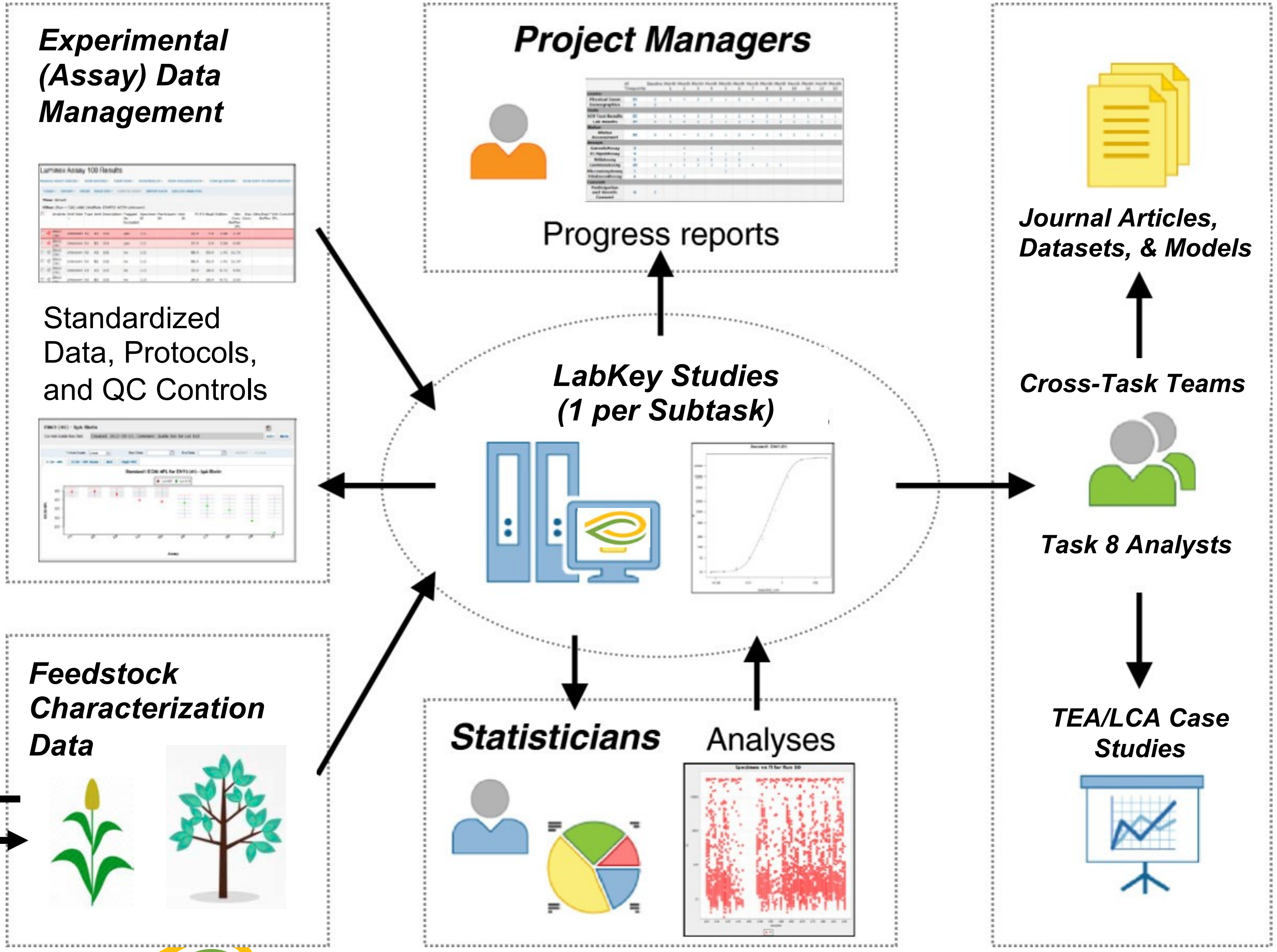
Communication strategy:

- Core Data Hub developers, data scientists, and IT staff coordinate via Microsoft Teams and LabKey's built-in issue tracking system.
- Funding is provided for "Data Liaisons" at all FCIC NLs to attend biweekly meetings and annual meeting.
- Eight hours of online LabKey user and developer training is made available to FCIC researchers each fiscal year.
- Feedback sessions with IAB and phased public roll-out during planned for FY21.

2 – Approach

Technical Approach:

- LabKey workflows for integrating bench-to-PDU scale experimental data and feedstock characterization data within ”Study” data structures have been created on the Data Hub.
- Each of the FCIC’s 35 Subtasks has its own Study folder wherein datasets, statistical views of the data, figures, and text may be collaboratively compiled into publications supported by curated, downloadable datasets.



Adapted from BMC Bioinformatics volume 14, Article number: 145 (2013)

2 – Approach

Technical Challenges:

- R&D teams have for generations worked within "silos" within their institutions; finding consensus on interoperable data formats, definitions, and protocols to enable the success of the FCIC is an ongoing but surmountable challenge.
- Data standards and definitions must originate with the researchers who are experts in their field—some with decades of experience.
- Striking a balance to encourage FCIC researchers to share data via the Data Hub while still allowing for publishing advantage and patenting.

Technical Metrics:

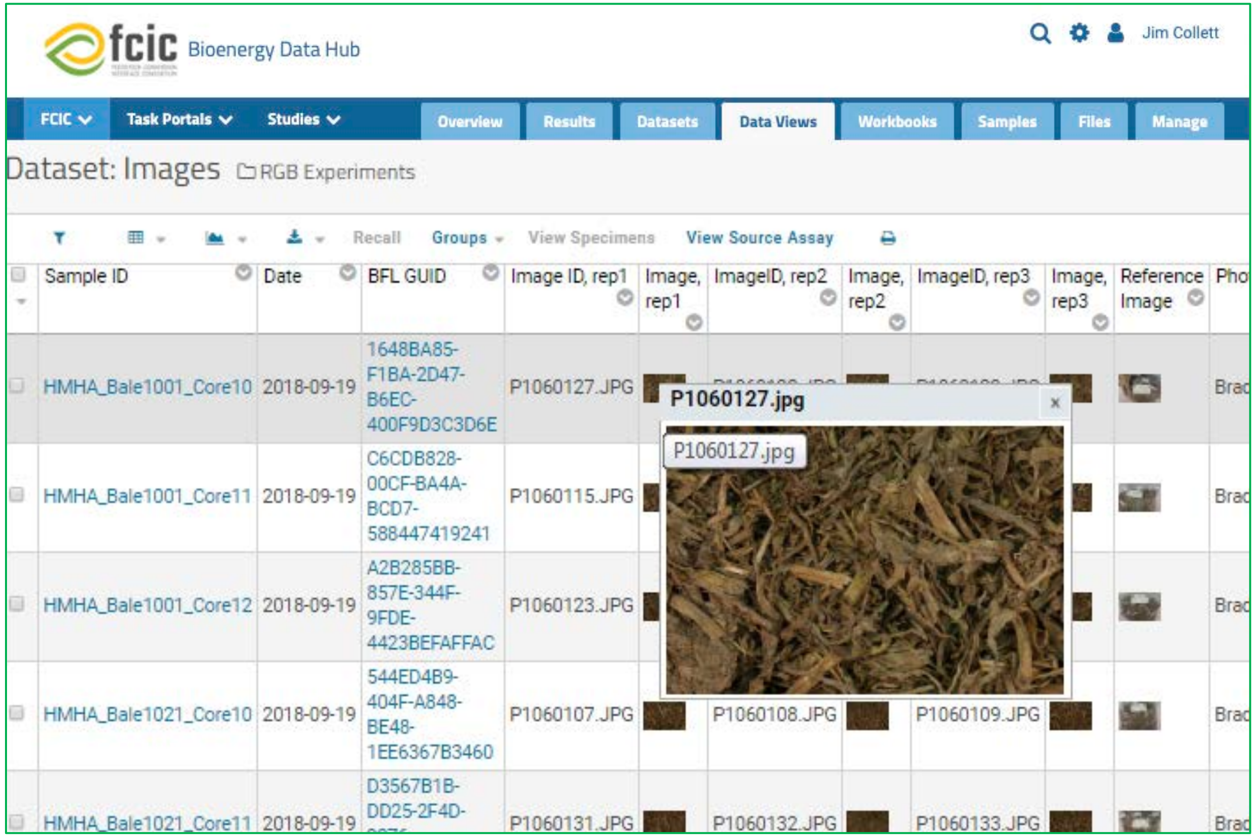
- LabKey Server is an FDA-compliant database with password protected, role-based permissions and extensive activity auditing tools that describe in detail who is using the Data Hub, how often they are using it, and what data and information they are accessing.
- LabKey's built-in survey tools will be used to collect feedback from stakeholders during public roll-out in FY21.

3 – Impact

Impact:

- FCIC research teams across 9 National Labs now have:
  - A modern, commercial-quality collaborative platform to integrate, standardize, and archive datasets for advanced analytics and TEA/LCA.
  - A web portal with advanced search tools for sharing FCIC datasets with industrial stakeholders.

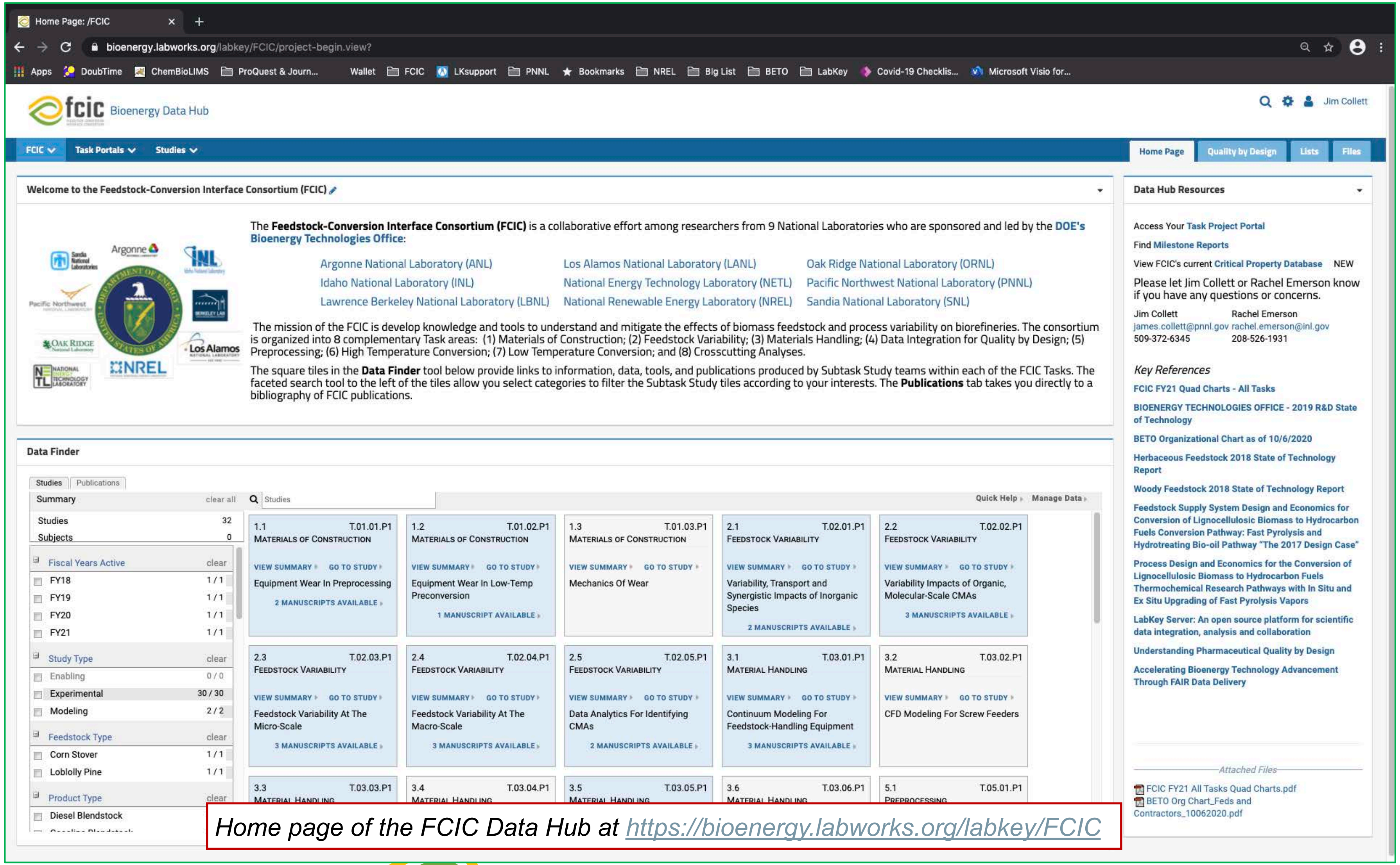
- The Data Hub's QbD database and searchable Subtasks and publications will provide well-documented data and evidence for:
  - Material properties for biorefinery feedstocks, intermediates, and products.
  - Process parameters for commercial unit operations to facilitate biorefinery scale-up.
  - Data that can support market specifications for feedstocks, intermediates, and products throughout the value chain.



