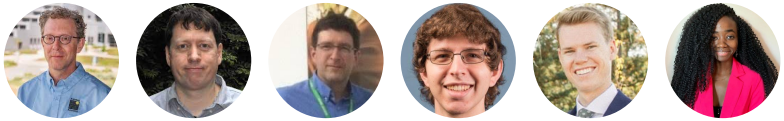
Multiscale Characterization and Multimodal Analysis of Photovoltaic Performance with FAIR Data

Mengjie Li[1], Dylan Colvin[1], Manju Matam[1], Erika I. Barcelos[2], Max Liggett[1], Andrew Ballen[1], Jared Wilson[1], Dominique Yao[2], Philip Knodle[3], Andrew Gabor[3], Greg Horner[4], Craig Neal[1], Sudipta Seal[1], Laura S. Bruckam[2], Roger H. French[2], Hubert Seigneur[1], Kris Davis[1]

[1] University of Central Florida
[2] Case Western Reserve University
[3] BrightSpot Automation
[4] Tau Science

Acknowledgement

Time Series Team: 

Field Inspection Team: 

Module Indoor Characterization: 

Materials Characterization: 

Data Analytics Team: 

PIs: 

SOLAR ENERGY TECHNOLOGIES OFFICE
U.S. Department Of Energy

SETO award number:
DE-EE009347

UCF

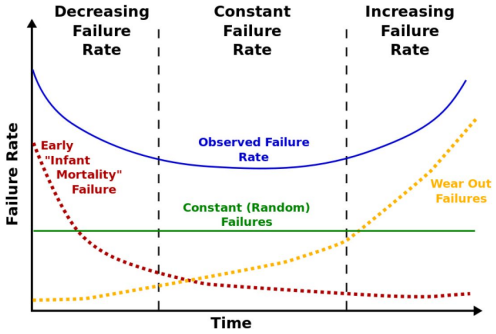
Motivation

Day-to-day

PV modules - only effective when deployed in direct sunlight

Inherently harsh service environment

- **UV** (especially for high efficiency modules)
- **Heat**
- **Humidity**
- Shading*



Extreme weather

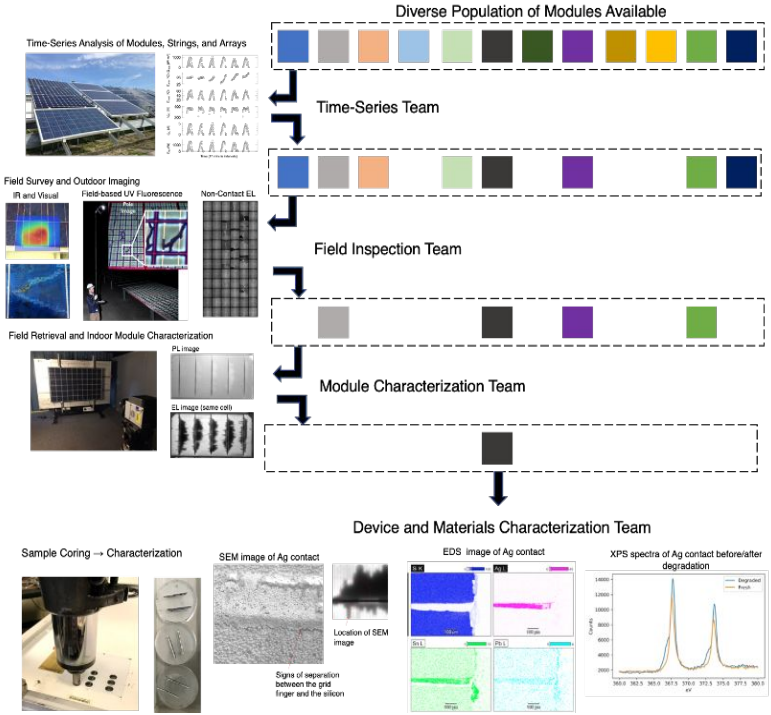
Challenge of severe weather

- **Hail**
- **High Wind**
- **Hurricane**
- Heavy snow
- Cold & Heat shock

UCF

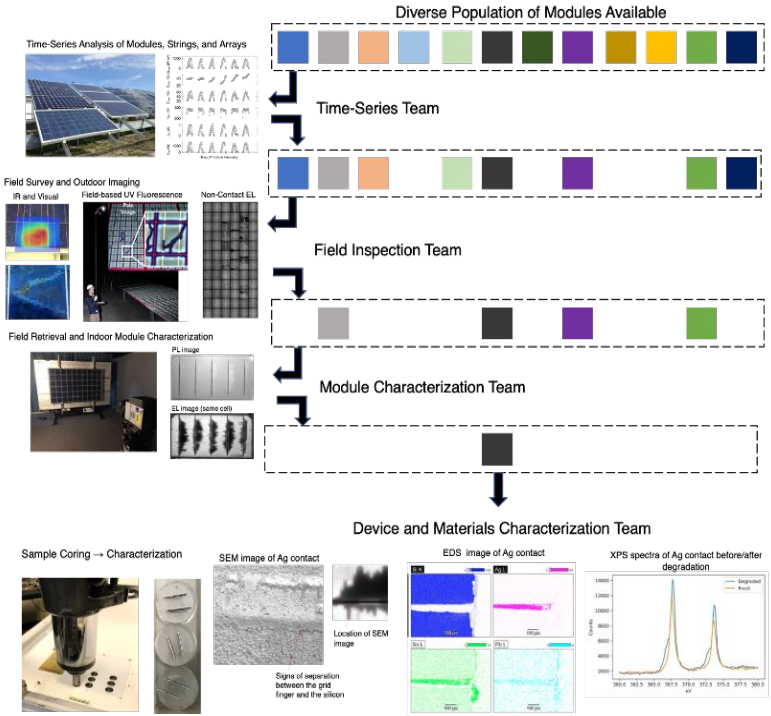
DE-EE0009347: Multiscale Characterization

Project title: Gaining Fundamental Understanding of **Critical Failure Modes** and **Degradation Mechanisms** in **Fielded** Photovoltaic Modules via **Multiscale Characterization**



DE-EE0009347: Multiscale Characterization

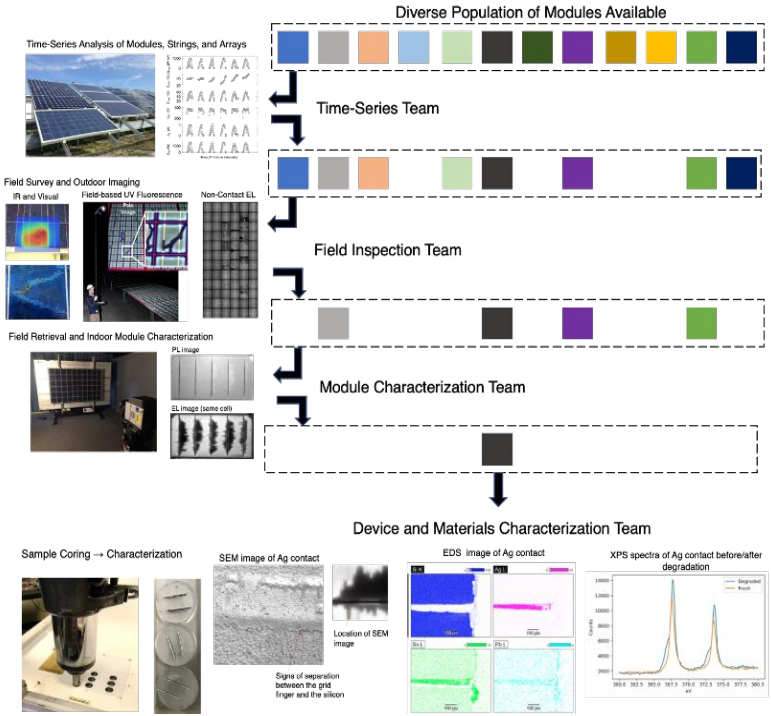
Project title: Gaining Fundamental Understanding of **Critical Failure Modes** and **Degradation Mechanisms** in **Fielded** Photovoltaic Modules via **Multiscale Characterization**



