

eXtremeMAT

accelerating the development of extreme environment materials



U.S. DEPARTMENT OF
ENERGY

Extreme Environment Materials - Data Analytics

Ram Devanathan, Jennifer R. Bauer, Gary D. Black, Michael Gao, Michael Glazoff, Jeffrey A. Hawk, Carina Lansing, Thomas Lograsso, Lianshan Lin, Turab Lookman, Pratik Ray, Vyacheslav Romanov, Kelly Rose, Arun Sathanur, Dongwon Shin, and Yuki Yamamoto

2018 Review Meeting For Crosscutting Research, Pittsburgh, PA
April 10, 2018

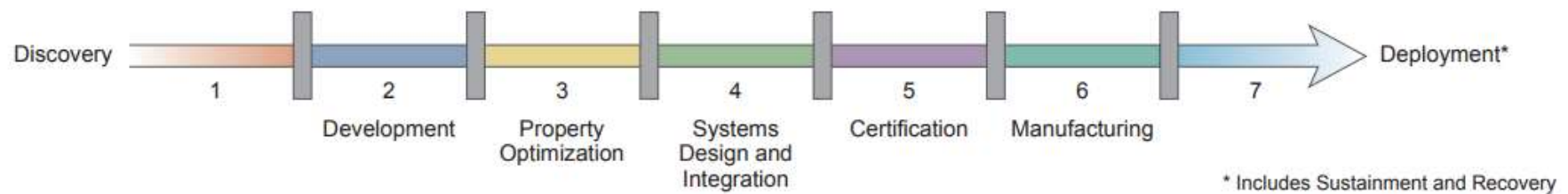


eXtremeMAT

accelerating the development of extreme environment materials



Traditional empirical materials development



Traditional materials development takes 10-20 years (source: OSTP MGI White Paper, 2011)

Empirical lifetime prediction is unreliable and not transferable to new alloys.

Solution: Use data management and analytics to integrate materials development.



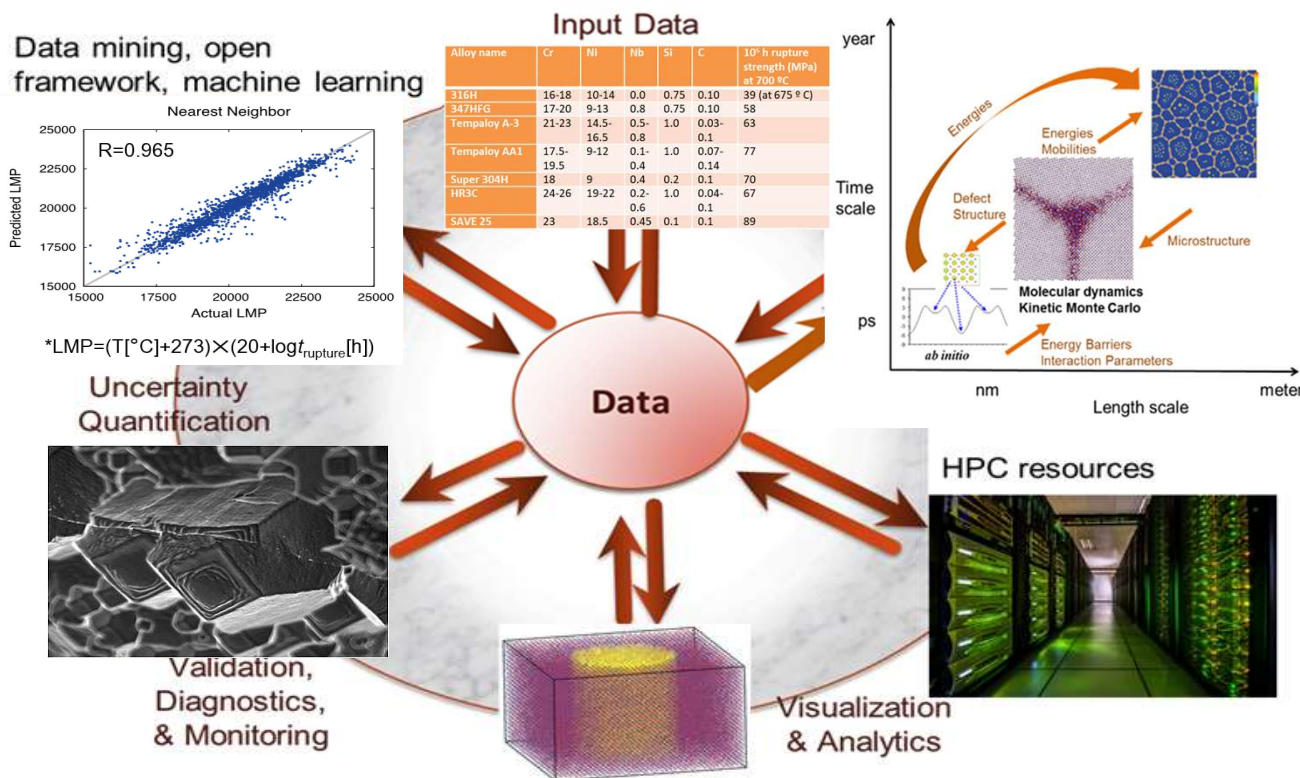
eXtremeMAT

accelerating the development of extreme environment materials



U.S. DEPARTMENT OF
ENERGY

Aim: Use data management and analytics to reduce the time and cost for alloy development and predict lifetime reliably.



eXtremeMAT

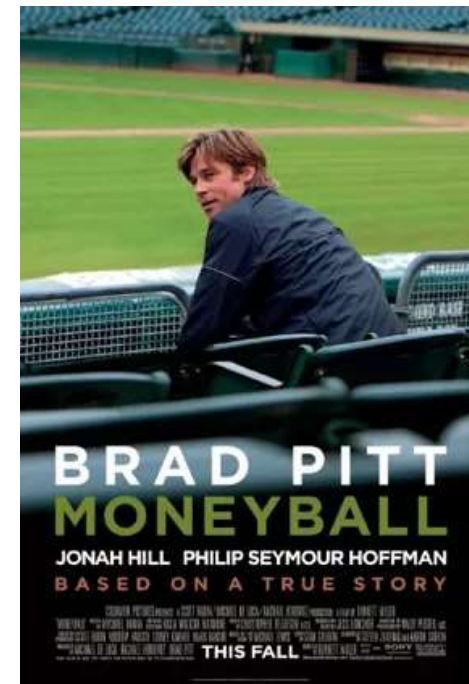
accelerating the development of extreme environment materials



U.S. DEPARTMENT OF
ENERGY

Playing 'Money Ball' with materials is challenging.

Issues	Baseball	Materials Science
Data availability	Reliable and readily available	Scattered, hidden (proprietary), or unpublished.
Standards	Uniform	No data standards
Variability	Unambiguous: Hits, RBI, HR etc.	Uncertainty can be large; variability between groups
Relations	Rules, positions etc. well defined.	Processing-microstructure-property relations nonexistent.
Stability	Rules don't change during the season	Material microstructure changes during processing and service; surface degrades.