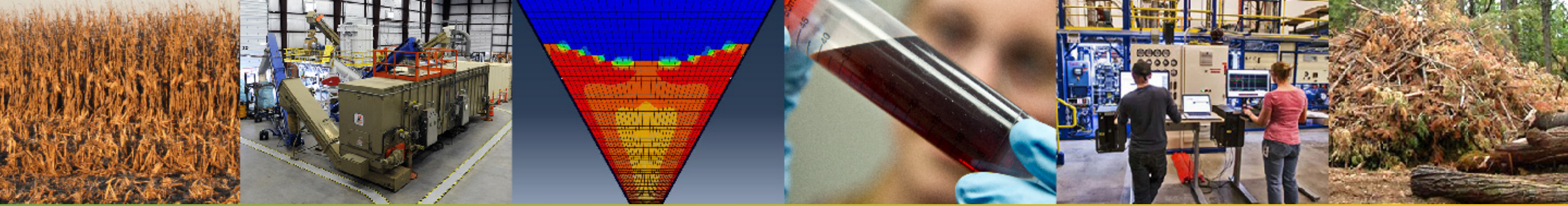
LabKey dataset being used as training data for machine learning to dissect corn stover bale quality in real time within Subtask 2.4, Feedstock Variability At The Macro-Scale. "BFL GUID" column entries directly link to samples in the Bioenergy Feedstock Library at INL.

Dissemination:

- We will promote the FCIC Data Hub as the "go-to" platform for datasets on feedstock variability and biorefinery unit operations via webinars, publications, and conference presentations.
- High quality, curated datasets on the Data Hub will be convincing "calling cards" for industrial partnerships that lead to commercialization and deployment of BETO-funded research.



Home page of the FCIC Data Hub at https://bioenergy.labworks.org/labkey/FCIC



4 – Progress and Outcomes

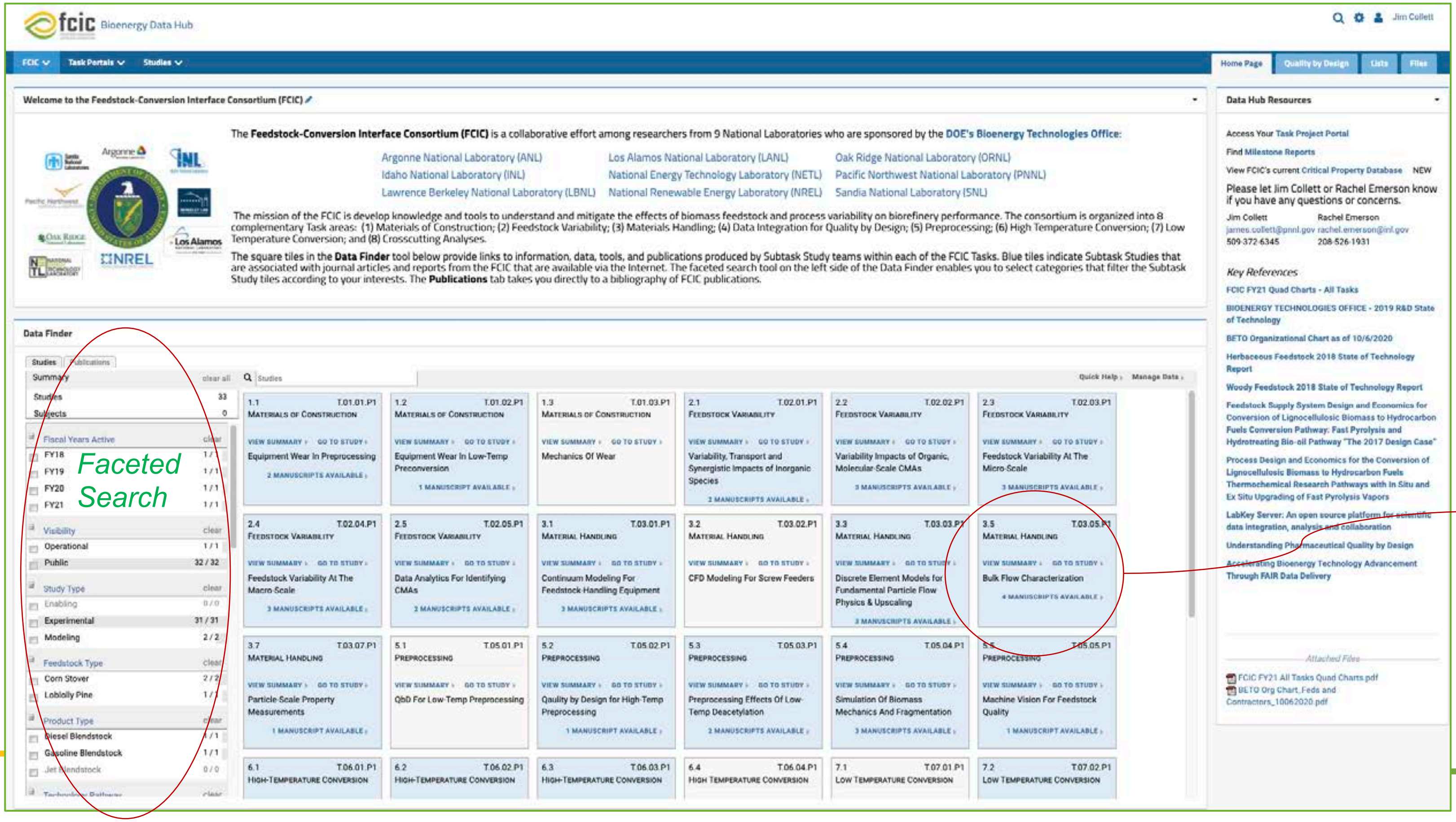
fcic
FEEDSTOCK-CONVERSION
INTERFACE CONSORTIUM

4 – Progress: Finding Knowledge and Tools on the FCIC Data Hub

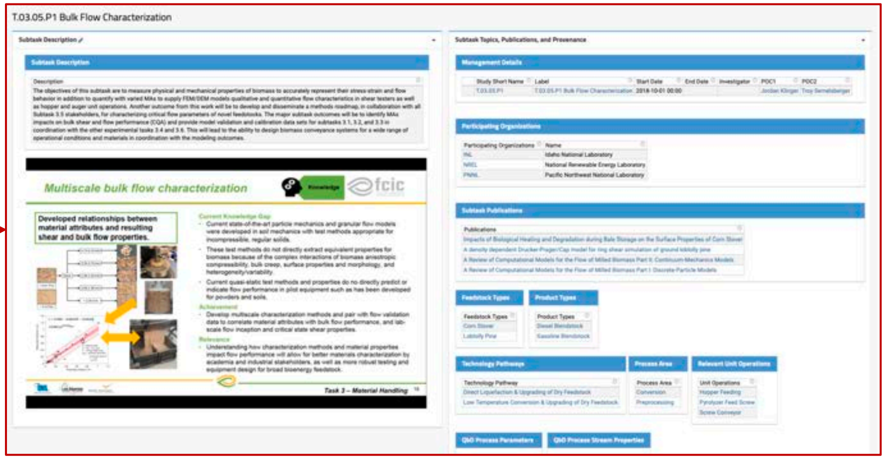
Tools

fcic FEEDSTOCK-CONVERSION INTERFACE CONSORTIUM

- The Data Hub Web Portal is now online and equipped with a “Data Finder” dashboard for accessing “Study” folders for each FCIC Subtask. FCIC publications may be accessed via the “Publications” tab.
- Data accruing on the Data Hub are easily **Findable** and **Accessible** (per FAIR data guidelines) via a faceted search tool that allows users to click on metadata tags to rapidly down-select to Subtasks covering specific R&D focus areas.



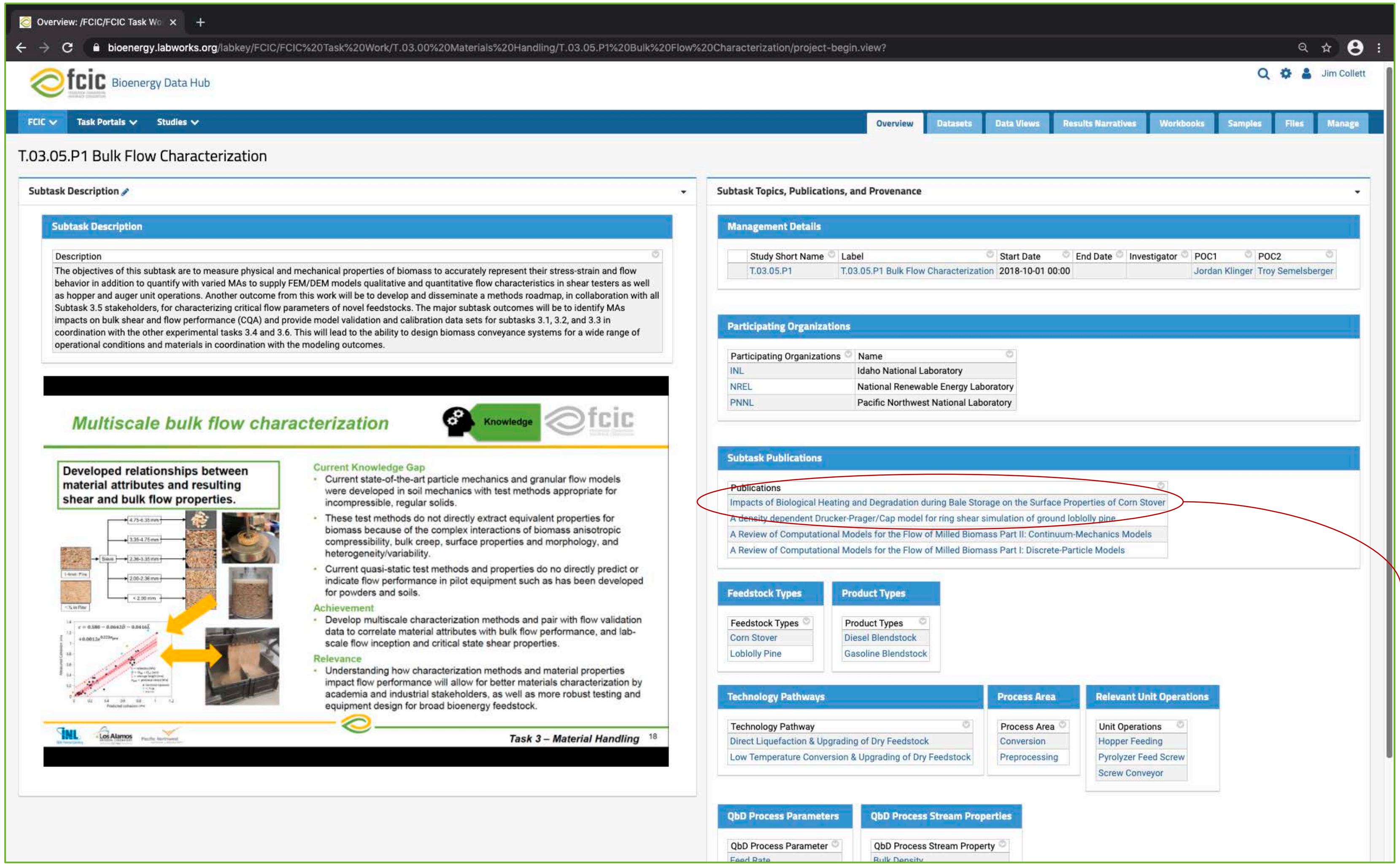
Subtask metadata tags now include a range of 260 descriptors for analytical methods, feedstock types, product types, technology pathways, process areas, unit operations, process parameters, process stream properties, fiscal years active, and participating organizations.