- Time Series: Remote time-series electrical performance and weather data

DE-EE0009347: Multiscale Characterization

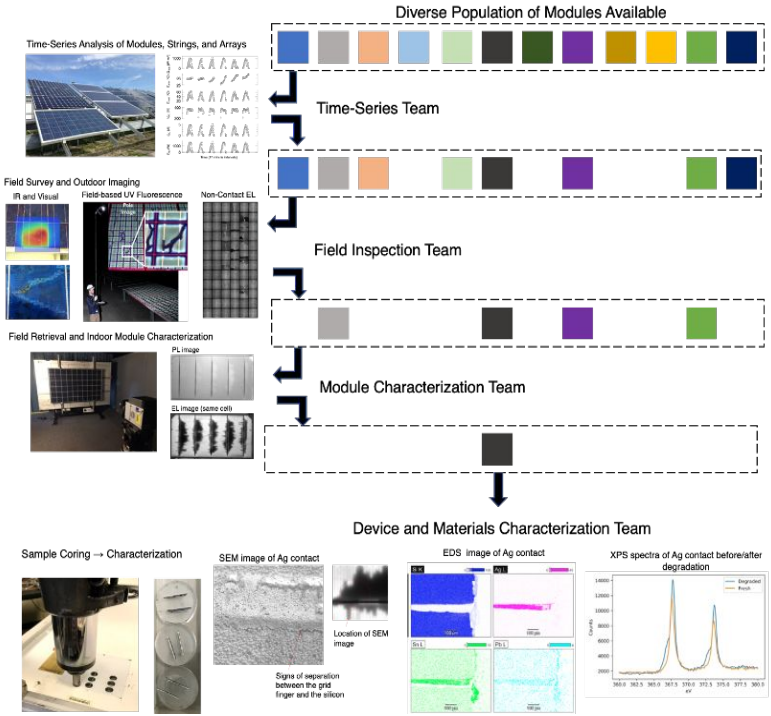
Project title: Gaining Fundamental Understanding of **Critical Failure Modes** and **Degradation Mechanisms** in **Fielded** Photovoltaic Modules via **Multiscale Characterization**



- Time Series: Remote time-series electrical performance and weather data
- Field Inspection :
  - Pole-mounted IR imaging
  - Pole-mounted UVF imaging
  - Drone-based UVF imaging
  - High resolution scanning PL/Non-contact EL

DE-EE0009347: Multiscale Characterization

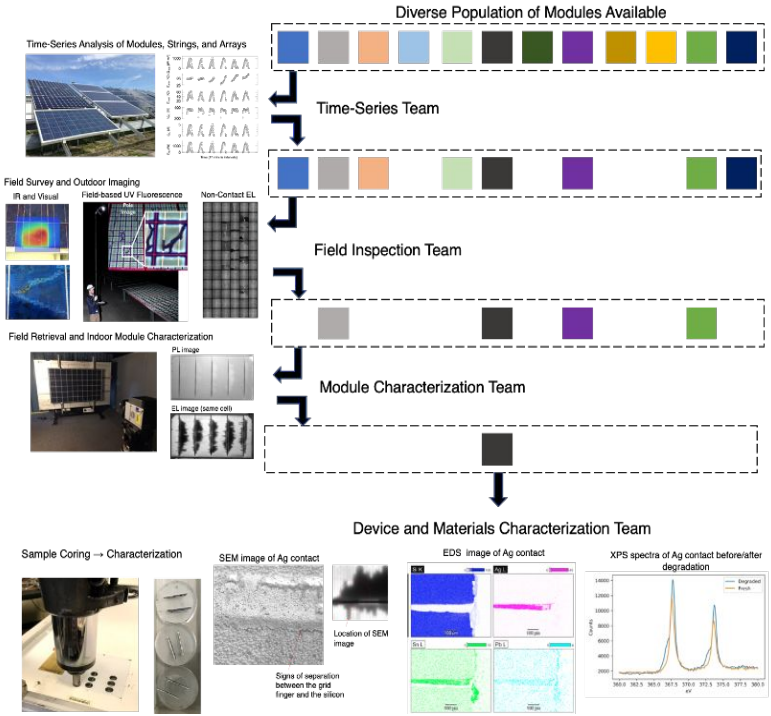
Project title: Gaining Fundamental Understanding of **Critical Failure Modes** and **Degradation Mechanisms** in **Fielded** Photovoltaic Modules via **Multiscale Characterization**



- Time Series: Remote time-series electrical performance and weather data
- Field Inspection :
  - Pole-mounted IR imaging
  - Pole-mounted UVF imaging
  - Drone-based UVF imaging
  - High resolution scanning PL/Non-contact EL
- Indoor Module Characterization
  - *IV* and SunsV_{oc} curve
  - EL imaging
- Materials Characterization
  - SEM
  - XPS

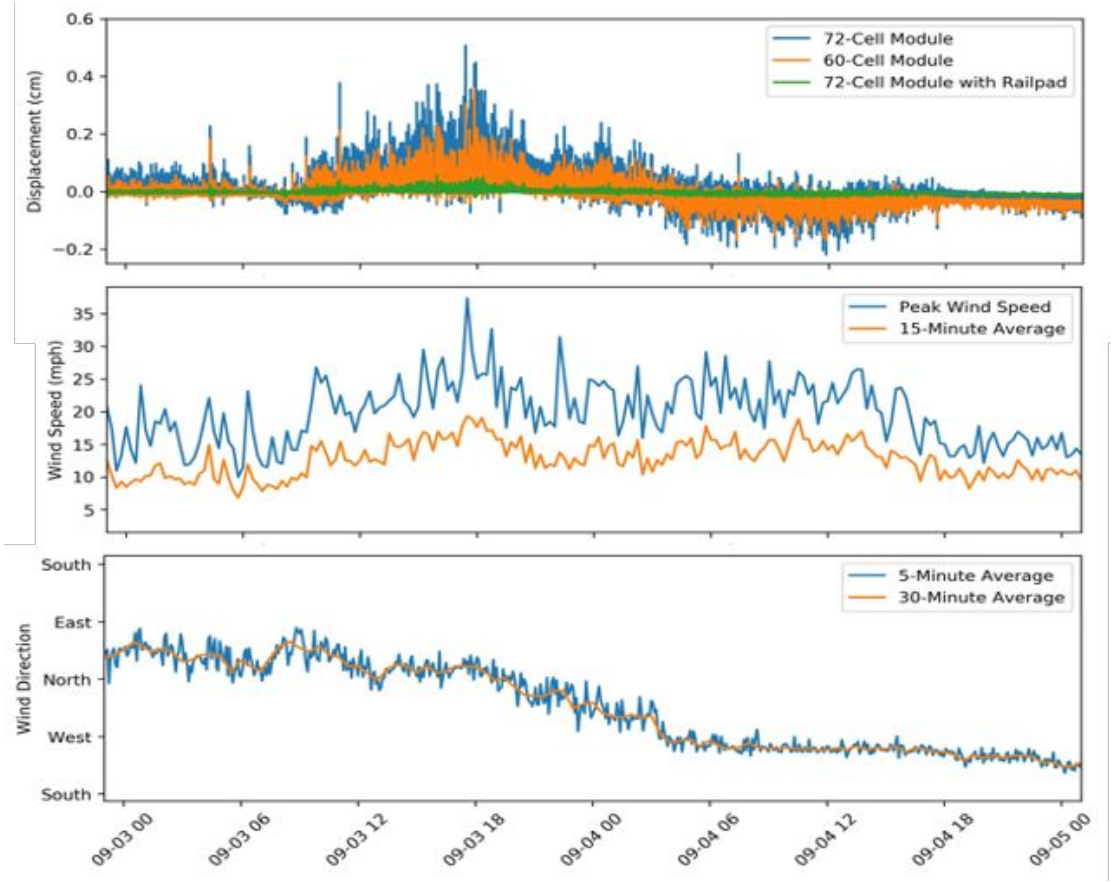
DE-EE0009347: Multiscale Characterization