Alloy data analysis challenges

- Data is sparse and expensive to obtain.
- Metadata and provenance are missing.
- Existing frameworks focus on 0 K properties.
- Complex descriptors (e.g., microstructure)
- Going beyond confirmation of one's biases
- Interpretability of models is essential.



eXtremeMAT

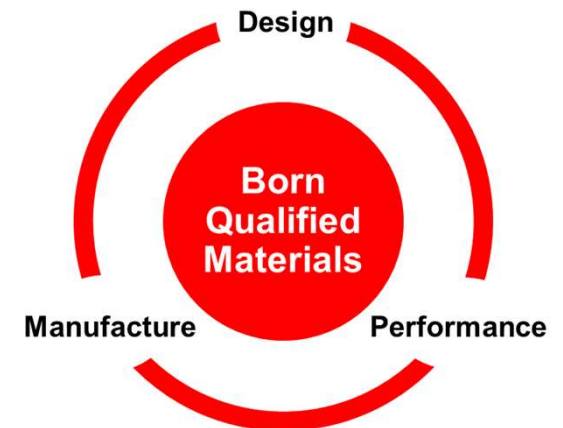
accelerating the development of extreme environment materials



U.S. DEPARTMENT OF
ENERGY

Creep-resistant alloy development issues

- Lack of composition-processing-structure-property relations
- Lack of validated physics-based models (especially processing)
- Gaps in microstructural and thermodynamic data
- Alloy development knowledge and data are not widely shared
- Parameter space is large (minor alloying element optimization)
- Extrapolation from short-term tests to long term life estimation
- Doing this by experiment alone would be very long and expensive



eXtremeMAT

accelerating the development of extreme environment materials



Scope of Task 3: Data Management and Analytics

- Subtask 3.1. Data assessment
 - Collect existing data on austenitic steels and identify gaps for extreme environments
 - Composition, microstructure, mechanical properties, thermodynamics and kinetics
- Subtask 3.2. Data management
 - Develop open standards and formats, curate data, and include metadata
 - Develop robust and flexible user interface to integrate with simulation and analysis tools
 - Integrate and enhance existing frameworks to support collaboration; Sustain data long term
- Subtask 3.3: Data analytics, machine learning, and rapid simulation tools
 - Identify and evaluate available data analytics tools and gaps
 - Couple machine learning with rapid simulation tools (Thermocalc, DICTRA, etc.)
 - Develop data-driven models, integrate with physics-based models, and visualize data
 - Predict lifetime in extreme environments (chromia-forming), validate results, and refine models
 - **Design a new alumina-forming alloy for extreme environments**



eXtremeMAT

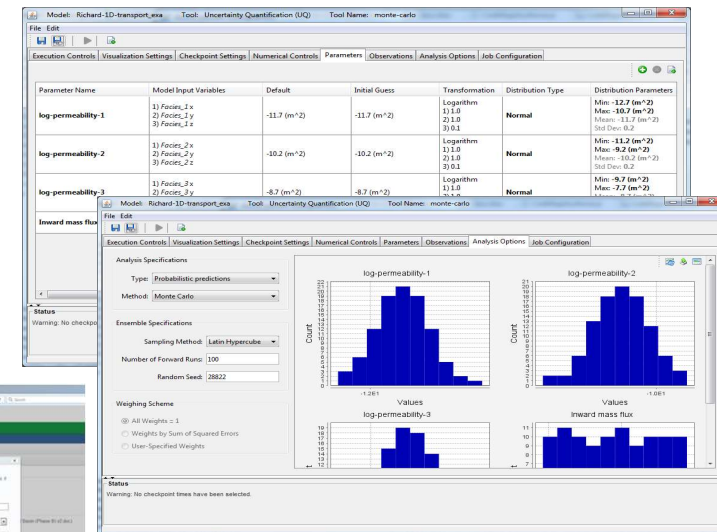
accelerating the development of extreme environment materials



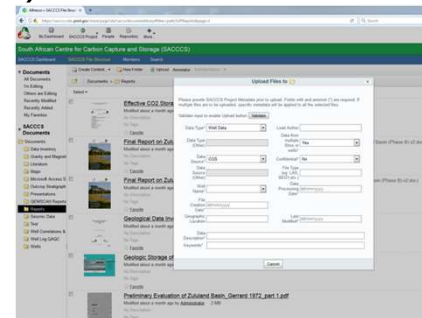
U.S. DEPARTMENT OF
ENERGY

3.2 Data Infrastructure Attributes

- Customizable and powerful web portal-based user interface (UI)
- Tool integration including client-based UIs via web browser
- Simulation management (HPC job launching/monitoring)
- Hierarchical data organization with role-based access control
- Standards-based security and authentication
- Support for structured data
- Local and cloud-based data storage with version control
- Bulk and large file uploads (e.g., microstructure data)
- Open source or no license fees for software/data
- Local deployment behind company firewalls
- Metadata and provenance
- Visualization
- Publishing



Ensemble Job Launching



Metadata capture, display and search



eXtremeMAT

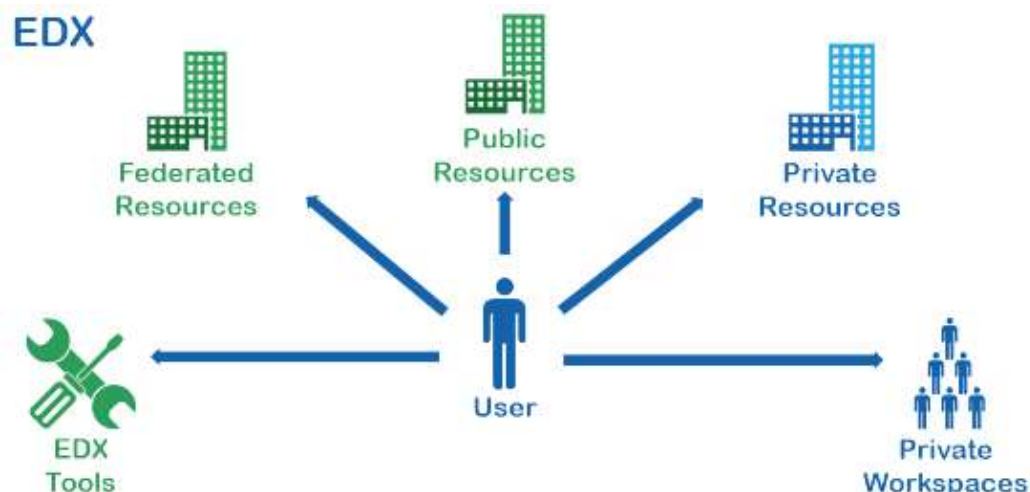
accelerating the development of extreme environment materials



U.S. DEPARTMENT OF
ENERGY

3.2 Energy Data eXchange (EDX)

- EDX is a data management and R&D platform for DOE FE.
- It is a secure, online coordination and collaboration platform that supports energy research, knowledge transfer and data needs.
- Reliable access to historic and current R&D data, data driven products, and tools
- Both public and secure, private functionalities
- Enables knowledge transfer, data preservation, reuse and discovery.