4 – Progress: LabKey Subtask Studies

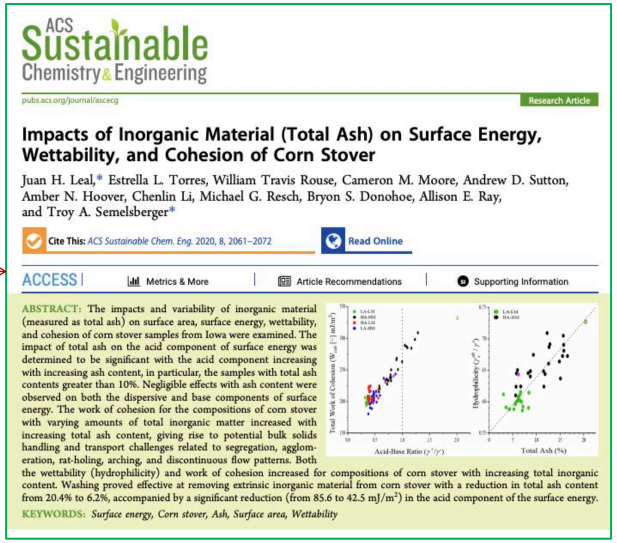
Tools

fcic FEEDSTOCK-CONVERSION INTERFACE CONSORTIUM



Technical Approach:

- Each of the FCIC's 35 Subtasks has its own "Study" home page on the Data Hub
- Studies have links for datasets, analyzed data views, results narratives, provenance, publications, and metadata associated with the Subtask.
- Subtask Studies provide supporting data and information for the 66 "Knowledge" and "Tool" products developed thus far by the FCIC.



ACS Sustainable Chemistry & Engineering

pubs.acs.org/journal/ascecg — Research Article

Impacts of Inorganic Material (Total Ash) on Surface Energy, Wettability, and Cohesion of Corn Stover

Juan H. Leal,* Estrella L. Torres, William Travis Rouse, Cameron M. Moore, Andrew D. Sutton, Amber N. Hoover, Chenlin Li, Michael G. Resch, Bryon S. Donohoe, Allison E. Ray, and Troy A. Semelsberger*

Cite This: ACS Sustainable Chem. Eng. 2020, 8, 2061–2072 — Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: The impacts and variability of inorganic material (measured as total ash) on surface area, surface energy, wettability, and cohesion of corn stover samples from Iowa were examined. The impact of total ash on the acid component of surface energy was determined to be significant with the acid component increasing with increasing ash content, in particular, the samples with total ash contents greater than 10%. Negligible effects with ash content were observed on both the dispersive and base components of surface energy. The work of cohesion for the compositions of corn stover with varying amounts of total inorganic matter increased with increasing total ash content, giving rise to potential bulk solids handling and transport challenges related to segregation, agglomeration, rat-holing, arching, and discontinuous flow patterns. Both the wettability (hydrophilicity) and work of cohesion increased for compositions of corn stover with increasing total inorganic content. Washing proved effective at removing extrinsic inorganic material from corn stover with a reduction in total ash content from 20.4% to 6.2%, accompanied by a significant reduction (from 85.6 to 42.5 mJ/m^2) in the acid component of the surface energy.

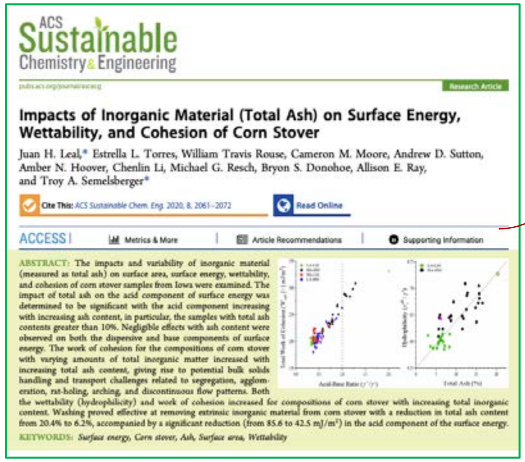
KEYWORDS: *Surface energy, Corn stover, Ash, Surface area, Wettability*

4 – Progress: Subtask Data Access and Export

Tools

fcic FEEDSTOCK-CONVERSION INTERFACE CONSORTIUM

- Harmonized, integrated datasets supporting FCIC publications may be easily found, filtered, sorted, and downloaded to Excel or to TSV or CSV text files.
- The integrated LANL, NREL and INL dataset shown below highlights how the Data Hub enables self-service access to the past work of others to supply **Interoperable** and **Reusable** FCIC data in accordance with FAIR data principles.


ACS Sustainable Chemistry & Engineering

Research Article

Impacts of Inorganic Material (Total Ash) on Surface Energy, Wettability, and Cohesion of Corn Stover

Juan H. Leal,* Estrella L. Torres, William Travis Rouse, Cameron M. Moore, Andrew D. Sutton, Amber N. Hoover, Chenlin Li, Michael G. Resch, Bryon S. Donohoe, Allison E. Ray, and Troy A. Semelsberger*

Cite This: ACS Sustainable Chem. Eng. 2020, 8, 2061−2072 | Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: The impacts and variability of inorganic material (measured as total ash) on surface area, surface energy, wettability, and cohesion of corn stover samples from Iowa were examined. The impact of total ash on the acid component of surface energy was determined to be significant with the acid component increasing with increasing ash content, in particular, the samples with total ash contents greater than 10%. Negligible effects with ash content were observed on both the dispersive and base components of surface energy. The work of cohesion for the compositions of corn stover with varying amounts of total inorganic matter increased with increasing total ash content, giving rise to potential bulk solids handling and transport challenges related to segregation, agglomeration, rat-holing, arching, and discontinuous flow patterns. Both the wettability (hydrophilicity) and work of cohesion increased for compositions of corn stover with increasing total inorganic content. Washing proved effective at removing extrinsic inorganic material from corn stover with a reduction in total ash content from 20.4% to 6.2%, accompanied by a significant reduction (from 85.6 to 42.5 mJ/m²) in the acid component of the surface energy.

KEYWORDS: Surface energy, Corn stover, Ash, Surface area, Wettability

FCIC publication accessible on Web

fcic Bioenergy Data Hub | Jim Collett

FCIC | Task Portals | Studies | Overview | Datasets | Data Views | Results Narratives | Workbooks | Samples | Files | Manage

Corn Stover N2 Surface Area and Compositional Analysis | T.03.05.P1 Bulk Flow Characterization

1 - 37 of 37

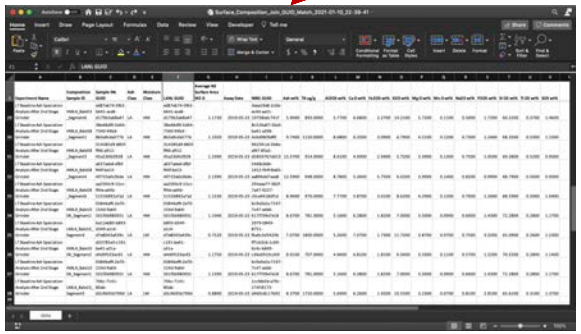
Excel | Text | Script

- Excel Workbook (.xlsx) Maximum 1,048,576 rows and 16,384 columns.
- Excel Old Binary Workbook (.xls) Maximum 65,536 rows and 256 columns.
- Refreshable Web Query (.iqy)

Column headers: Caption

Export

Experiment Name	Composition Sample ID	Sample INL GUID	Ash Class	Moisture Class	LANL GUID	Average N2 Surface Area M2 G	Assay Date	NREL GUID	Ash wt%	TX ug/g	Al2O3 wt%	Ca O wt%	Fe2O3 wt%	K2O wt%	Mg O wt%	Mn O wt%	Na2O wt%	P2O5 wt%	Si O2 wt%	Ti O2 wt%	SO3 wt%
LT Baseline Ash Speciation Analysis After 2nd Stage Grinder	LMLA_Bale11_Segment1	ca03f467-348d-e944-94c4-9b4dff2cc3fc	LA	LM	ca03f467-348d-e944-94c4-9b4dff2cc3fc	0.883	2019-05-23	61d60b5d-87ad-804e-b9e3-59acfc2e5a62	7.95	1720.0	5.49	6.04	1.79	10.66	3.58	0.06	0.79	2.49	64.82	0.26	1.35
LT Baseline Ash Speciation Analysis After 2nd Stage Grinder	LMLA_Bale11_Segment4	fa8f3731-0154-1b41-8f89-b4b20808bc93	LA	LM	fa8f3731-0154-1b41-8f89-b4b20808bc93	0.746	2019-05-23	61f4e57e-ece4-5848-a0b4-10f86f7e698d	8.94	1750.0	6.32	5.92	1.99	8.94	3.27	0.08	0.99	2.56	65.8	0.28	1.31
LT Baseline Ash Speciation Analysis After 2nd Stage Grinder	LMLA_Bale11_Segment6	f2772330-12f2-bc40-a6d0-218d1051bb4e	LA	LM	f2772330-12f2-bc40-a6d0-218d1051bb4e	0.732	2019-05-23	dcbcc9a4-525d-294d-a85d-aa0c5a018536	8.07	1730.0	5.57	6.09	1.71	11.09	3.22	0.06	0.91	2.74	67.17	0.25	1.38
LT Baseline Ash Speciation Analysis After 2nd Stage Grinder	LMLA_Bale12_Segment2	4c383541-55dc-2441-91e0-a79f5a6c9060	LA	LM	4c383541-55dc-2441-91e0-a79f5a6c9060	0.842	2019-05-23	372bf554-6f1f-d742-bc74-8c85daba5cdf	9.48	1630.0	6.58	5.41	2.02	9.5	3.0	0.08	1.08	2.37	68.18	0.29	1.26
LT Baseline Ash Speciation Analysis After 2nd Stage Grinder	LMLA_Bale12_Segment3	cfe01cb9-8d28-f34c-8987-970c4aafc541	LA	LM	cfe01cb9-8d28-f34c-8987-970c4aafc541	0.667	2019-05-23	ad54c824-1ee6-5642-8981-6aec8bdb566a	5.57	1170.0	5.07	7.52	1.61	13.22	4.4	0.07	0.78	3.41	59.12	0.23	1.62
LT Baseline Ash Speciation Analysis After 2nd Stage Grinder	LMLA_Bale12_Segment4	1c2a8dcf-1ea4-a448-8a9b-591589f2502d	LA	LM	1c2a8dcf-1ea4-a448-8a9b-591589f2502d	0.772	2019-05-23	4fa560ea-b90d-da4f-a5ab-077afca7d535	6.73	1490.0	5.29	6.9	1.77	12.93	3.71	0.08	0.78	3.56	61.55	0.26	1.62
LT Baseline Ash Speciation Analysis After 2nd Stage Grinder	LMLA_Bale12_Segment5	cb28b1a0-8251-cf4e-8188-caf0ba6fee5b	LA	LM	cb28b1a0-8251-cf4e-8188-caf0ba6fee5b	0.79	2019-05-23	7a885eba-e618-6d41-a56c-ad8036844c37	9.21	1690.0	6.18	5.65	1.94	9.46	3.02	0.07	0.91	3.01	66.39	0.29	1.33
LT Baseline Ash Speciation Analysis After 2nd Stage Grinder	LMLA_Bale12_Segment6	49d5d3a1-798b-3f4c-9dbb-39733790aa59	LA	LM	49d5d3a1-798b-3f4c-9dbb-39733790aa59	0.839	2019-05-23	b89f5cf9-0bde-044e-9241-9ae552c936a7	8.04	1470.0	6.03	6.0	1.99	10.9	3.16	0.08	0.84	3.14	65.31	0.3	1.47