Project title: Gaining Fundamental Understanding of Critical Failure Modes and Degradation Mechanisms in Fielded Photovoltaic Modules via Multiscale Characterization



- **<u>Time Series: Remote time-series electrical performance and weather data</u>**
- Field Inspection :
  - Pole-mounted IR imaging
  - **<u>Pole-mounted UVF imaging</u>**
  - Drone-based UVF imaging
  - High resolution scanning PL/Non-contact EL
- Indoor Module Characterization
  - *IV* and SunsV_{oc} curve
  - EL imaging
- Materials Characterization
  - SEM
  - XPS

Time Series


Displacement (cm)
0.6
0.4
0.2
0.0
−0.2
72-Cell Module
60-Cell Module
72-Cell Module with Railpad
Wind Speed (mph)
35
30
25
20
15
10
5
Peak Wind Speed
15-Minute Average
Wind Direction
South
East
North
West
South
5-Minute Average
30-Minute Average
09-03 00
09-03 06
09-03 12
09-03 18
09-04 00
09-04 06
09-04 12
09-04 18
09-05 00

Time Series - Overview

SOPO Target			Actual	
Month	Total # of System transferred to CRADLE	Total # of System made public	Total # of System transferred to CRADLE	Total # of System made public
6	6	0	**39**	**0**
18	10	0	**76***	**0**
24	-	10	**147***	**15**
30	15	15		

* FSEC systems: 6, FL Schools & Shelters systems: 141

UCF

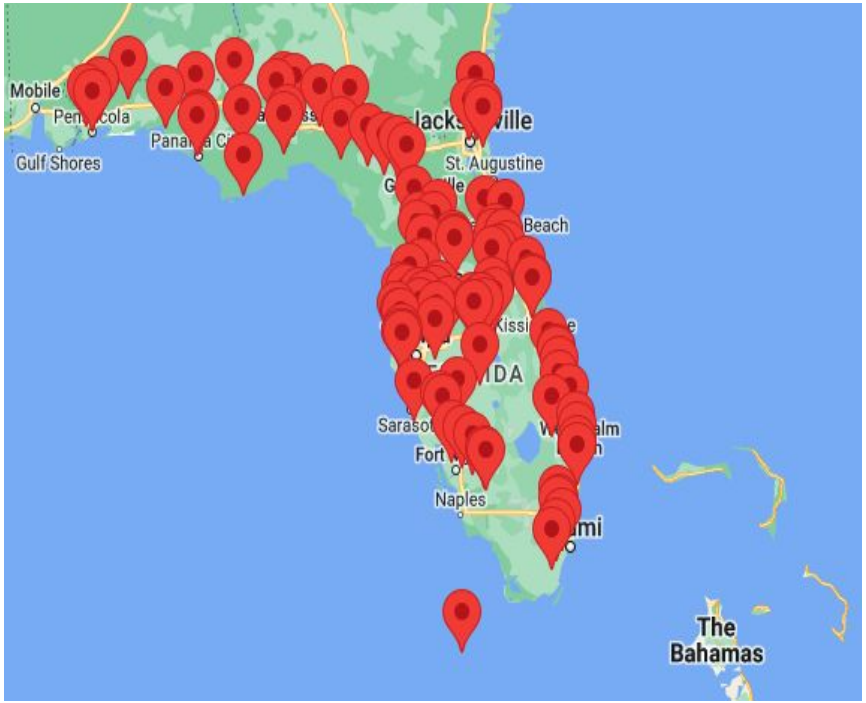
SunSmart Schools - IEA PVPS T13 ST1.4

First to outfit emergency shelter schools
- With solar + storage

More than 115, 10-kW, 3 Inverter PV systems
- Currently installed in schools designated
- Emergency shelters throughout Florida

FSEC, Florida's premier energy research center
- at University of Central Florida,
- coordinated the installations, beginning in 2000

Only open PV + Storage dataset in 29 nations within IEA PVPS!



Explore the interactive map to see the schools with solar power installed on them.