Built by researchers for research



eXtremeMAT

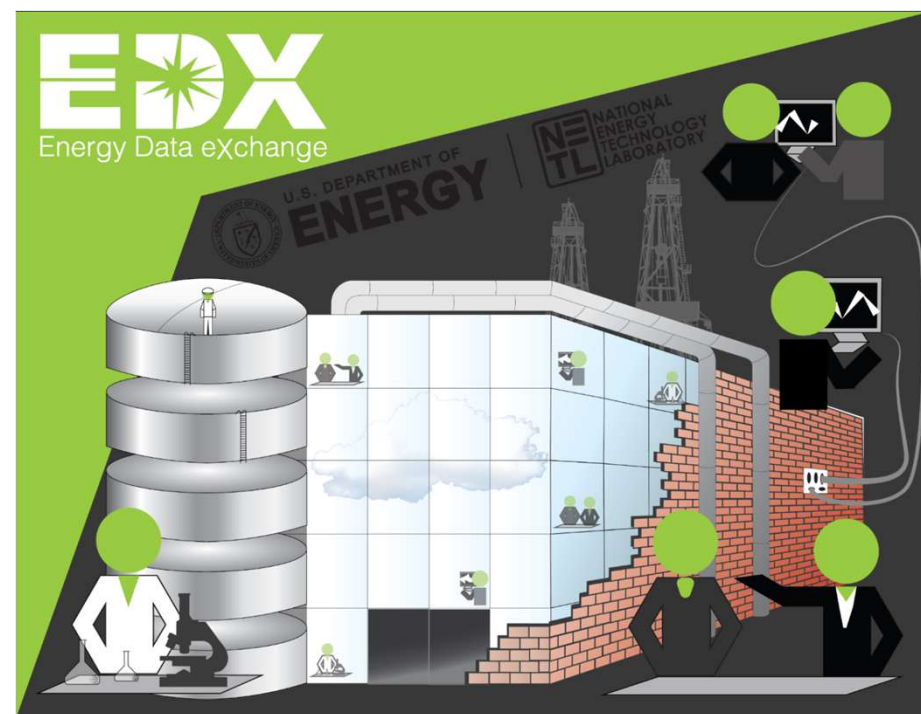
accelerating the development of extreme environment materials



U.S. DEPARTMENT OF
ENERGY

3.2 Energy Data eXchange (EDX)

- All data is contributed by registered EDX users
- Data can be contributed as public resources or private resources shared with a user defined research team (Collaborative Workspaces)
- Over 50,000 EDX Resources (16,299 public, 35,282 private)
- Advanced data discovery and search functionality
- Custom Smart Search Tool and Related Resources functionality
- Enhanced Search within data itself (v3.0)



eXtremeMAT

accelerating the development of extreme environment materials



U.S. DEPARTMENT OF
ENERGY

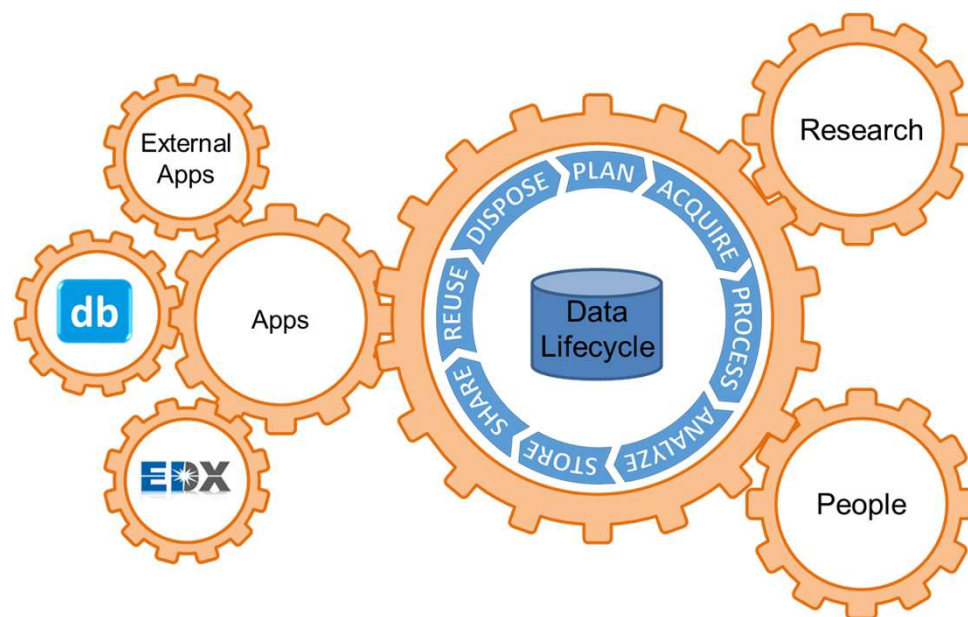
Integrating EDX with Collaborative Frameworks and Repositories

EDX

- Access/search outside data sources
- Data mining
- Collaboration beyond data (forums, calendar)
- DataBook (smart lab notebook)
- Metadata capture/search
- Smart/living database (structured data support)

Open collaborative framework

- Tool integration including client-based
- Simulation management
- Provenance support (tying tools to data)
- Local data storage and deployment



eXtremeMAT

accelerating the development of extreme environment materials



U.S. DEPARTMENT OF
ENERGY

Accommodating Different Data Types with Materials Handbook

Test Type	Creep or Stress-Rupture (CSR)	
Testing Organization	Idaho National Laboratory	
Test Standards	ASTM E139-96	
General Information Notes		
Creep rupture test was performed on specimen damaged by 180 prescribed fatigue-stress relaxation cycle strain range of 1% and an R-ratio of -1, holding at tensile and compressive peaks for 0 minute).		
Specimen Information		
Specimen ID / Number	A-1	
Material Trade Name	A508	
Batch / Heat Number	16567	
Product Form	Plate	
Testing Conditions		
Test Load (Constant Mode)	552.0206	MPa
Test Temperature	350	°C
Raw Data		
<p>Creep Strain vs Time</p>		

Component ID	Component Name	Component Assembly Location	Component Function	Linking value (Component ID)	Linked record
B1	Graphite Brick B1	X=0, Y=Top of B2		B1	Comp
B2	Graphite Brick B2	X=0, Y=Top of B3		B2	Comp
A-G01-X1	Acceleration Sense X1	B1	Measuring acceleration signal in X direction	A-G01-X1	Comp
A-G01-Y1	Acceleration Sense Y1	B1	Measuring acceleration signal in Y direction	A-G01-Y1	Comp
HR	Strain Gauge HR	B5 Dowel 1, 0° direction	Measuring graphite stress in radial and tangential directions	HR	Comp
HT	Strain Gauge HT	B18 Dowel 2, 270° direction	Measuring graphite stress in radial and tangential directions	HT	Comp
V-G01-T-01	Camera 01		Measuring the displacement of brick B1 and base plate.	V-G01-T-01	Comp