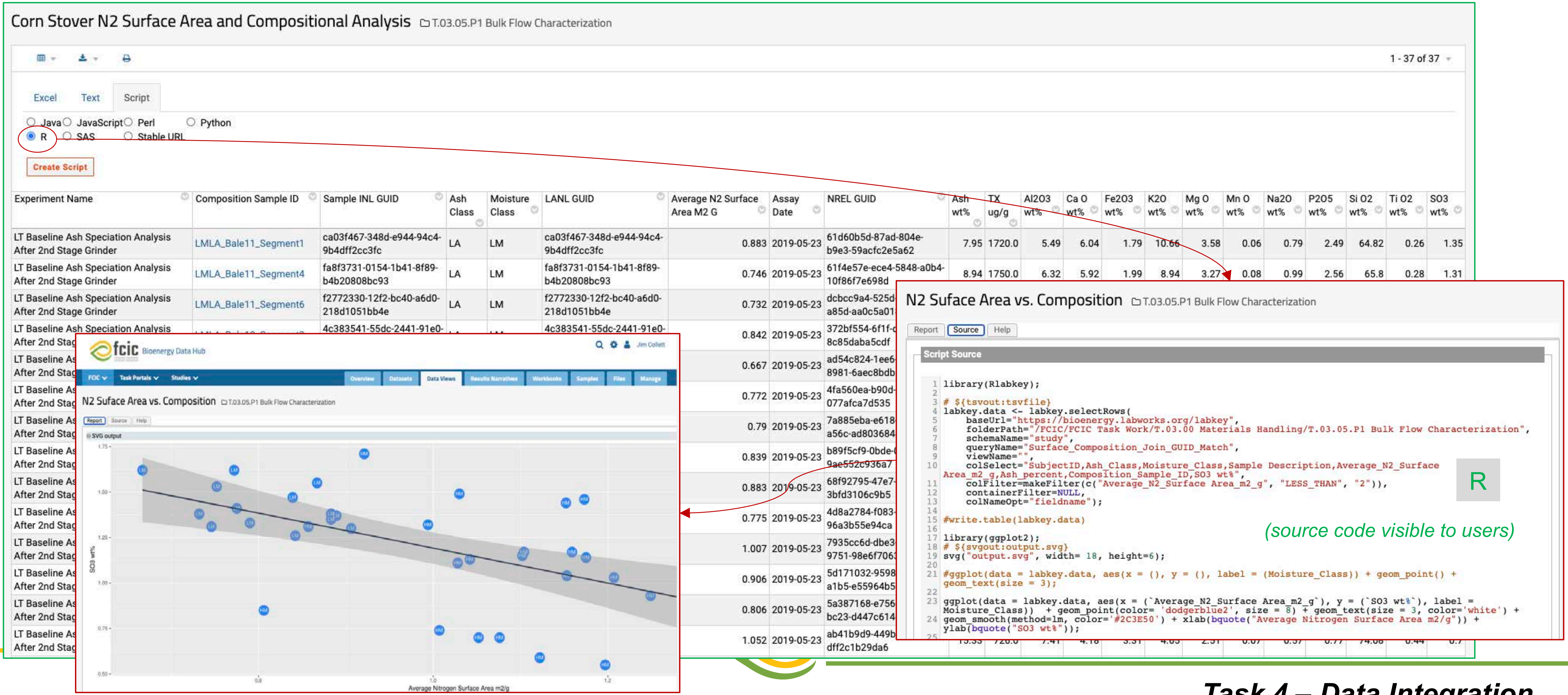


Data exported to Excel

4 – Progress: Data Hub Analytical Tools

Tools

Data Analysis: The Task 4 Team is developing reusable data analysis and automated data processing tools using LabKey's extensive Application Programming Interfaces (APIs) for JMP/SAS, R, Python, Perl, Java, and JavaScript.



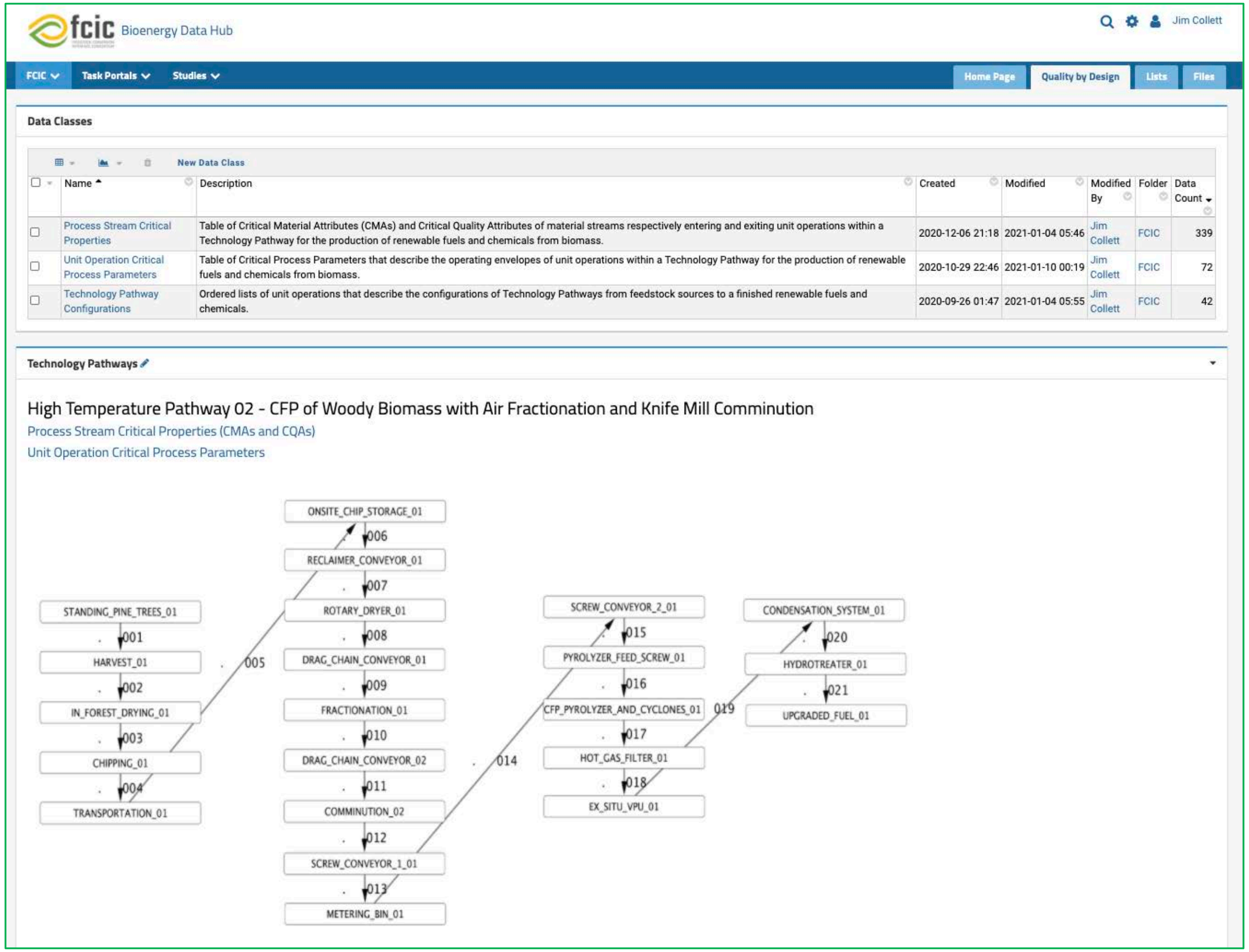
4 – Progress: Quality by Design Data Integration

Tools

fcic FEEDSTOCK-CONVERSION INTERFACE CONSORTIUM

Mapping QbD Properties onto Technology Pathways

- Process Stream Critical Properties
- Unit Operation Critical Process Parameters
- Collaborative QbD property assignment supports:
  - self-organization of metadata into ontologies within the FCIC community.
  - consensus for defining controlled vocabularies and units of measure for use across the growing bioeconomy.



4 – Progress: Quality by Design Data Integration

Tools

fcic
FEEDSTOCK-CONVERSION INTERFACE CONSORTIUM

- A Process Stream Critical Properties data table was created for managing the assignment of QbD critical material properties to feedstocks, intermediates, and products within the High- and Low-Temperature Conversion pathways.
- A related Unit Operation Critical Process Parameters data table was created for assignment of critical process parameters to unit operations within the High- and Low-Temperature Conversion technology pathways.
- Subtask teams and the FCIC PI have reached consensus on a first batch of 164 input material attributes, 85 output quality attributes, and 79 process parameters; these data are now being harmonized and uploaded to the database.

1 - 12 of 12

default This grid view has been modified. Revert Edit Save

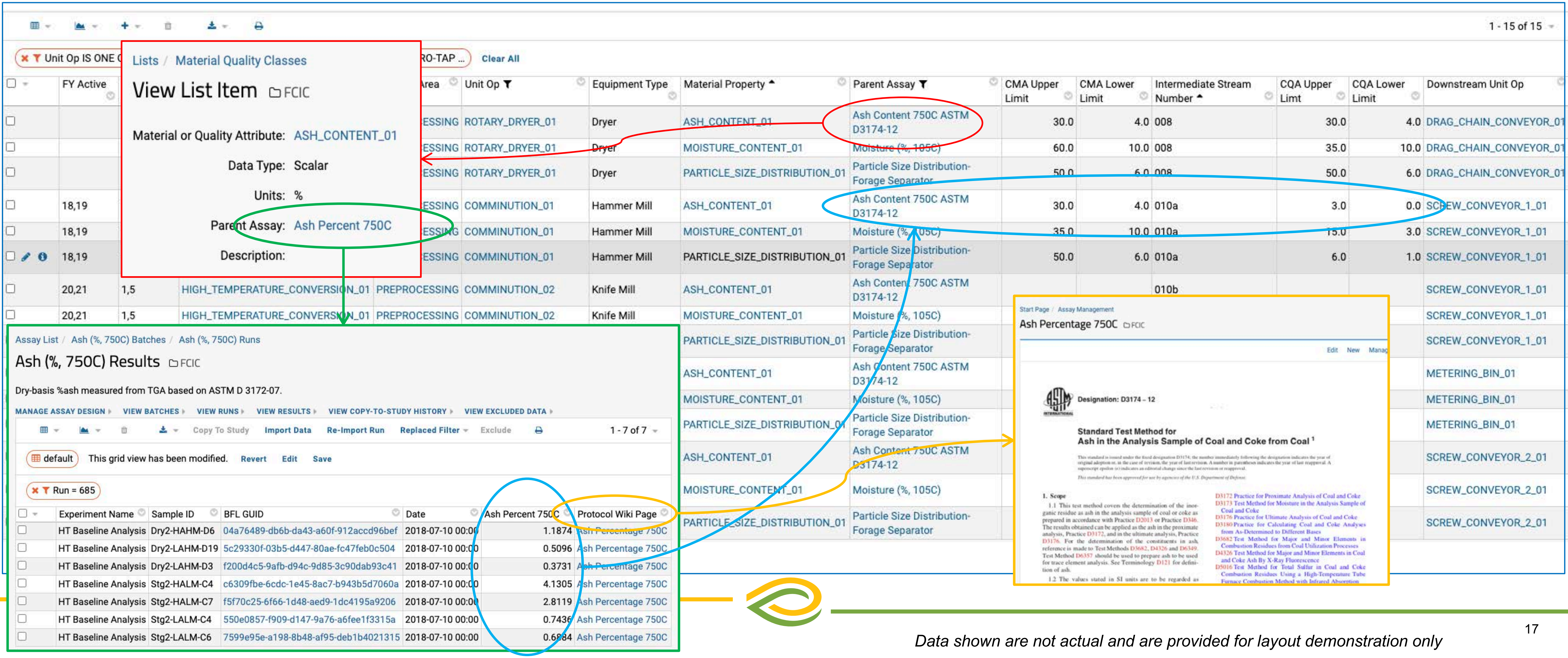
Material Property IS ONE OF (MOISTU… Unit Op IS ONE OF (DRAG_CHAIN_CON… Clear All