SunSmart Schools - IEA PVPS T13 ST1.4

PV on Schools

- Installation: 2000 - 2004
- 28 systems
- **Over 20 years of data**

SunSmart I

- Installation: 2007 - 2009
- 20 systems
- **Over 15 years of data**

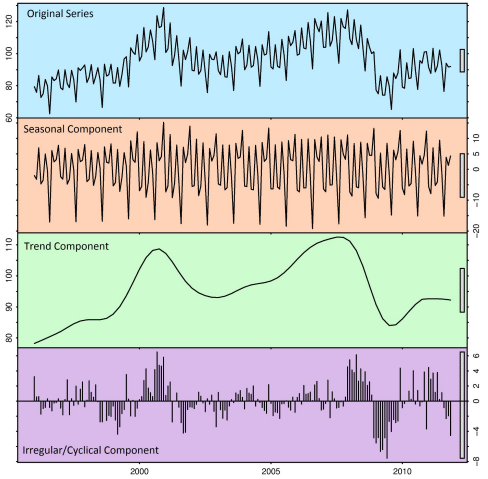
E-Shelter

- Installation: 2012 - 2013
- 90 systems
- **Over 10 years of data**

Time Series

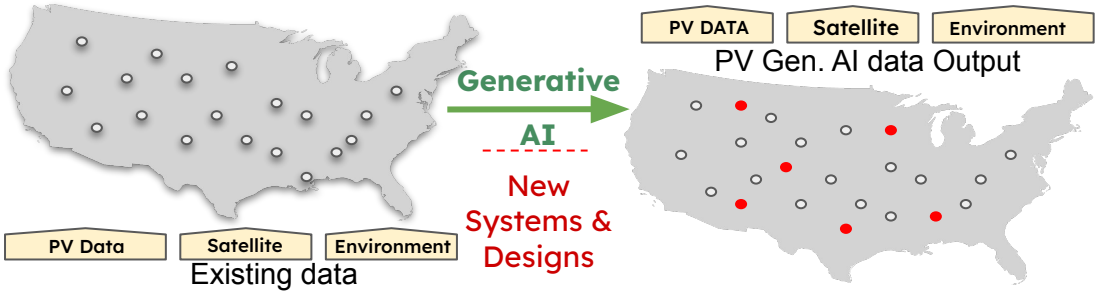
Time Series Analysis Tasks

- PLR
- Anomaly Detection
- Forecasting

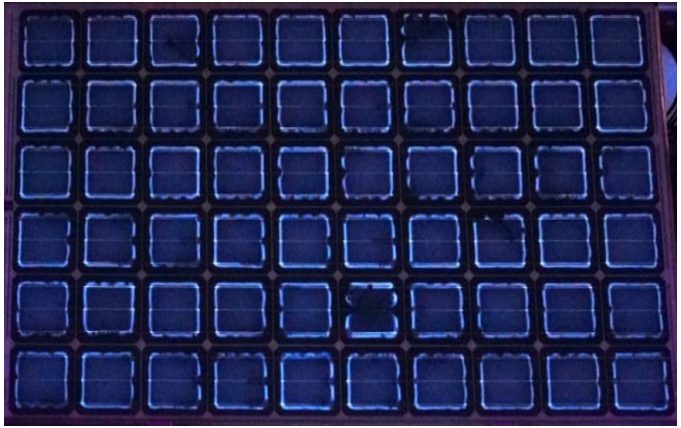


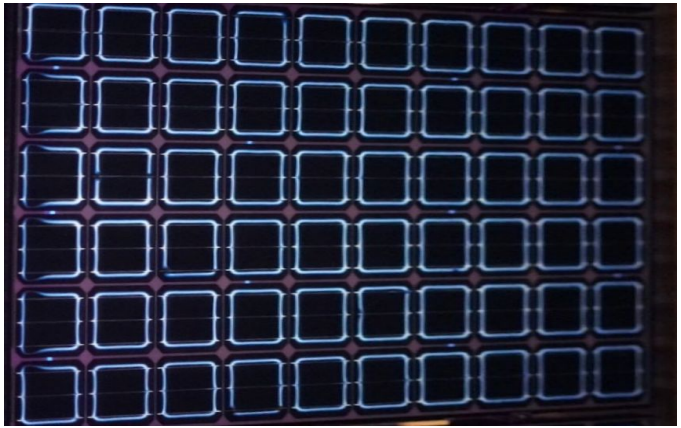
Time Series Data Related Challenges

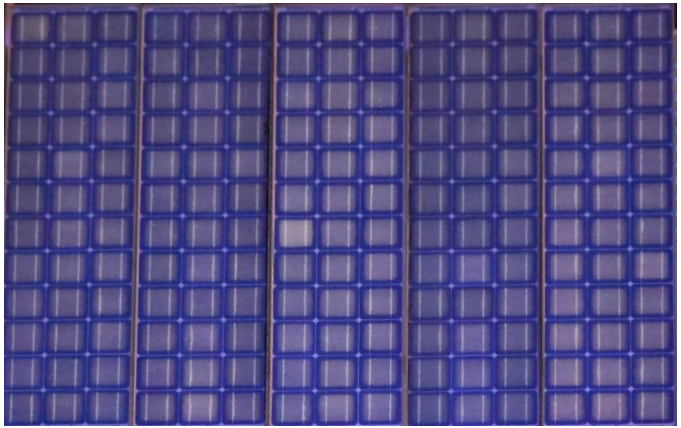
- Missing data
- Outlier
- Length of data
- **Lack of open data**
- **Data management**
- **Proper documentation**
- **Effective data reuse**



UVF Imaging







UVF

FSEC field inspections

- VCAD: 14
- Shell: 160
- ENKI: 30
- SLTE: 15
- LCOE: 24
- PVL: 112
- RTC Suniva: 26
- SunPower: 63
- RTC SunPower: 72
- FGCU: 1,352
- Total: **1,869** imaged modules of 25+ unique module types

CWRU field inspections

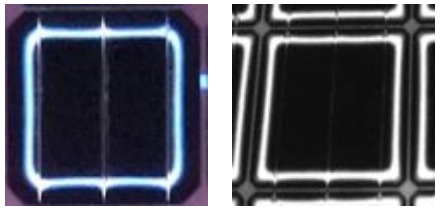
- Test site: 148
- 1MW power plant: 4000
- Total: over **2000** images of 20+ brands

Case Western Reserve University EST. 1826

UCF

UVF - Unique Patterns

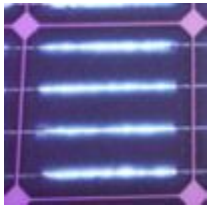
Ring



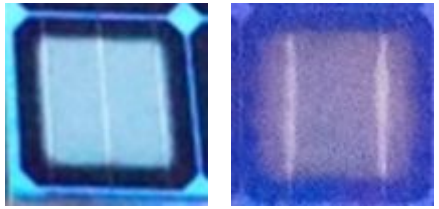
Dark Cracks

Corrosion Delamination

Bright Busbar

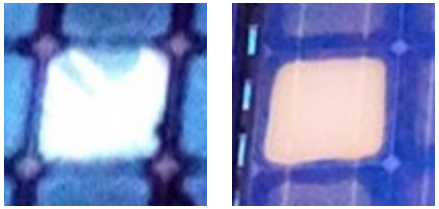


Square

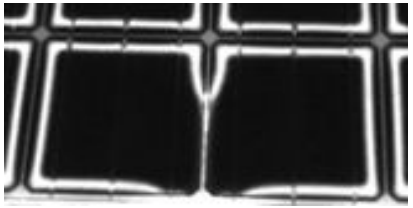


Dark Cracks - Ring

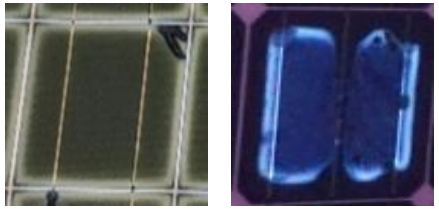
Hot Spots



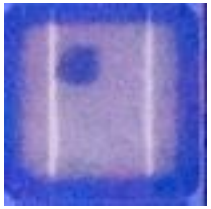
Junction Box Sealing



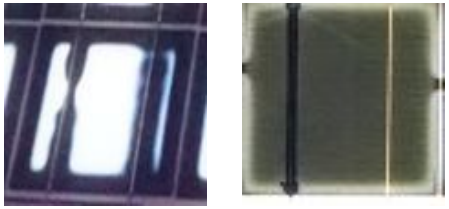
Ring + Square



Dot Cracks

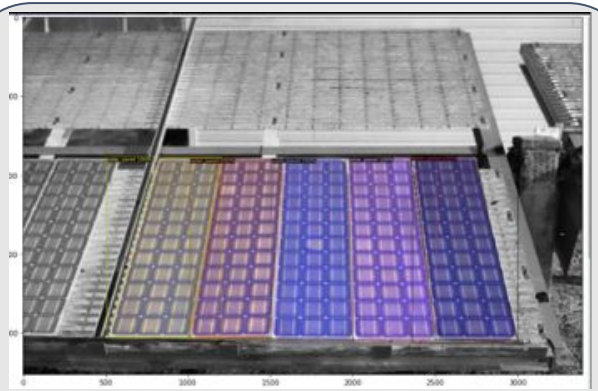


Under Busbar Cracks

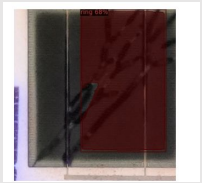


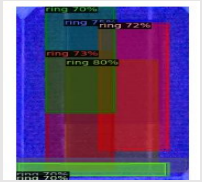
UVF - Image Processing Pipeline

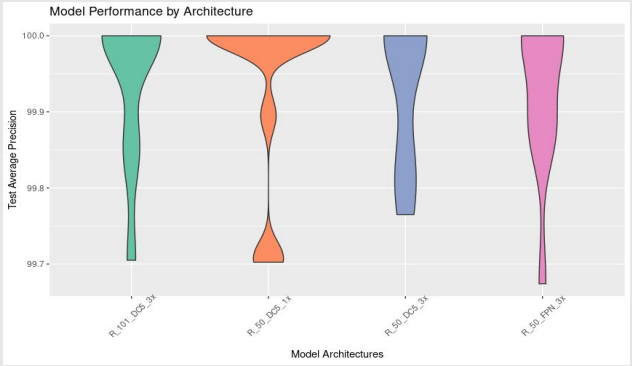
Module Segmentation



Feature Classification







Actionable Next Step

Class	Cracks (%)	Hot Spots(%)	Delamination (%)	Suggested follow up
A	5	1	10	No action
B	5 - 15	1 - 5	10 - 20	Close monitoring
C	15 - 30	5 - 10	20 - 40	Time series cross-validation, visual inspection, IR
D	Above 30	Above 10	Above 40	Further inspection (e. g. with field EL, field IV)