Procedure ID	Load Case (Boundary Condition)	Direction	Excitation Type	Excitation Amplitude (g)	Linking value (Procedure ID)
Test Procedure LCC1 P1	LCC1 - Top and Side Springs	X	White Noise	0.05	1.1
Test Procedure LCC1 P2	LCC1 - Top and Side Springs	Y	White Noise	0.05	1.2
Test Procedure LCC1 P3	LCC2 - Side Springs	X	White Noise	0.05	2.1

Source 1 Data Source Editor Name	R. W. SWINDEMAN
Source 1 Data Package File 1	8092297 ann 16mm plate FILE.xls
Source 1 Data Package File 2	8092297 mill ann 16mm plate FILE.xls
Source 1 Data Package File 3	316 test summary.xlsx
Record Management Information	
Handbook Record ID	316SS 8092297 Ann Plate Tensile Cree
Record Contributing Signatory	DOE United States
Information Category	Background Public Information
Record Distribution Scope	Unlimited
Record Edited by	Lianshan LIN and Weiju REN
Digitization Status	In Progress
Further Information	
Generic Alloys for this Data Package	316H SS
Alloy Pedigrees for this Data Package	316H SS_8092297_Plate_Annealed_Republic_Steel_1974
Creep Test Data for this Data Package	41 Links Show All

Microanalysis Organization	Oak Ridge National Laborat
Specimen Information	
Material Trade Name	Inconel 617
Batch / Heat Number	XX01A3US
Product Form	Plate
Product Dimensions	13 mm (1/2")
Heat Treatment History	Solution annealed by vendor
Aging History	593 C x 10,000 hours aging
Testing History	Creep testing at 593 C and
Microstructural Characteristics	
Characterization Image 1	
Image 1 ID	M-10818
Comments on Image 1	
Creep fracture surface of specimen aged and creep tested in helium. The particles or	

Materials data along with pedigree, chemical and physical properties, specimen information, technical specification etc.



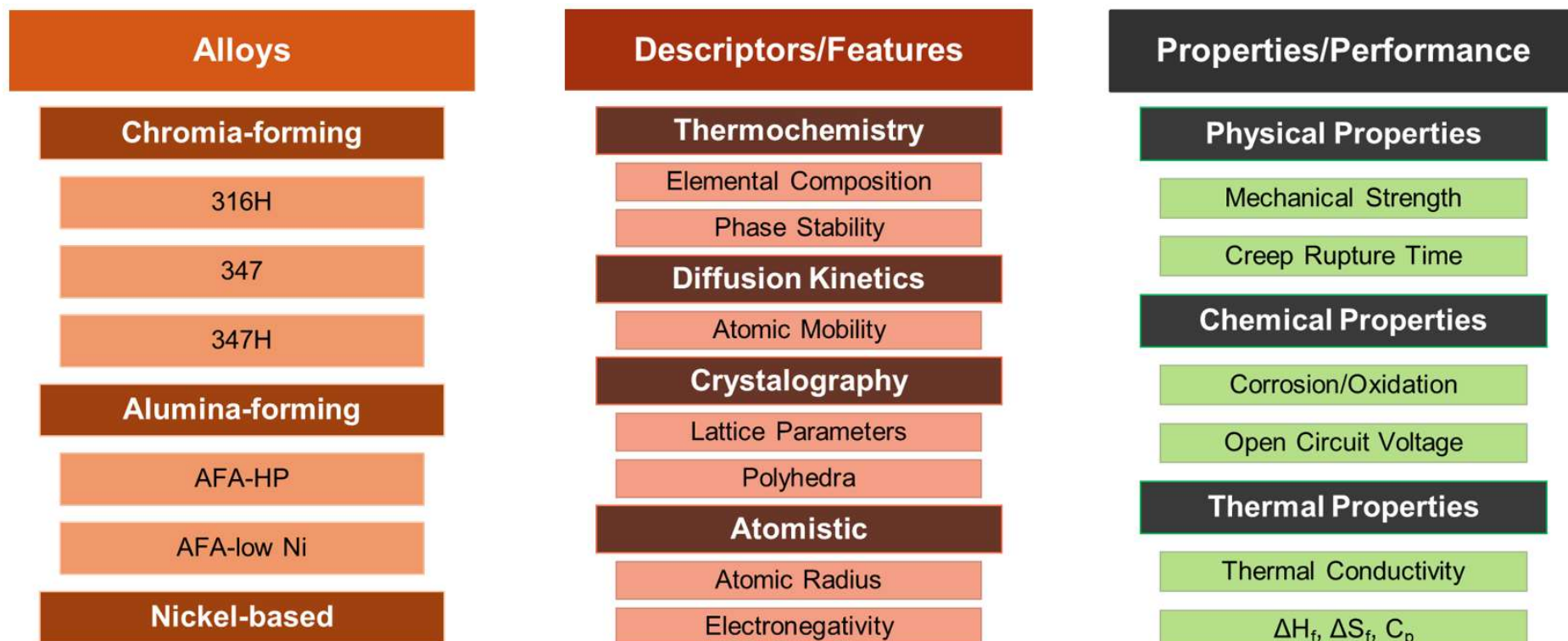
eXtremeMAT

accelerating the development of extreme environment materials



U.S. DEPARTMENT OF
ENERGY

3.3 Inputs and Outputs for Data Analytics



eXtremeMAT

accelerating the development of extreme environment materials



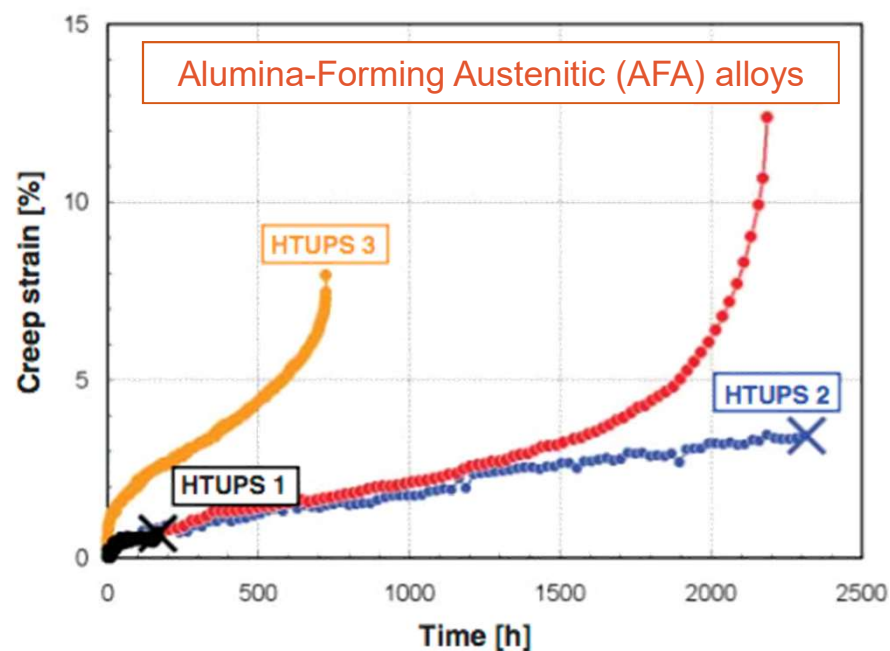
U.S. DEPARTMENT OF
ENERGY

High-Dimensional Optimization

Compositions (wt %)