Tech Pathway	Unit Op	Intermediate Stream	Intermediate State	Material Property	Units	CMA Upper Limit	CMA Lower Limit	CQA Upper Limit	CQA Lower Limit	Downstream Unit Op	Critical Downstream Unit Op
HIGH_TEMPERATURE_CONVERSION_02	DRAG_CHAIN_CONVEYOR_01	009	Chips	ASH_CONTENT_01	%	30.0	4.0	30.0	4.0	FRACTIONATION_01	COMMINUTION_01
HIGH_TEMPERATURE_CONVERSION_02	DRAG_CHAIN_CONVEYOR_01	009	Chips	PARTICLE_SIZE_DISTRIBUTION_01	mm	50.0	6.0	50.0	6.0	FRACTIONATION_01	COMMINUTION_01
HIGH_TEMPERATURE_CONVERSION_02	DRAG_CHAIN_CONVEYOR_01	009	Chips	MOISTURE_CONTENT_01	%	35.0	10.0	35.0	10.0	FRACTIONATION_01	COMMINUTION_01
HIGH_TEMPERATURE_CONVERSION_02	FRACTIONATION_01	010	Chips	PARTICLE_SIZE_DISTRIBUTION_01	mm					DRAG_CHAIN_CONVEYOR_01	
HIGH_TEMPERATURE_CONVERSION_02	FRACTIONATION_01	010	Chips	ASH_CONTENT_01	%					DRAG_CHAIN_CONVEYOR_01	
HIGH_TEMPERATURE_CONVERSION_02	FRACTIONATION_01	010	Chips	MOISTURE_CONTENT_01	%					DRAG_CHAIN_CONVEYOR_01	
HIGH_TEMPERATURE_CONVERSION_02	DRAG_CHAIN_CONVEYOR_01	011	Chips	PARTICLE_SIZE_DISTRIBUTION_01	mm	50.0	6.0		1.0	COMMINUTION_02	PYROLYZER_FEED_SCREW_01
HIGH_TEMPERATURE_CONVERSION_02	DRAG_CHAIN_CONVEYOR_01	011	Chips	ASH_CONTENT_01	%	30.0	4.0		0.0	COMMINUTION_02	PYROLYZER_FEED_SCREW_01
HIGH_TEMPERATURE_CONVERSION_02	DRAG_CHAIN_CONVEYOR_01	011	Chips	MOISTURE_CONTENT_01	%	35.0	10.0		3.0	COMMINUTION_02	PYROLYZER_FEED_SCREW_01
HIGH_TEMPERATURE_CONVERSION_02	COMMINUTION_02	012	Particles	PARTICLE_SIZE_DISTRIBUTION_01	mm					SCREW_CONVEYOR_1_01	
HIGH_TEMPERATURE_CONVERSION_02	COMMINUTION_02	012	Particles	ASH_CONTENT_01	%					SCREW_CONVEYOR_1_01	
HIGH_TEMPERATURE_CONVERSION_02	COMMINUTION_02	012	Particles	MOISTURE_CONTENT_01	%					SCREW_CONVEYOR_1_01	

Data shown are not actual and are provided for layout demonstration only

4 – Progress: Access to Data for QbD Critical Property Assignment

- The QbD tables support links to experimental data that may be considered in the assignment of sets of critical properties to process intermediate streams or unit operations.



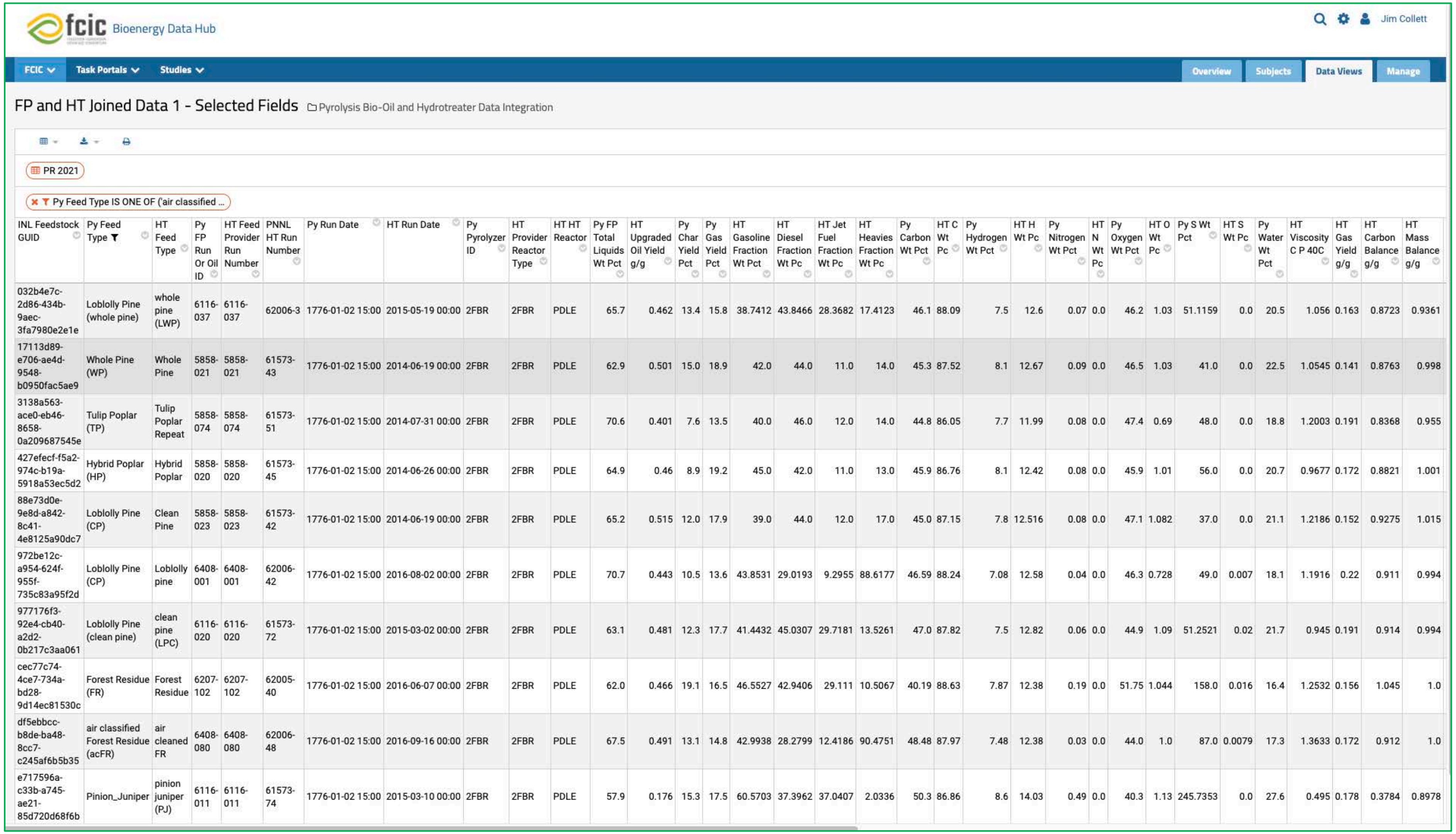
Data shown are not actual and are provided for layout demonstration only

4 – Progress: Experimental Data Integration

Tools

FEEDSTOCK-CONVERSION INTERFACE CONSORTIUM

- Screenshot shows part of a dataset containing 216 records with 224 data columns integrated via a simple SQL query from 3 uploaded data files containing:
  - Feedstock property data from INL
  - Bio-oil property data and pyrolyzer run metadata from NREL
  - Fuel property data and hydrotreater run metadata from PNNL
- Stable, reusable SQL code and user access to source files promotes trust in large datasets on the Data Hub for data mining analysis and machine learning.
- Performing such integrations manually in Excel is time-consuming and error-prone.