CASE WESTERN RESERVE UNIVERSITY

BrightSpot Automation

UCF

UVF - Advantage & Challenges

Advantages

- Compare to EL imaging, doesn't need to connect modules to power source
- Provide much higher resolution than IR images
- Unique feature detection such as:
  - differentiating pre-existing crack vs. newly formed crack;
  - different BOM;
  - under busbar hidden cracks

Challenges

- Labor
  - Measurement
  - Image processing
- **Lack of open data**
- **Data management**
- **Proper documentation**
- **Effective data reuse**

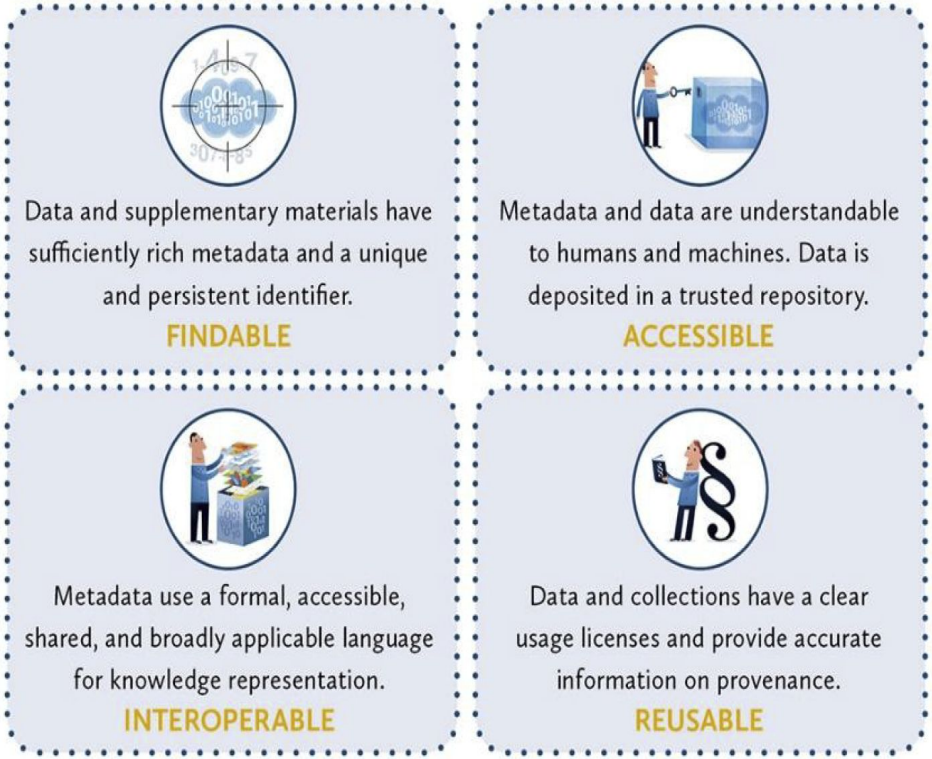
Throughput

- Pole mounted system: ~1500 modules per hour
- Drone based system: ~5000 modules per hour

UCF

FAIR Digital Objects

- **Persistent identifier**
- **Humans** and **machines readable** and **actionable**
- Conform to **community standards**
- Open protocol for **authentication** and **authorization**



Wilkinson M., *et al.* The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* 3, 160018 (2016). https://doi.org/10.1038/sdata.2016.18

Knowledge Graph

What is a knowledge Graph?

- **Nodes:** entities w. types & attributes;
- **Edges:** relations
- capture (factual) knowledge as graphs

Where do KGs come from?

- Structured data: sensors, tables, Wiki infoboxes, databases, social nets, …
- Unstructured data: text, images, videos

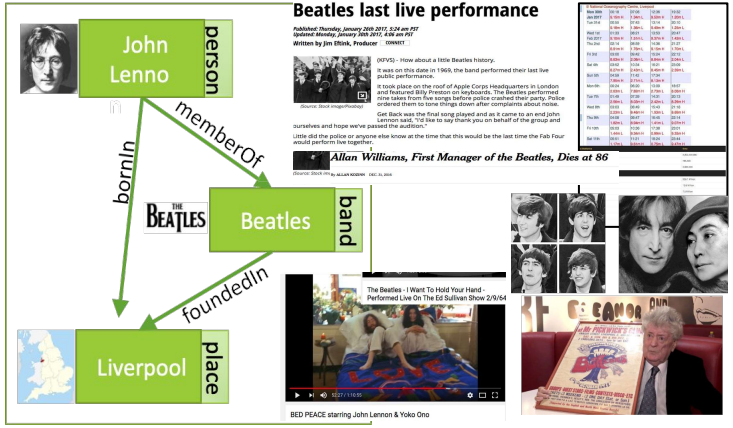
Why (Knowledge) Graphs?

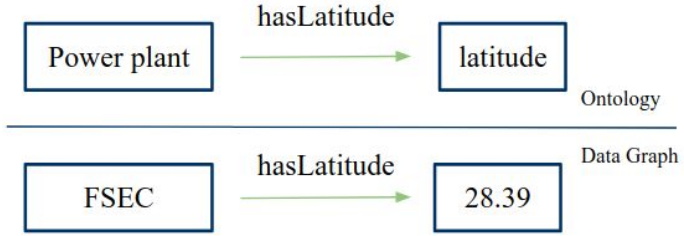
Humans:

- Explore data via intuitive/processible structure
- Combat information overload
- Tool for supporting knowledge-driven tasks

Also:

- Key ingredient for many AI tasks
- Bridge from data to human semantics
- Use decades of work on graph analysis