Elements	HTUPS	HTUPS 1	HTUPS 2	HTUPS 3	ORNL AFA
Fe	64.27	60.25	57.73	56.58	57.78
Ni	16	19.97	20	19.98	19.95
Cr	14	14.15	14.2	14.21	14.19
Al	-	-	2.4	3.67	2.48
Si	0.15	0.15	0.15	0.1	0.15
Mn	2	1.95	1.95	1.92	1.95
Mo	2.5	2.47	2.46	2.46	2.46
Nb	0.15	0.14	0.14	0.14	0.86
Ti	0.3	0.28	0.31	0.31	-
V	0.5	0.49	0.5	0.49	-
C	0.08	0.068	0.076	0.079	0.075
B	0.01	0.007	0.011	0.011	0.01
P	0.04	0.042	0.044	0.04	0.043

Yamamoto et al., *Science* 316:322, 2007.



Complex high-dimensional statistics problem (>10 elements) to optimize/maximize target properties



eXtremeMAT

accelerating the development of extreme environment materials

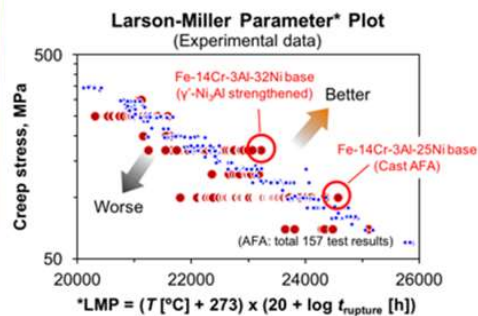


U.S. DEPARTMENT OF
ENERGY

Fitting High-Dimensional Datasets

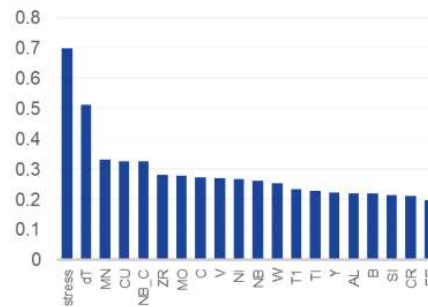
Data Collection/Population

- Need consistent experimental data
- What do we want to predict?
 - Strength? Corrosion?
- Material descriptors
 - Composition
 - Thermodynamics and kinetics
 - Microstructure etc.



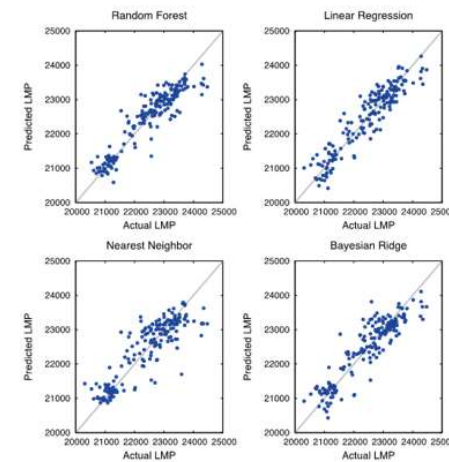
Correlation Analysis

- Feature selection
- High ranking descriptors
- Superficial vs scientific descriptors
- **Participation of domain experts**
- Iterative process



Machine Learning

- 'Just fit' the curve...
- Different ML models



Idaho National Laboratory

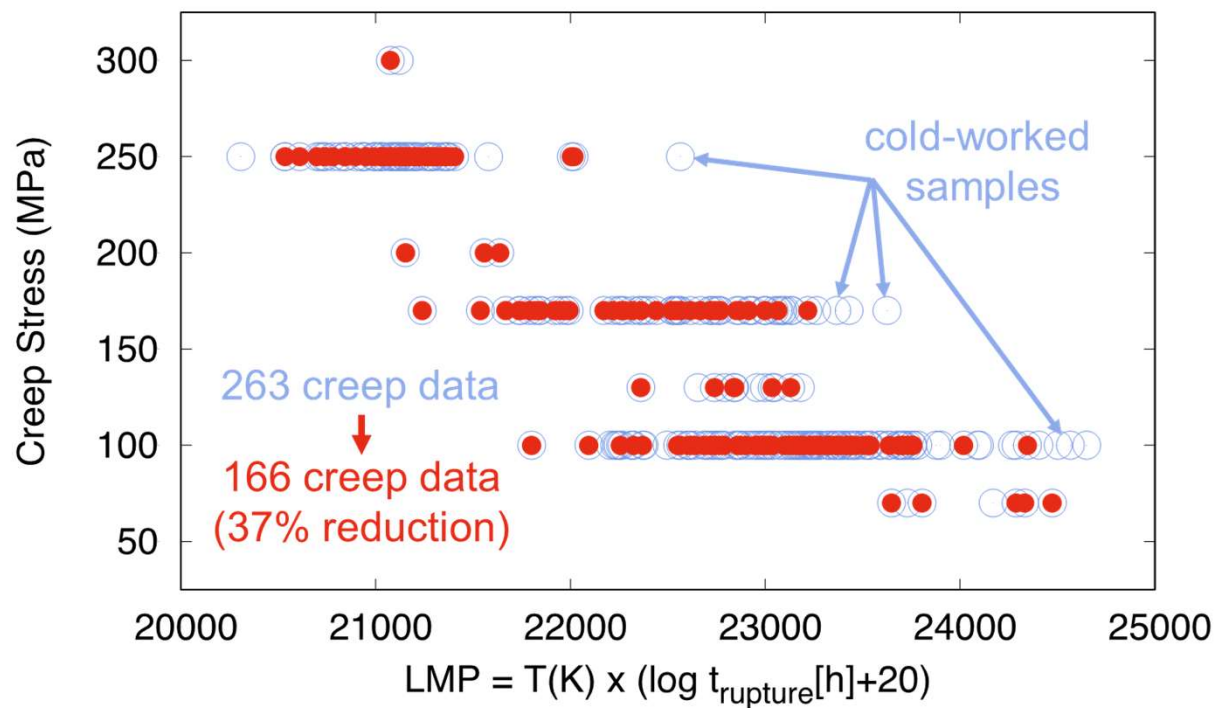


eXtremeMAT

accelerating the development of extreme environment materials



Knowing Pedigree of Data is Crucial to Develop Reliable Models



eXtremeMAT

accelerating the development of extreme environment materials

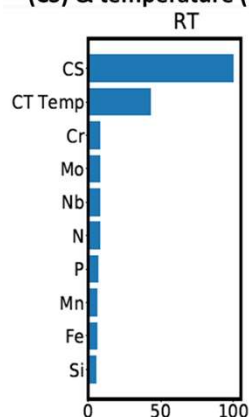


U.S. DEPARTMENT OF
ENERGY

Analysis of Rupture Time for NETL 9Cr Data Set

We have constructed good ML models that could be used to predict chemistry with targeted rupture time.