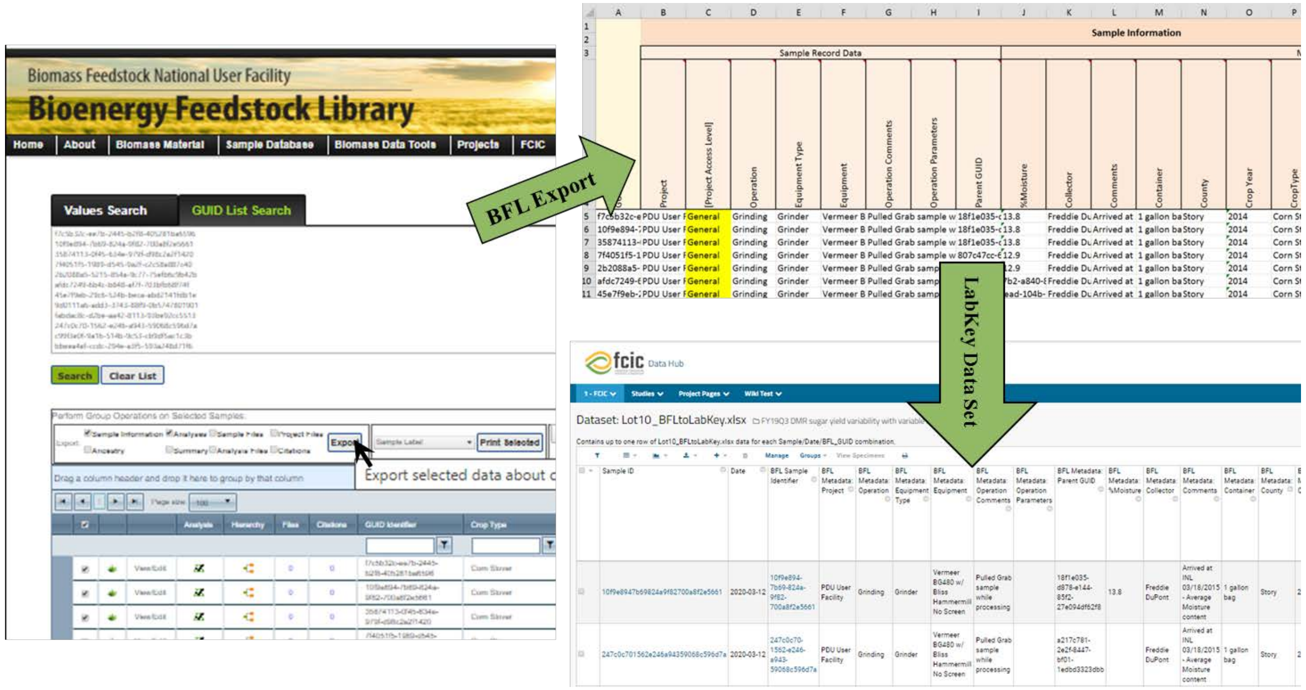


INL Feedstock GUID	Py Feed Type	HT Feed Type	Py FP Run Or Oil ID	HT Feed Provider Run Number	PNNL HT Run Number	Py Run Date	HT Run Date	Py Pyrolyzer ID	HT Provider Reactor Type	HT HT Reactor	Py FP Total Liquids Wt Pct	HT Upgraded Oil Yield g/g	Py Char Yield Pct	Py Gas Yield Pct	HT Gasoline Fraction Wt Pct	HT Diesel Fraction Wt Pc	HT Jet Fuel Fraction Wt Pc	HT Heavies Fraction Wt Pc	Py Carbon Wt Pct	HT C Wt Pc	Py Hydrogen Wt Pct	HT H Wt Pc	Py Nitrogen Wt Pct	HT N Wt Pc	Py Oxygen Wt Pct	HT O Wt Pc	Py S Wt Pct	HT S Wt Pc	Py Water Wt Pct	HT Viscosity C P 40C	HT Gas Yield g/g	HT Carbon Balance g/g	HT Mass Balance g/g
032b4e7c-2d86-434b-9aec-3fa7980e2e1e	Loblolly Pine (whole pine)	whole pine (LWP)	6116-037	6116-037	62006-3	1776-01-02 15:00	2015-05-19 00:00	2FBR	2FBR	PDLE	65.7	0.462	13.4	15.8	38.7412	43.8466	28.3682	17.4123	46.1	88.09	7.5	12.6	0.07	0.0	46.2	1.03	51.1159	0.0	20.5	1.056	0.163	0.8723	0.9361
17113d89-e706-ae4d-9548-b0950fac5ae9	Whole Pine (WP)	Whole Pine	5858-021	5858-021	61573-43	1776-01-02 15:00	2014-06-19 00:00	2FBR	2FBR	PDLE	62.9	0.501	15.0	18.9	42.0	44.0	11.0	14.0	45.3	87.52	8.1	12.67	0.09	0.0	46.5	1.03	41.0	0.0	22.5	1.0545	0.141	0.8763	0.998
3138a563-ace0-eb46-8658-0a209687545e	Tulip Poplar (TP)	Tulip Poplar Repeat	5858-074	5858-074	61573-51	1776-01-02 15:00	2014-07-31 00:00	2FBR	2FBR	PDLE	70.6	0.401	7.6	13.5	40.0	46.0	12.0	14.0	44.8	86.05	7.7	11.99	0.08	0.0	47.4	0.69	48.0	0.0	18.8	1.2003	0.191	0.8368	0.955
427efecf-f5a2-974c-b19a-5918a53ec5d2	Hybrid Poplar (HP)	Hybrid Poplar	5858-020	5858-020	61573-45	1776-01-02 15:00	2014-06-26 00:00	2FBR	2FBR	PDLE	64.9	0.46	8.9	19.2	45.0	42.0	11.0	13.0	45.9	86.76	8.1	12.42	0.08	0.0	45.9	1.01	56.0	0.0	20.7	0.9677	0.172	0.8821	1.001
88e73d0e-9e8d-a842-8c41-4e8125a90dc7	Loblolly Pine (CP)	Clean Pine	5858-023	5858-023	61573-42	1776-01-02 15:00	2014-06-19 00:00	2FBR	2FBR	PDLE	65.2	0.515	12.0	17.9	39.0	44.0	12.0	17.0	45.0	87.15	7.8	12.516	0.08	0.0	47.1	1.082	37.0	0.0	21.1	1.2186	0.152	0.9275	1.015
972be12c-a954-624f-955f-735c83a95f2d	Loblolly Pine (CP)	Loblolly pine	6408-001	6408-001	62006-42	1776-01-02 15:00	2016-08-02 00:00	2FBR	2FBR	PDLE	70.7	0.443	10.5	13.6	43.8531	29.0193	9.2955	88.6177	46.59	88.24	7.08	12.58	0.04	0.0	46.3	0.728	49.0	0.007	18.1	1.1916	0.22	0.911	0.994
977176f3-92e4-cb40-a2d2-0b217c3aa061	Loblolly Pine (clean pine)	clean pine (LPC)	6116-020	6116-020	61573-72	1776-01-02 15:00	2015-03-02 00:00	2FBR	2FBR	PDLE	63.1	0.481	12.3	17.7	41.4432	45.0307	29.7181	13.5261	47.0	87.82	7.5	12.82	0.06	0.0	44.9	1.09	51.2521	0.02	21.7	0.945	0.191	0.914	0.994
cec77c74-4ce7-734a-bd28-9d14ec81530c	Forest Residue (FR)	Forest Residue	6207-102	6207-102	62005-40	1776-01-02 15:00	2016-06-07 00:00	2FBR	2FBR	PDLE	62.0	0.466	19.1	16.5	46.5527	42.9406	29.111	10.5067	40.19	88.63	7.87	12.38	0.19	0.0	51.75	1.044	158.0	0.016	16.4	1.2532	0.156	1.045	1.0
df5ebbcc-b8de-ba48-8cc7-c245af6b5b35	air classified Forest Residue (acFR)	air cleaned FR	6408-080	6408-080	62006-48	1776-01-02 15:00	2016-09-16 00:00	2FBR	2FBR	PDLE	67.5	0.491	13.1	14.8	42.9938	28.2799	12.4186	90.4751	48.48	87.97	7.48	12.38	0.03	0.0	44.0	1.0	87.0	0.0079	17.3	1.3633	0.172	0.912	1.0
e717596a-c33b-a745-ae21-85d720d68f6b	Pinion_Juniper	pinion juniper (PJ)	6116-011	6116-011	61573-74	1776-01-02 15:00	2015-03-10 00:00	2FBR	2FBR	PDLE	57.9	0.176	15.3	17.5	60.5703	37.3962	37.0407	2.0336	50.3	86.86	8.6	14.03	0.49	0.0	40.3	1.13	245.7353	0.0	27.6	0.495	0.178	0.3784	0.8978

4 – Progress: Bioenergy Feedstock Library integration via Python API

Tools

Creation of LabKey Datasets directly from Bioenergy Feedstock Library (BFL) exports



Description

LabKey's Python API was used to develop an automated process for creating Study datasets for batches of sample provenance and analytical data exported from the Bioenergy Feedstock Library (BFL).

Value of new tool

The tool reduces the need to manually create redundant data structures between the two systems and represents an additional method to establish interoperability between the BFL and LabKey.

Potential Customers & Outreach Plan

INL researchers can more easily import samples and data directly into LabKey that are being tracked for the FCIC in the BFL database system.

4 – Progress: Milestones

Task Status Snapshot

- LabKey Studies created for 35 Subtasks
- 30 publications indexed
- QbD database property assignment underway:
  - 164 material attributes
  - 85 quality attributes
  - 79 process parameters
- 119 Data Hub user accounts
- 8 hours of user and developer online training provided to 60 FCIC researchers during FY20

- FY19 Q2: **Initial funding received** for this new FY19-21 Task
- FY19 Q3: **FCIC LabKey Data Hub construction begins** on AWS cloud. LabKey also installed locally behind firewalls at PNNL, INNL, and NREL for internal data integration and workflow development.
- FY19 Q4: **Data from multiple NL teams integrated on Data Hub** within 6 “Pioneer” Subtask Studies; interoperability with **Bioenergy Feedstock Library** database at INL established.

- FY20 Q1: A **Quality by Design roadmap** for FCIC is developed with input from PI and Task leads.
- FY20 Q2: **Preprocessing Area data integration and analysis workflows** constructed within LabKey Assay and Study data structures on the Data Hub.
- FY20 Q3: **Pretreatment Area data integration and analysis workflows** constructed within LabKey Assay and Study data structures on the Data Hub.
- FY20 Q4: **Quality by Design database framework** deployed on Data Hub for assigning critical properties to unit operations and intermediate streams within biorefineries.

- FY21 Q1: **Data Hub Web Portal** ready for phased roll-out with FY18-FY20 reports, curated datasets, and publications organized into 35 Subtask Study folders accessible to a limited group of industry and academic stakeholders for Beta testing and feedback.
- FY21 Q2: **Support Task 8 Case Study 1** by uploading and transforming data generated by FCIC experimental tasks into datasets with harmonized units and variable names to facilitate analysis.
- FY21 Q3: **Support Task 8 Case Study 2** by uploading and transforming data generated by FCIC experimental tasks into datasets with harmonized units and variable names to facilitate analysis.
- FY21 Q4: **Data Hub Web Portal fully open to public** and updated with FY21 reports, curated datasets, and publications.

Summary

Management:

- LabKey Data Hub on AWS cloud is the central collaboration platform for integrating, standardizing, and publicly releasing datasets from the FCIC.
- Development/Admin team at PNNL, INL, and NREL; funding provided for Data Liaisons at other FCIC NLs to attend twice-monthly meetings, FCIC annual meeting, and 8 eight hours of online LabKey training each year.

Technical Approach:

- Each FCIC Subtask has its own LabKey Study folder into which experimental and feedstock characterization data may be uploaded and harmonized to support **FAIR** data sharing, reuse and integrated analysis across the FCIC.
- The Data Hub Web Portal provides faceted search tools to help users quickly down-select to Subtasks covering feedstocks, intermediates, products, unit operations, QbD properties, and publications of interest.

Impact: The QbD database will provide industrial stakeholders with hundreds of well-documented specifications for material properties of biorefinery feedstocks, intermediates, products, and process parameters for unit operations to help to define operating envelopes that ensure reliable biorefinery scale-up and performance.

Progress: Data Hub Web Portal is ready for welcoming 1st round of "beta" users such as the FCIC Industrial Advisory Board, DFO partners, and BETO-supported R&D teams outside of the FCIC. It will be continually updated with data and knowledge products to support new FCIC journal publications, Task 8 Case Studies, and QbD critical property assessments.

Quad Chart Overview- FCIC, Task 4

Timeline

- 10/1/2018 - 9/30/2021

	FY20	Active Project
DOE Funding	$559K	FY19- $795K FY20- $559K FY21- $489K Total- $1,843K

Project Partners (N/A)

Barriers addressed

Ct-N: Multiscale Computational Framework toward Accelerating Technology Development

Ft-E: Feedstock Quality: Monitoring and Impact on Preprocessing and Conversion Performance

ADO-A: Process Integration

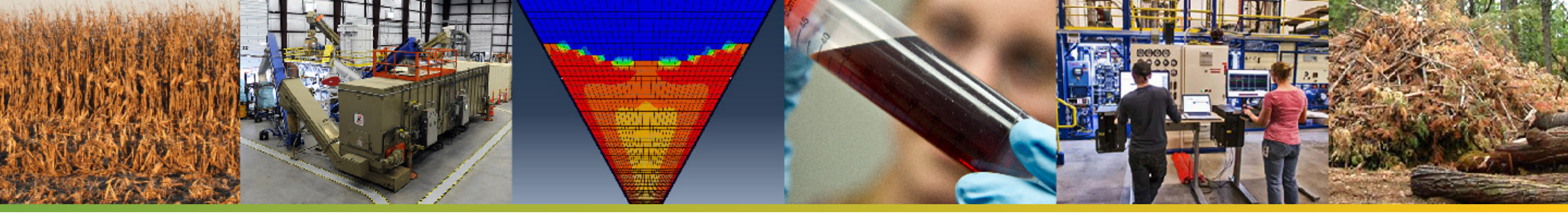
Project Goal

Provide a collaborative computational environment for hypothesis development, experimental and modeling workflow management, integration of datasets and metadata, and deliverables sharing between FCIC subtasks within a uniform Quality by Design framework, as well as a portal for public access to FCIC results, data, and software.

End of Project Milestone

Data Hub Web Portal goes live and opens to industry stakeholders and public with FY18-FY21 reports, curated datasets, and publications.

Funding Mechanism (N/A)



Thank you!

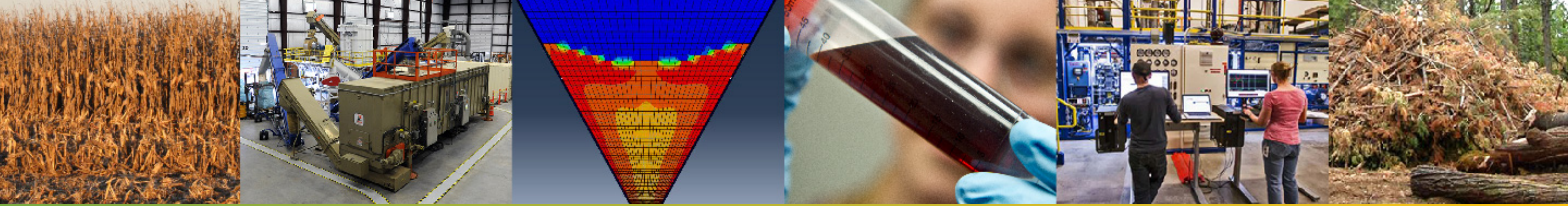
Task 4 Team: David Sievers (NREL), Robert Kinoshita (INL), Shaun O'Leary (PNNL), Matt MacDuff (PNNL), Prerna Prateek (INL), Lorenzo Vega-Montoto (INL), Stan Martin (ORNL), Bruce Wilson (ORNL), Alan Chappell (PNNL)

Data and Design Liaisons: Oslo Jacobson (LBNL), Nick Dylla (ANL), Troy Semelsberger (LANL), Ken Sale (SNL), Steve Phillips (PNNL), Erin Webb (ORNL), Dave Thompson (INL)