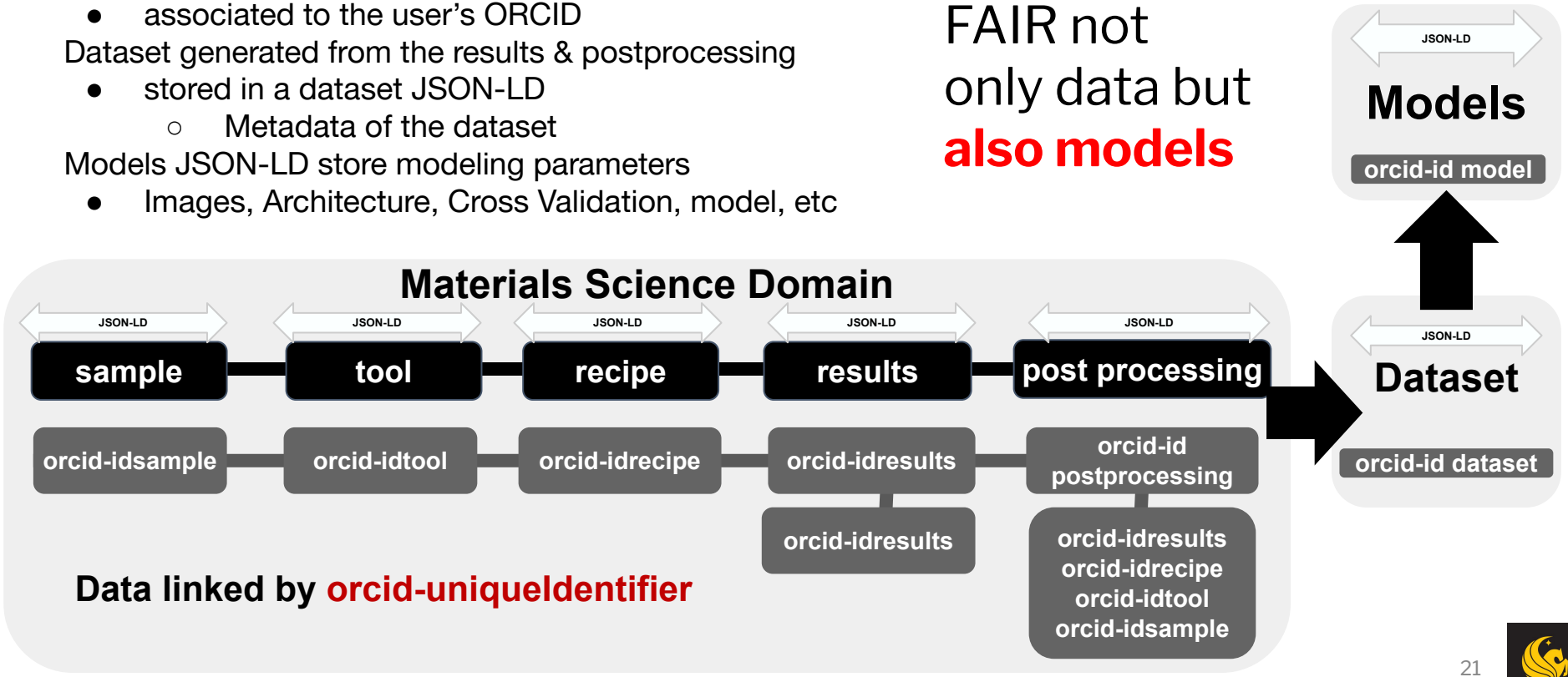
Linking Data in a Domain for Efficient Pipelining & Modeling

Metadata and data are linked by unique ids
- associated to the user's ORCID

Dataset generated from the results & postprocessing
- stored in a dataset JSON-LD
  - Metadata of the dataset

Models JSON-LD store modeling parameters
- Images, Architecture, Cross Validation, model, etc

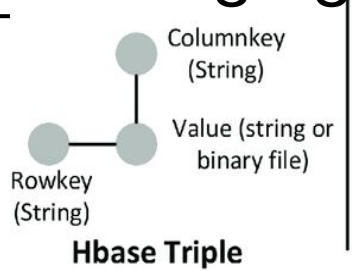
FAIR not only data but **also models**

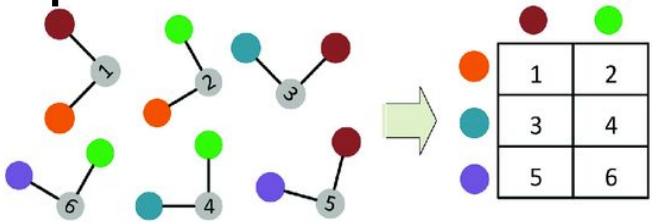


Domain Ontologies - Knowledge graphs

Apache HBase:

- Data Storage and
- Represented in RDF Triples





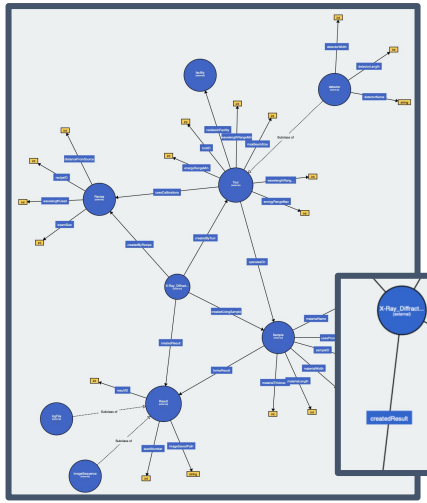
Ontologies created in OWL language

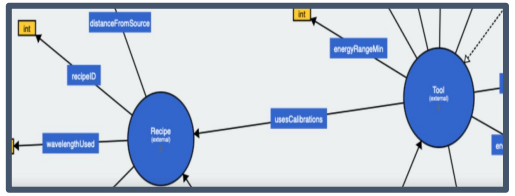
- Builds on top of RDF
- Extends RDF for complex knowledge & reasoning
- Provides a more expressive language
  - And larger vocabulary

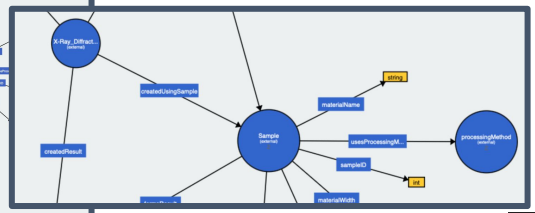
Creation of Ontology-driven Knowledge Graphs

- JanusGraph Distributed Database
  - Scalable graph database optimized for
    - storing and querying graphs
    - containing hundreds of billions
    - of vertices and edges
    - distributed across D/HPC CRADLE

JanusGraph



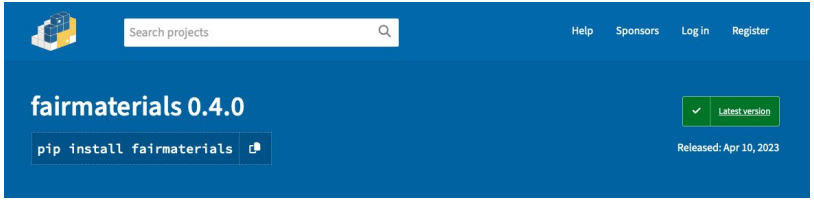


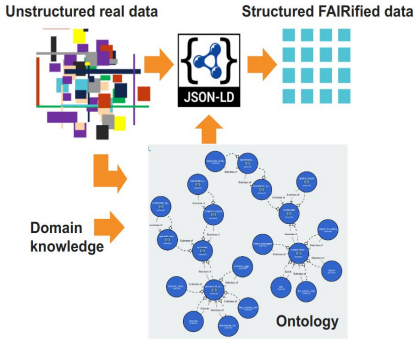


FAIRmaterials

An ontology is a formal dictionary

- **of terms for a given industry or field that**
  - shows how the properties are related
- **Terms are stored as object-relationship pairs**
  - Subject, predicate, and object





FAIRmaterials: Make Materials Data FAIR

We provide here tools used by the Solar Durability and Lifetime Extension Center (SDLE) for FAIRifying data from materials science projects. Functions have been created for numerous tools common in the field in order to make the metadata more Findable, Accessible, Interoperable, and Reproducible.