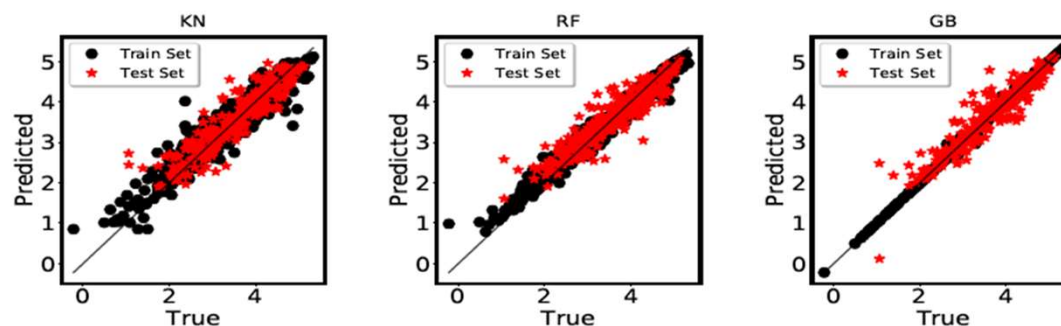
Descriptor ranking: Creep test stress (CS) & temperature (CT Temp) crucial



1248 data size, 26 descriptors

- Alloys compositions (21): Fe, C, Cr, Mn, Si, Ni, Co, Mo, W, Nb, Al, P, Cu, Ta, Hf, Re, V, B, N, O, S
- Temperatures (4): Homogenization heat treatment temp (**Homo**), Normalization temp (**Normal**), Heat treatment temp1 (**Temper1**), Creep Test Temp (**CT Temp**)
- Stress (1): Creep test stress (**CS**)

Comparison of Machine Learned Models



eXtremeMAT

accelerating the development of extreme environment materials



Ultimate Goal of the Data Management and Analytics Task

Accessible Databases

- Open standard, sharable database formats with curated eXtremeMAT data
- Flexible interface for access control and simulation/analytics job launching
- Connectivity to existing open databases (energy materials network standards)
- Use data to develop reduced order models and link modeling scales

Materials properties data for extreme environments

- Structure, volume fractions of phases, and phase diagrams
- New range of compositions beyond current standards
- Thermomechanical processing history
- Performance data : corrosion, fatigue, creep, UTS, LCF
- Metastable data: time and temperature, aging
- Uncertainty quantification

Analytical Tools

- Novel tools to predict time to failure (incipient or not) in a chromia-forming alloy
- Design of a new alumina-forming alloy using the data and tools of eXtremeMAT



eXtremeMAT

accelerating the development of extreme environment materials



U.S. DEPARTMENT OF
ENERGY

Ultimate Goal of the Data Management and Analytics Task

Accessible Databases

- Open standard, sharable database formats with curated eXtremeMAT data
- Flexible interface for access control and simulation/analytics job launching
- Connectivity to existing open databases (energy materials network standards)
- Use data to develop reduced order models and link modeling scales

Materials properties data for extreme environments

- Structure, volume fractions of phases, and phase diagrams
- New range of compositions beyond current standards
- Thermomechanical processing history
- Performance data : corrosion, fatigue, creep, UTS, LCF
- Metastable data: time and temperature, aging
- Uncertainty quantification

Analytical Tools

- Novel tools to predict time to failure (incipient or not) in a chromia-forming alloy**
- Design of a new alumina-forming alloy using the data and tools of eXtremeMAT**