LabKey Software Support and Developer Team: Steve Hanson, Hannah Brakke, Adam Rauch, Chet Chopra

energy.gov/fcic

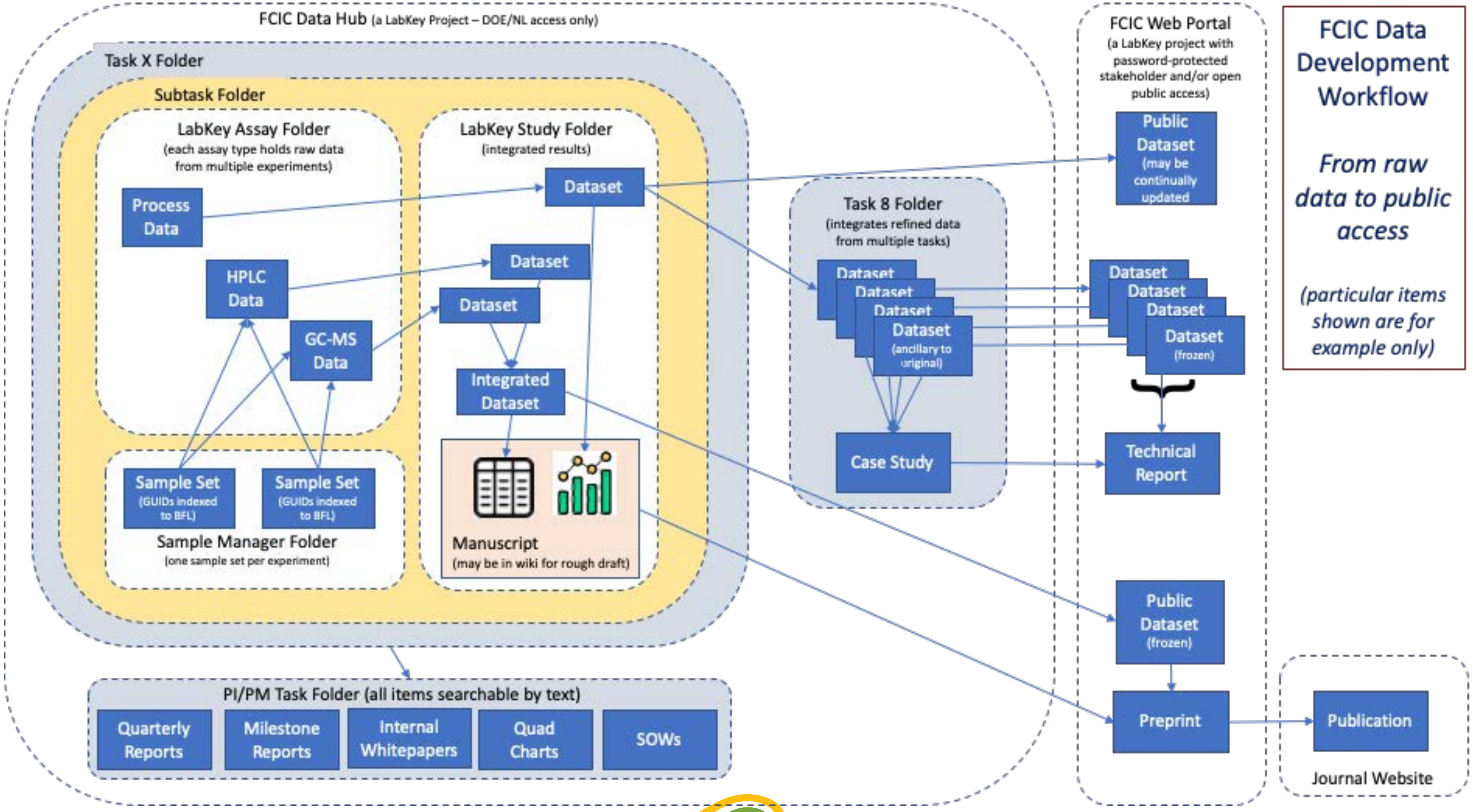
fcic
FEEDSTOCK-CONVERSION INTERFACE CONSORTIUM



Additional Slides

fcic
FEEDSTOCK-CONVERSION
INTERFACE CONSORTIUM

Task 4: FCIC Data Development Workflow



LabKey study standardization

Tools

Standardization of LabKey Study metadata

Studies

Task 5 Studies

1 - 6 of 6

Study	Start Date	Investigator	Biomass Type	Description	Laboratory	Task	Subtask	Unit Operation	Hypothesis	Publication DOI	Report
FY 18 HT Baseline Study	2019-06-05 00:00	INL	Loblolly Pine	This study contains all of the analytical data and INL PDU processing data for the samples collected from the preprocessing operations of the high temperature baseline runs.	INL	2, 5		INL PDU Grinders INL PDU Dryers		10.1021/acssuschemeng.9b06718	
FY18 LT Baseline Study	2019-05-23 00:00	INL	Corn Stover	This study contains all of the analytical data and INL PDU processing equipment data for the samples collected from the preprocessing operations of the low temperature baseline runs.	INL	2, 5		Grinders			
FY19 CQA HT Pathway Identification Study	2020-01-28 00:00	Vicki Thompson	Loblolly Pine	In this milestone, critical quality attributes (CQAs) for high-temperature pathways are identified. These CQAs are the final results from a yearlong effort to gather information on biomass material attributes and how they interact with preprocessing operations to produce quality attributes which are passed to downstream conversion operations. This milestone draws on FY18 baseline runs, FY19 experimental data, historical INL BFNUF data collected, NREL experimental data collected for both high and low temperature conversion processes and from a search of published data that describes the effects of biomass variation on both high and low temperature conversion processes. To provide context for each of the CQAs identified, the critical material attributes (CMAs) that impact those CQAs are presented as well as critical process	INL	5	5.2	HT conversion pathways	CQAs can be identified for various HT conversion pathways through consensus of conclusion drawn		FY19 Q4 Task 5

Investigator | Biomass Type | Description | Laboratory | Task | Subtask | Unit Operation | Hypothesis | Publication DOI | Report

Description

Creation of FCIC experimental study specific metadata parameters collected about each Study

Value of new tool

Can potentially facilitate higher level data queries and harmonization.

Potential Customers & Outreach Plan

FCIC researchers and leadership. This same concept will be applied to the Public Facing Web Portal (FY21) in order to standardize various studies and research hosted through this database.

Database Standardization & Streamlining

Tools

fcic FEEDSTOCK-CONVERSION INTERFACE CONSORTIUM

- **LabKey Core Data Infrastructure Improvements**

```
SELECT Data.sample_id,
Data.parent_id,
Data.Date,
Data.analyte,
MAX(Data.value) AS value
FROM Data
GROUP BY Data.sample_id,
Data.parent_id,
Data.Date,
Data.analyte
PIVOT value BY analyte
```

Analytical Assay Results (long format)

Sample ID	Parent sample ID	Date	Analyte	Parameter	Value	Unit
37795		2019-12-26 00:00	acetic_acid		3.07	g/L
37798		2019-12-26 00:00	acetic_acid		3.04	g/L
37801		2019-12-26 00:00	acetic_acid		2.98	g/L
37804		2019-12-26 00:00	acetic_acid		2.95	g/L

Pivoted Results (filtered for one sample ID)

			Analyte																
Sample ID	Parent sample ID	Date	acetic_acid	acetoin	arabinose	butyric_acid	ethanol	formic_acid	galactose	glucose	glycerol	lactic_acid	meso_BDO	pH	propionic_acid	soluble_lignin	SS_BDO	total_glucose	total_xylose
37795		2019-12-26 00:00	3.07													10.75		0.95	2.44

Description

Analytical data stored in efficient "long" tabular format consistent with industrial database standards. A specimen can have multiple entries that comprehensively documents each measurement done.

Value of new tool

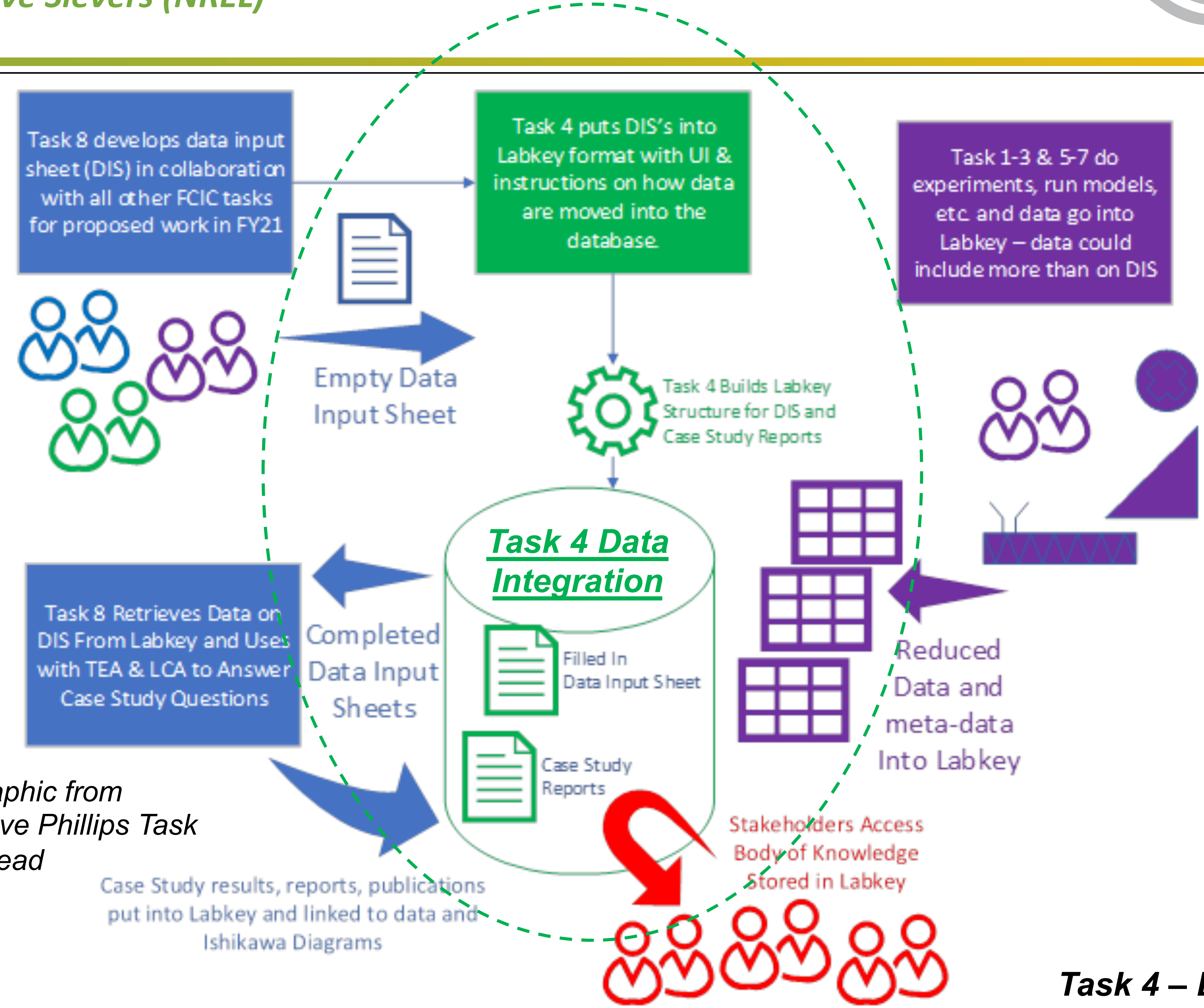
This database format offers future flexibility and documents other data in addition to simple analyte information such as type of measurement, units, and sample lineage.

Potential Customers & Outreach Plan

FCIC researchers benefit from comprehensive analysis documentation. This same database design can be applied within other projects at national labs that use LabKey.