Version:	0.3.0
Depends:	R (≥ 2.10)
Imports:	stringr, jsonlite, svDialogs, tidyjson, RColorBrewer, igraph, dplyr, graphics, jsonld, utils
Suggests:	knitr, rmarkdown, testthat (≥ 3.0.0)
Published:	2023-02-11

Through relating terms together,

- an ontology can create a densely interconnected web
- of many terms all related to one another

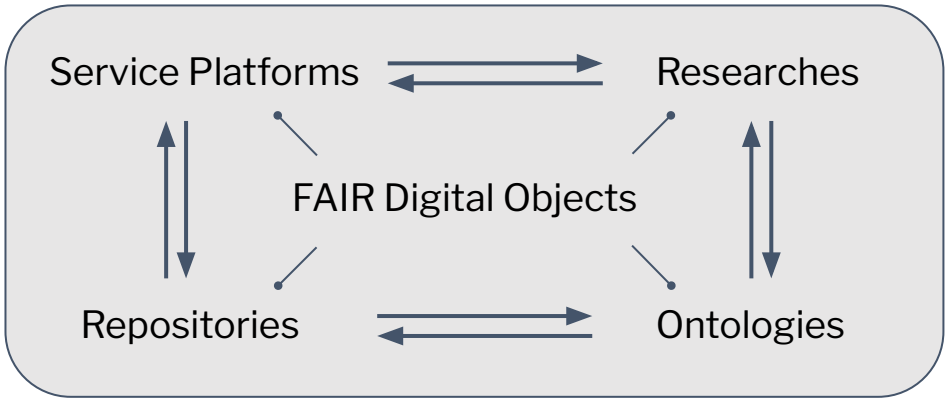
- **fairmaterials**
  - Latest release: Apr 10, 2023
- **2 languages**
  - R and Python
- **32 contributors**
- **42 domain ontologies**
  - Defined by OWL file
  - Documented with json-ld template
- **26 domain vignette examples**
  - How to FAIRify for that domain

Open FAIR Data & FAIR Ecosystem

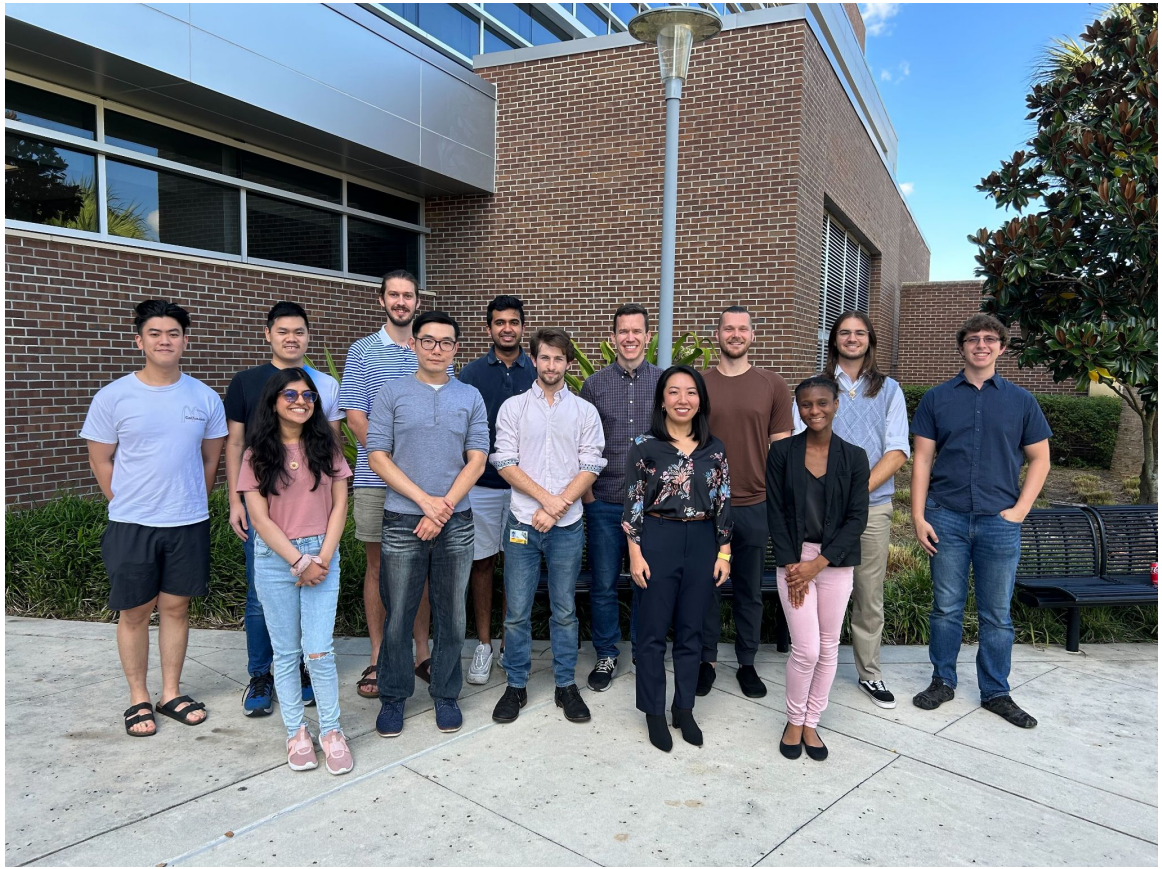
What is the Open Science Framework?

The **Open Science Framework (OSF)** is an open source cloud-based project management platform. It is designed to help teams collaborate in one centralized location; it is also used to share part or all of a research project or its outputs, such as preprints, open access articles, and data. Teams can connect third-party services that they already use (such as GitHub, Google Drive, and Amazon Web Services) directly to the OSF workspace. OSF provides version control, persistent URLs, and DOI registration. Cross-institutional collaboration is easy, as is controlling who has access to projects. OSF also allows researchers to connect to research services, like GitHub.

Name	Modified
Time-Series Data of Photovoltaic Systems Installed in Florida	
OSF Storage (United States)	
BAY.json	2023-04-26 02:55 PM
BAY_inv1.csv	2023-04-26 02:51 PM
BAY_inv2.csv	2023-04-26 02:51 PM
CEL.json	2023-04-26 02:55 PM
CEL_inv1.csv	2023-04-26 02:51 PM
CEL_inv2.csv	2023-04-26 02:51 PM
EHS.csv	2023-04-26 02:51 PM
EHS.json	2023-04-26 02:55 PM
IN2.csv	2023-04-26 02:51 PM
IN2.json	2023-04-26 02:55 PM
IN3.csv	2023-04-26 02:51 PM
IN3.json	2023-04-26 02:55 PM



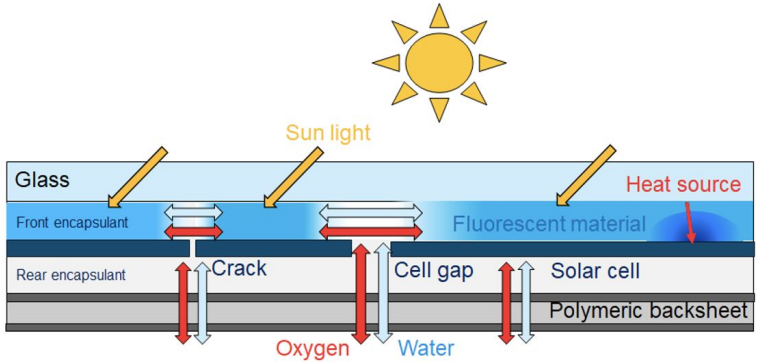
Thanks! Questions?