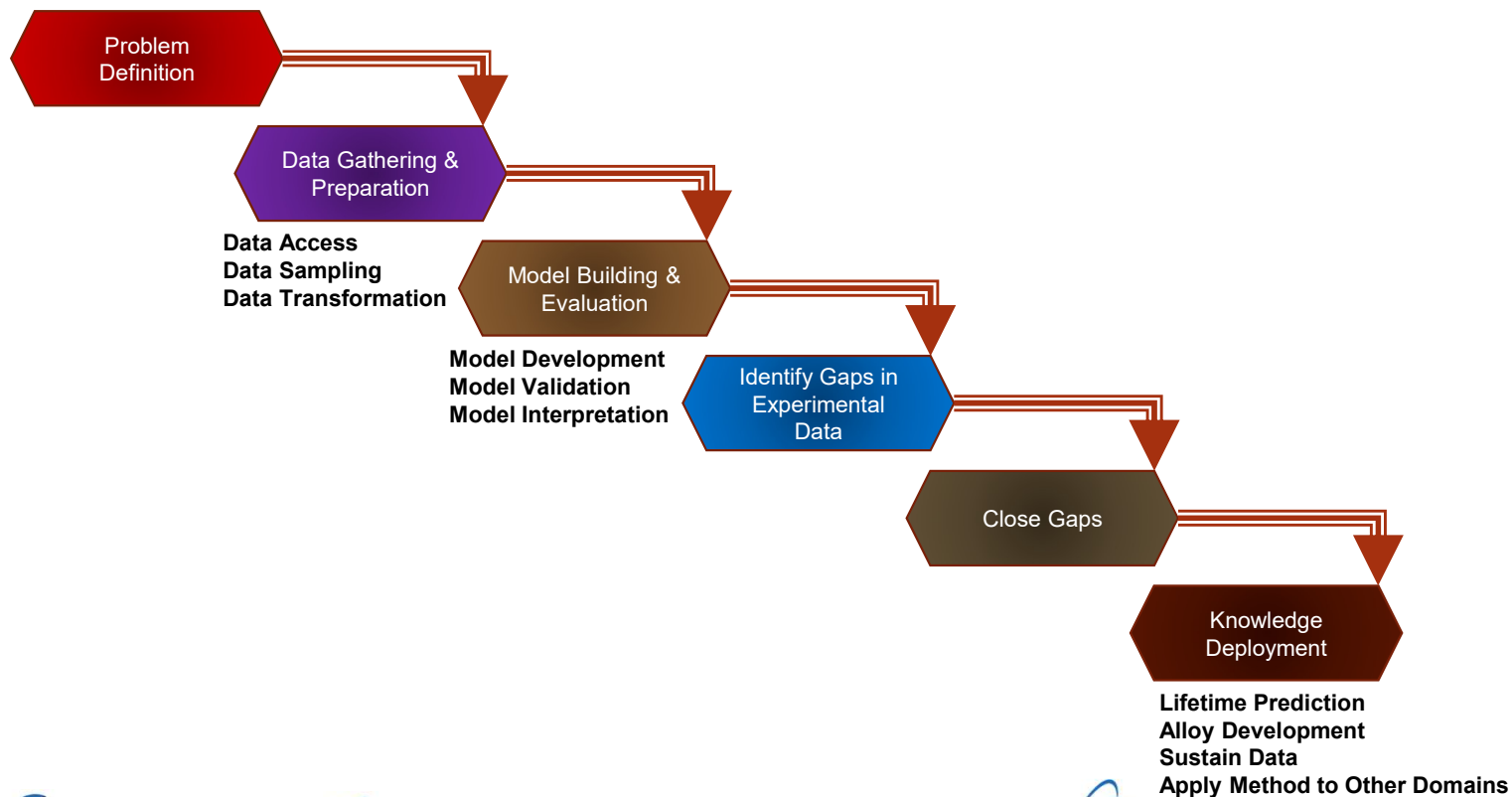
eXtremeMAT

accelerating the development of extreme environment materials



U.S. DEPARTMENT OF
ENERGY

Effective Integration Across Labs and Tasks



eXtremeMAT

accelerating the development of extreme environment materials



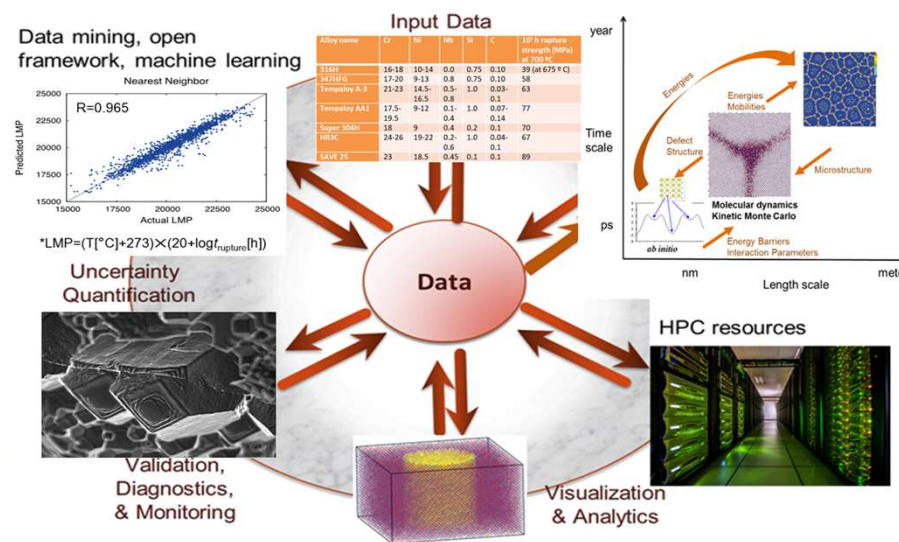
U.S. DEPARTMENT OF ENERGY

Questions?

We will use machine learning to develop data-driven models, develop linkages between physical models (Task 2) and connect them to experimental validation (Task 4).

These tools will reduce the time and cost for alloy development and lifetime prediction.

Task 4



Task 2



eXtremeMAT

accelerating the development of extreme environment materials



U.S. DEPARTMENT OF
ENERGY

Extra slides

- **Variability is a huge issue**
 - Composition and performance of the same alloy made by different teams could vary.
- **Brute force simulations may not get us to our goal**
- **Embed data analytics and VV & UQ at every stage**
- **Need to decide on data and tools platform**
 - Framework should integrate metadata recording, job submission, publication, visualization, access control etc.
 - Get past IP concerns and open source tools
 - Access control and security; File versioning; Metadata and Provenance
 - Bulk uploads, large files (microstructure; 10s of GB)
 - Job launching (simulation management)
 - Support for other repositories
 - User interface must be intuitive for power and infrequent users



eXtremeMAT

accelerating the development of extreme environment materials



- **Need to put some thought into machine learning**
 - ML often confirms your biases
 - Need ML that develops new and unanticipated insights
 - Genetic programs to develop constitutive laws on the fly
 - New analytic forms without prior assumptions
- **Modeling should focus on relevant conditions**
 - Previous alloy development was for room temp (no microstructural evolution)
 - Bond region in heat exchanger (don't model bulk for this)
- **Multi scale models: Coarsening may throw out key data or features.**
 - Don't throw the baby out with the bath water. Use human expertise with ML.
- **CALPHAD is an inadequate database**
- **Classical thermodynamic data does not have a length scale (can't predict microstructure evolution)**
- **Identify relevant descriptors**
 - Multiple descriptors; go beyond simple descriptors



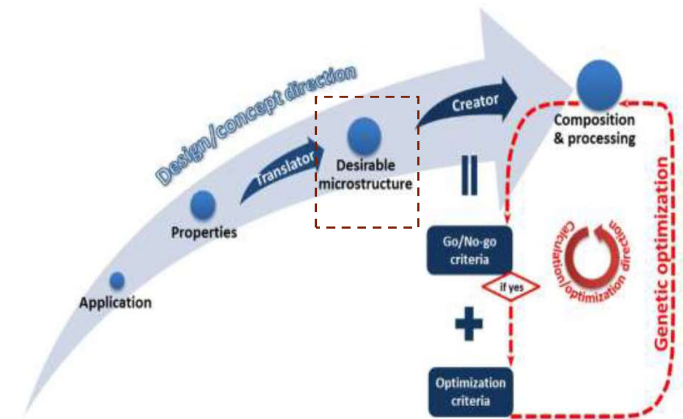
eXtremeMAT

accelerating the development of extreme environment materials



Microstructural data and thermodynamic data are missing

- Data is available for composition and mechanical properties, but not for microstructure.
- Microstructure is treated as a black box between chemistry and properties.
- Microstructure changes during service.
- Need to develop key descriptors of microstructure for quantification.