Subtask 4.4: Facilitating Task 8 TEA

Rachel Emerson (INL) and Dave Sievers (NREL)



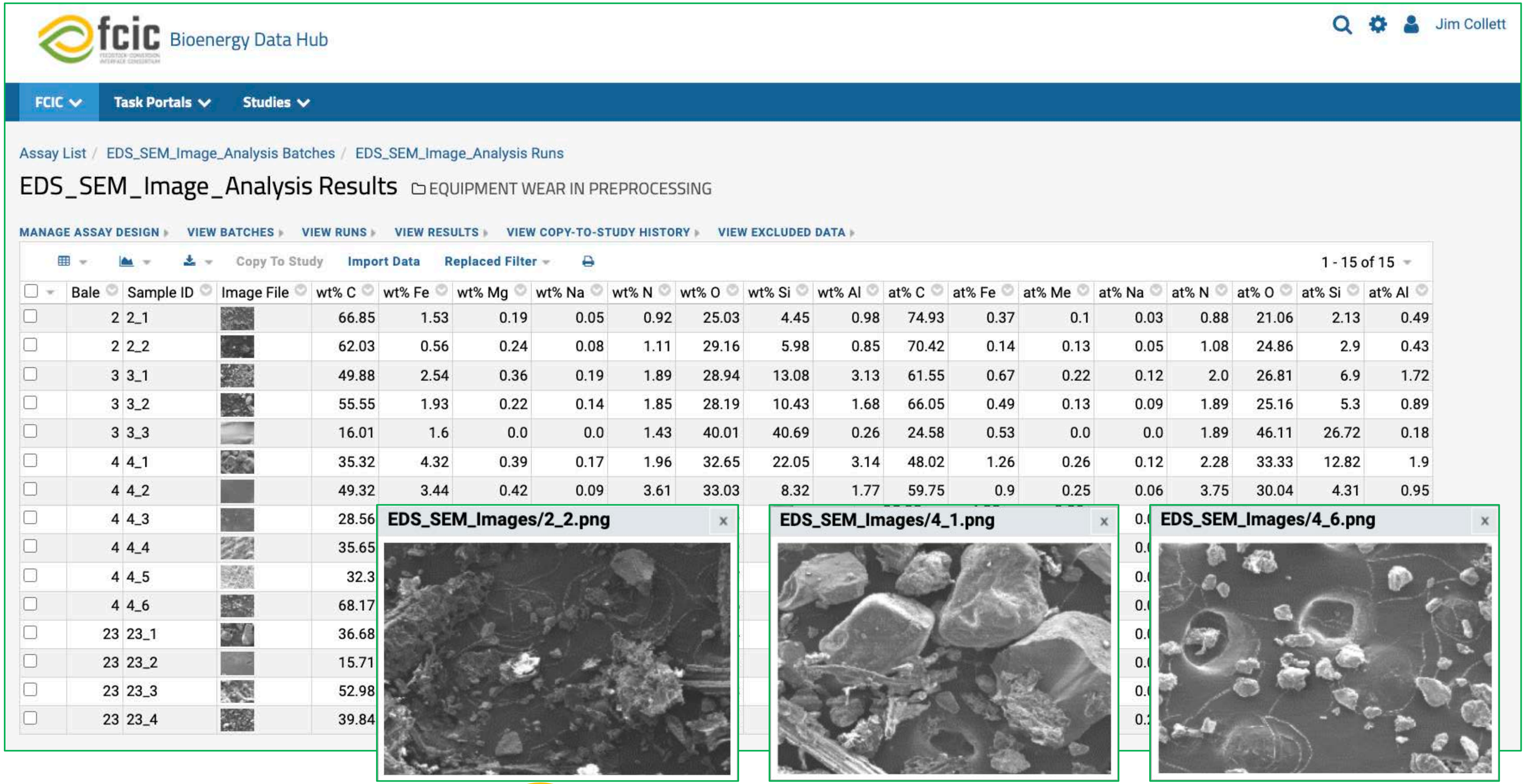
Graphic from Steve Phillips Task 8 Lead

Task 4 – Data Integration

4 – Progress: Assay Data Integration

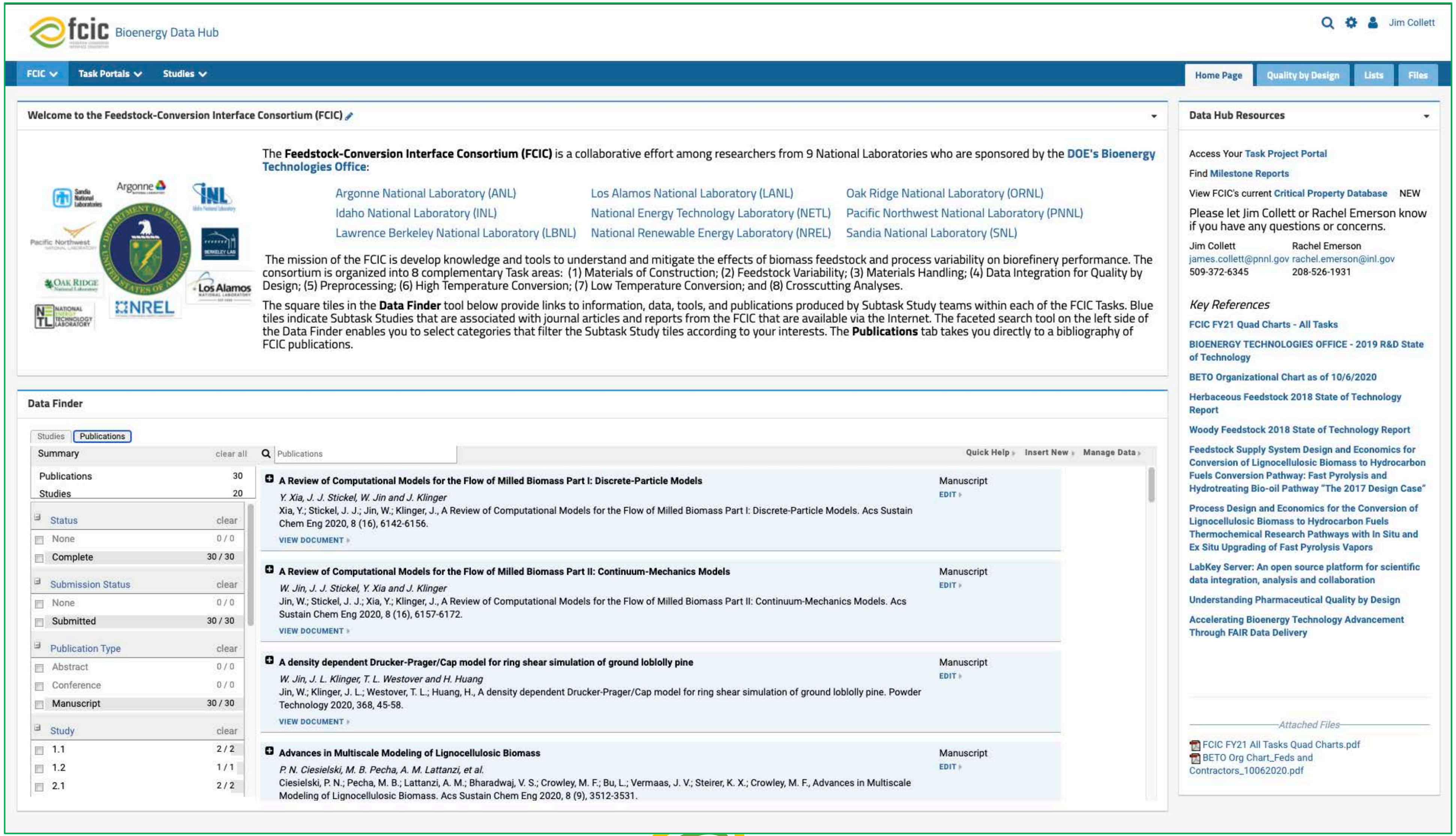
Getting Data into LabKey

- Assay workflows and data structures for uploading experimental data were developed for 6 “Pioneer” Subtasks.
- Stable data table structures and harmonized data types facilitated rapid integration of experimental data.



Bale	Sample ID	Image File	wt% C	wt% Fe	wt% Mg	wt% Na	wt% N	wt% O	wt% Si	wt% Al	at% C	at% Fe	at% Me	at% Na	at% N	at% O	at% Si	at% Al
2	2_1		66.85	1.53	0.19	0.05	0.92	25.03	4.45	0.98	74.93	0.37	0.1	0.03	0.88	21.06	2.13	0.49
2	2_2		62.03	0.56	0.24	0.08	1.11	29.16	5.98	0.85	70.42	0.14	0.13	0.05	1.08	24.86	2.9	0.43
3	3_1		49.88	2.54	0.36	0.19	1.89	28.94	13.08	3.13	61.55	0.67	0.22	0.12	2.0	26.81	6.9	1.72
3	3_2		55.55	1.93	0.22	0.14	1.85	28.19	10.43	1.68	66.05	0.49	0.13	0.09	1.89	25.16	5.3	0.89
3	3_3		16.01	1.6	0.0	0.0	1.43	40.01	40.69	0.26	24.58	0.53	0.0	0.0	1.89	46.11	26.72	0.18
4	4_1		35.32	4.32	0.39	0.17	1.96	32.65	22.05	3.14	48.02	1.26	0.26	0.12	2.28	33.33	12.82	1.9
4	4_2		49.32	3.44	0.42	0.09	3.61	33.03	8.32	1.77	59.75	0.9	0.25	0.06	3.75	30.04	4.31	0.95
4	4_3		28.56															
4	4_4		35.65															
4	4_5		32.3															
4	4_6		68.17															
23	23_1		36.68															
23	23_2		15.71															
23	23_3		52.98															
23	23_4		39.84															

Data Finder Publications Tab



Data Hub Wiki Pages

- LabKey Server provides a wide range of "webparts" that enable data owners to customize their environment.
- The screenshot to the right shows an editable Wiki webpart that allows users to integrate text and graphics into narratives to summarize research results.

CMA Identification for TCPDU Pyrolysis Performance FY19Q3 Results Summary

From the FY18 HT baseline runs, seven of the supersacks of loblolly pine chips and residues that were preprocessed at INL and then converted to pyrolysis oils at NREL were labeled for three conversion processes efficiency metrics: 'On Stream', a continuous variable representing the ratio of the time on stream divided by the time on stream plus the downtime; 'Char Removal', a categorical binary variable indicating whether a run needed to be stopped to clean out either cyclones or char bridging; and 'Feedtrain Bridging', a categorical binary variable indicating whether there were bridging problems in the feedtrain during sample processing. The properties of the raw materials were assessed as potential explanatory variables for these three efficiency metrics including: moisture, total ash, ash speciation, elemental C, H, N, O, and S, volatiles, and particle size distribution factors (D10, D50 and D90).

There was not an obvious visual correlation or pattern between any of these factors alone and the three efficiency metrics. When using a multivariate least squares linear regression approach the 'On Stream' metric could be linearly correlated to a combination of three factors: moisture, particle size factor of D50, and total carbon. All three of these factors contribute significantly (considering p

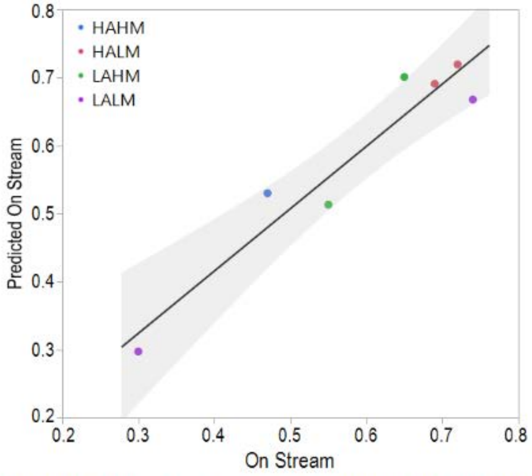


Figure 1. Multivariate linear fit explaining 'On Stream' variability using moisture, particle size factor of D50, and total carbon.

For the 'Char' and 'Feedtrain' efficiency metrics, a Response Screening methodology was used to asses any strongly relating factors; however, as the data was limited, a good fit was not obtained for either metric. It is instructive to look at the factors (Table 1) that did contribute to those models (p

Table 1. Response screening results using a binary response for char (Yes: char formation resulted in shutdown or No: char formation did not result in shutdown) and Feedtrain (Yes, bridging occurred during a run causing shutdown or No: bridging did not occur to an extent requiring shutdown)

Char Response Factor	PValue	Feedtrain Response Factor	PValue
Titanium	0.002	Nitrogen	0.004
D90	0.03	Sodium	0.02
		Halogens	0.03

Attached Files

Publications, Patents, Presentations, Awards, and Commercialization

- Collett, J., *LabKey for Multicenter R&D on Biofuels and Bio-based Products*, in *LabKey User Conference*. 2019: Seattle, WA.