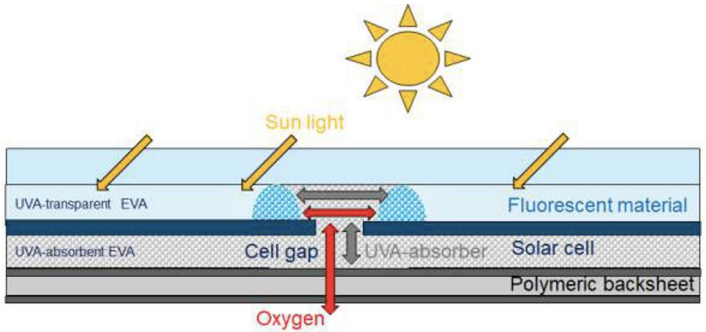


Email:
mengjie.li@ucf.edu

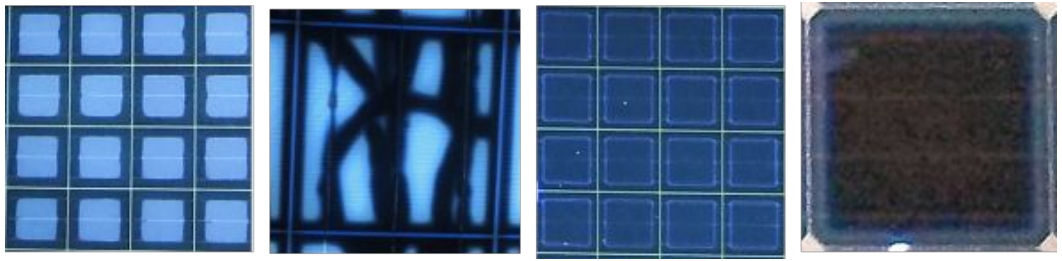
UVF - Mechanism *

- Fluorescent material: Polymer, impurities, cross linker decomposition products and additives and their degradation products (UVA absorbing EVA, UVA absorbing polyolefin)
- Impact factors: temperature, UV dose
- Competing process: photobleaching (requires both oxygen and light)





Square Pattern



Ring Pattern

FAIR Principles

It should be possible for others to discover your data. Rich **metadata** should be available online in a **searchable resource**, and the data should be assigned a **persistent identifier** (e.g. DOI, Handle…)

Lots of documentation is needed to support data interpretation and reuse. It needs to be clear how, why & by whom data were created & processed (**provenance**). The data should conform to **community standards** and be properly **licensed** so others know what kinds of reuse are permitted.

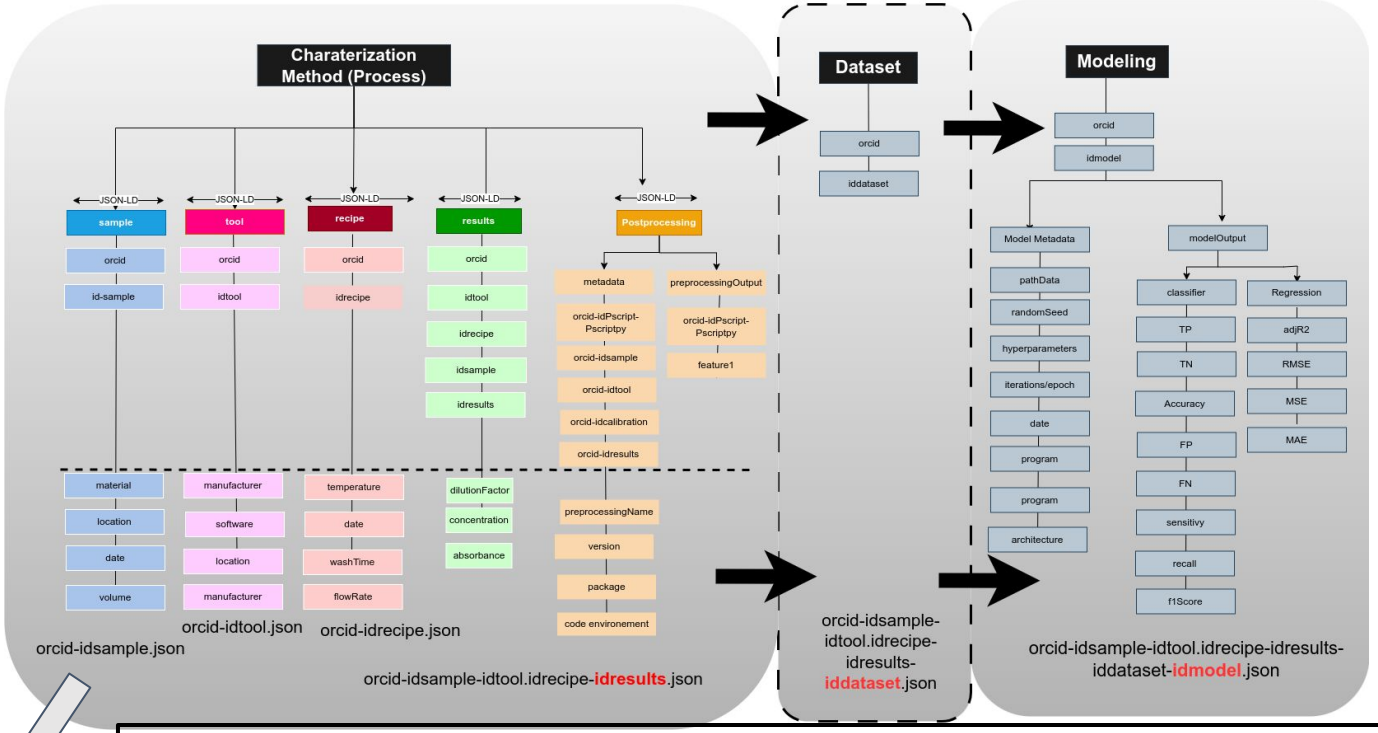
It should be possible for **humans** and **machines** to retrieve your (meta)data by their identifier using standardized communication **protocol**. The protocol should be open, free and universally implementable. The protocol allows for **authentication** and **authorization**, when necessary.

(Meta)data should use a **formal, and broadly recognized standards** to allow them to be combined & exchanged (file formats, metadata schemas, controlled vocabularies, keywords, ontologies, qualified references & links to other related data)

Wilkinson M., *et al.* The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* 3, 160018 (2016). https://doi.org/10.1038/sdata.2016.18

UCF

FAIRmaterials - Ontologies



FAIR not only data but **also models**

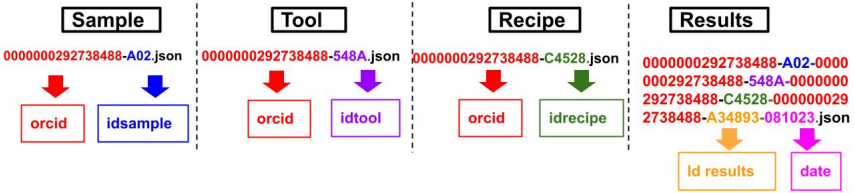
Global unique persistent identifier

Idsample : **User defined id** + **hash**

Ex: 0005-9394-0000-0000-**PVOhio2023**-**24d470987fda1278c63c3f97ab30869b821906449f3ecf290ee48086b8215668**

FAIRmaterials - Ontologies

How to link JSON-LDs in a domain



In BFO, **all entities are divided into continuants and occurrents.** BFO adopts a view of reality as comprising
(1) **continuants,** entities that continue or persist through time, such as objects, qualities, and functions - objects we see around us every day.
(2) **occurrents**, processes conceived as extended through (or spanning) time the events or happenings in which continuants participate.