Q. Lu et al, J. Mat. Sci. Tech, 2017



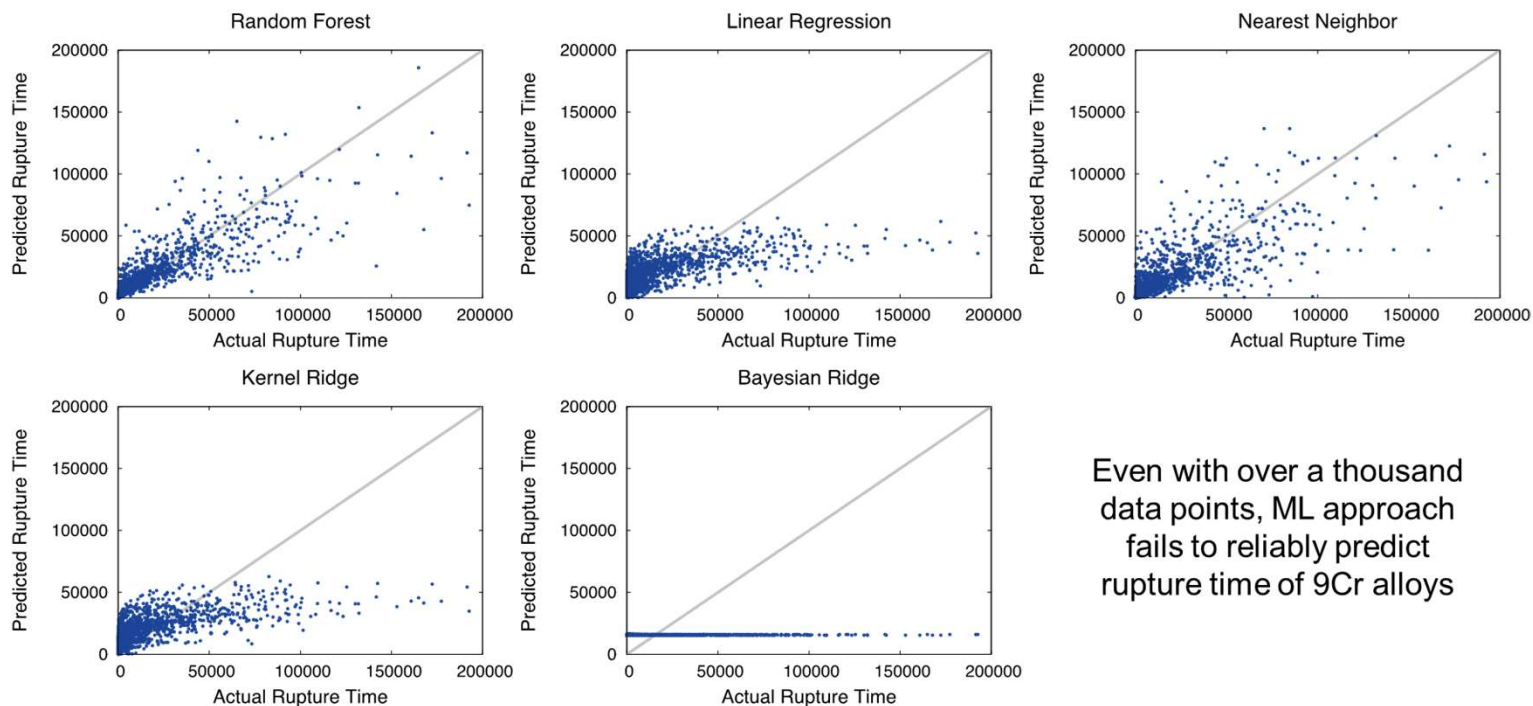
eXtremeMAT

accelerating the development of extreme environment materials



U.S. DEPARTMENT OF
ENERGY

Preliminary analysis of NETL 9Cr alloy data



Even with over a thousand data points, ML approach fails to reliably predict rupture time of 9Cr alloys



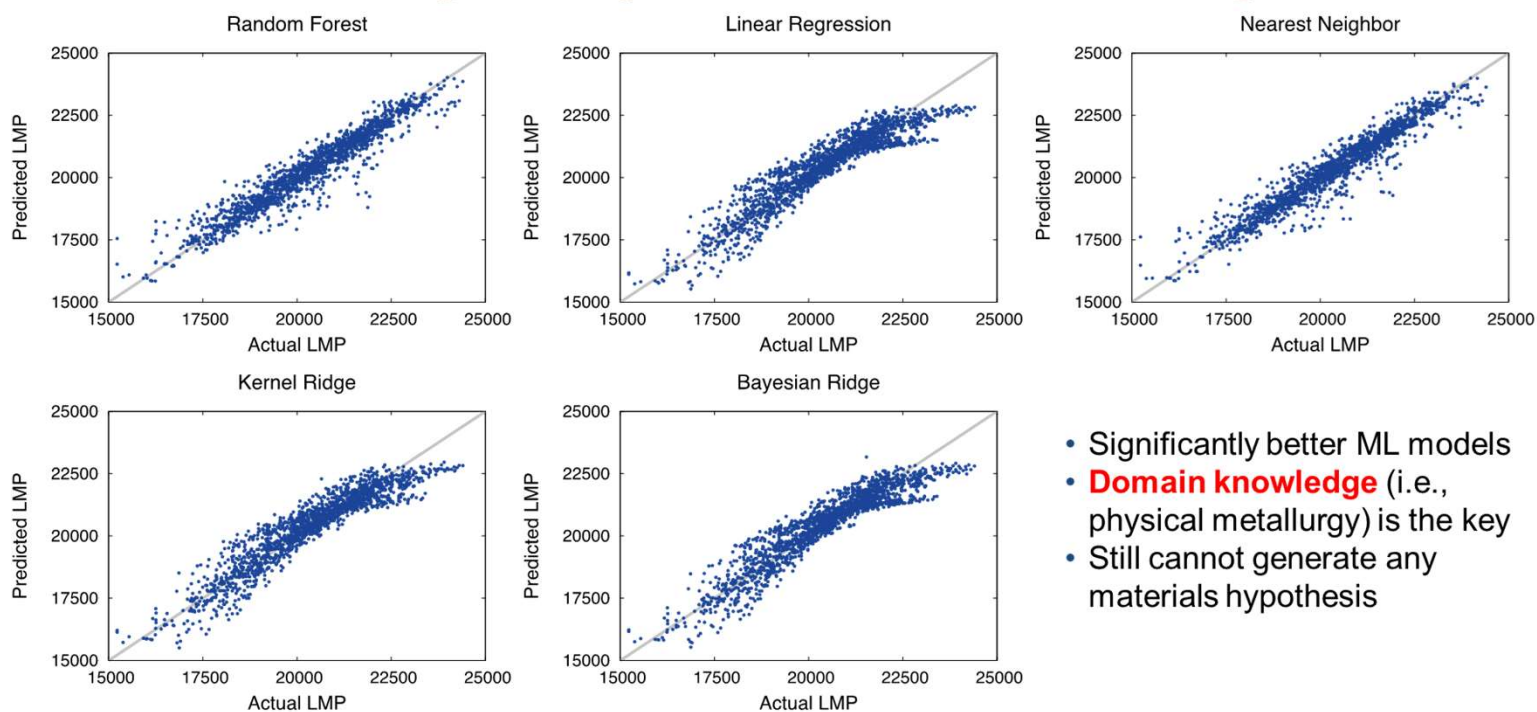
eXtremeMAT

accelerating the development of extreme environment materials



U.S. DEPARTMENT OF
ENERGY

Preliminary analysis of NETL 9Cr alloy data



- Significantly better ML models
- **Domain knowledge** (i.e., physical metallurgy) is the key
- Still cannot generate any materials hypothesis

