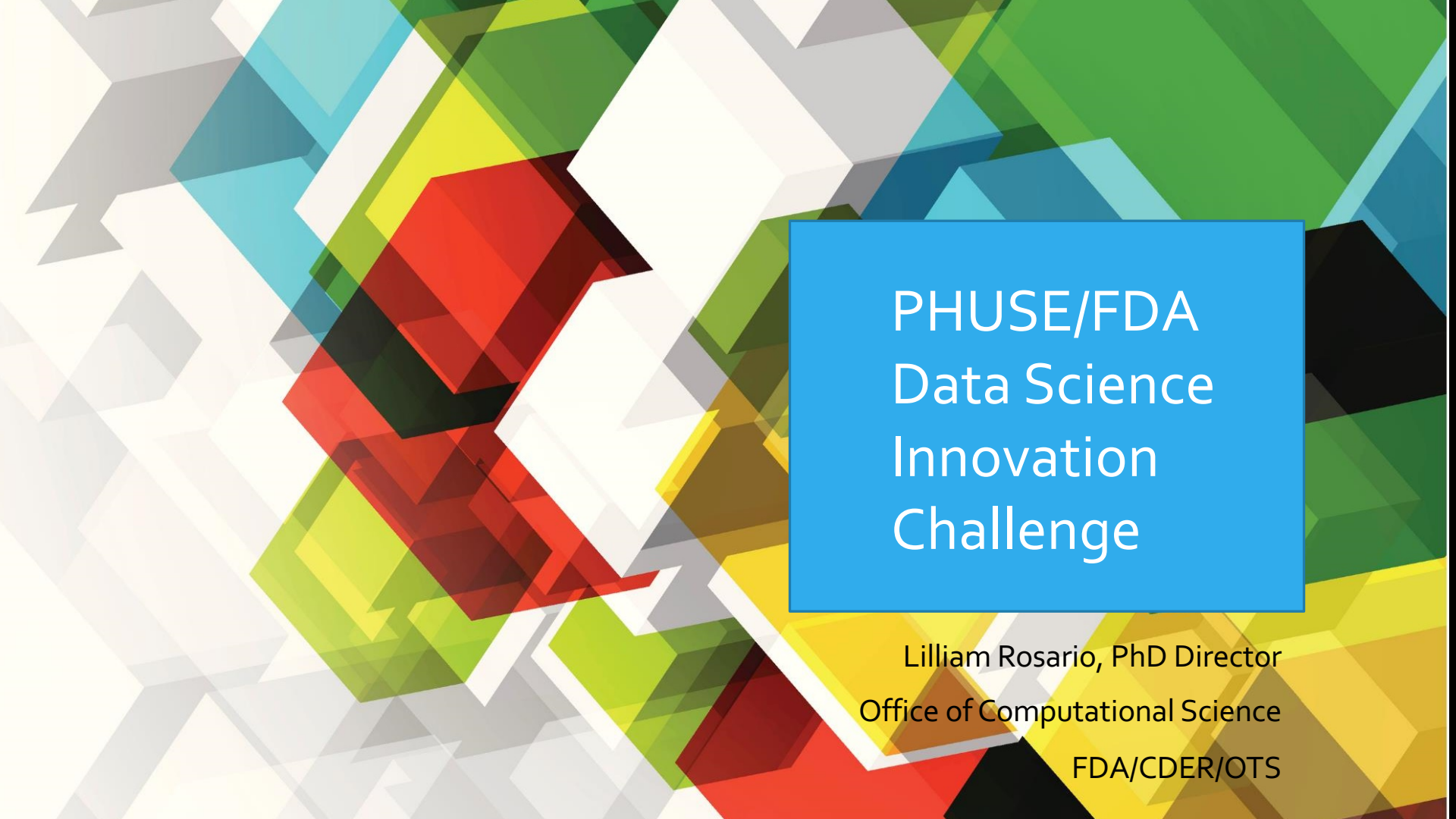


The Computational Science Symposium

September 21st–23rd

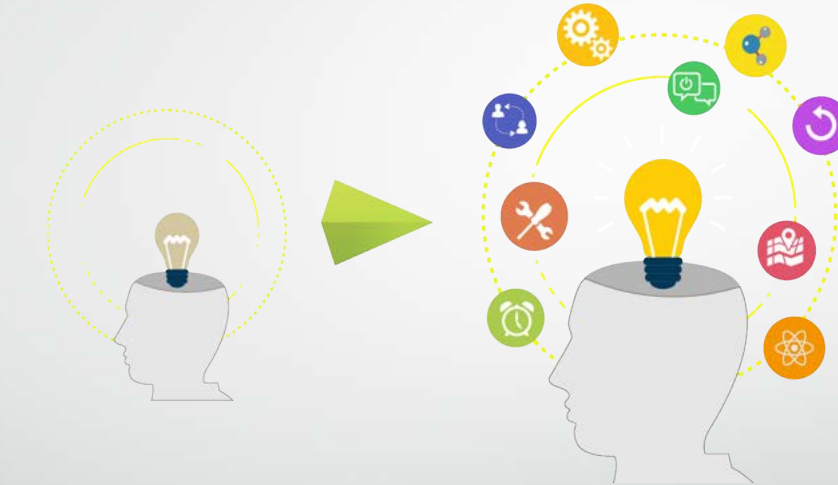
Panel Discussion: Outcomes of the First
PHUSE/FDA Data Science Innovation
Challenge: Collaborate and Innovate to
Shape the Future of Healthcare



PHUSE/FDA Data Science Innovation Challenge

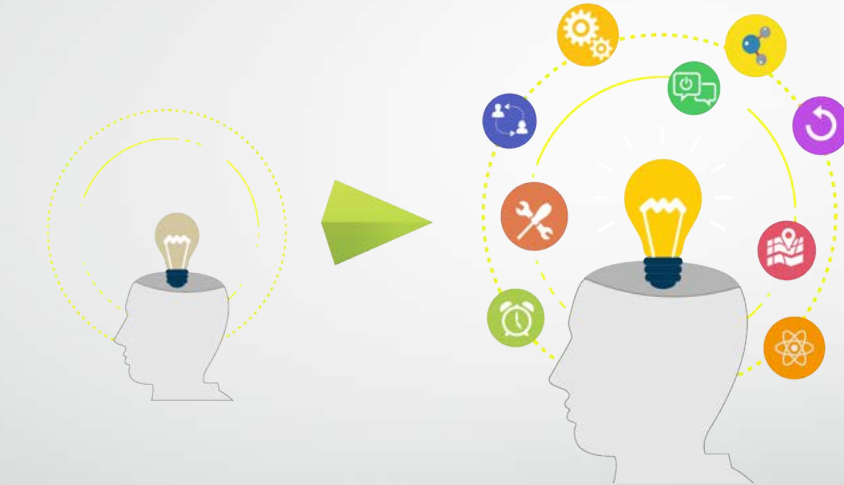
Lilliam Rosario, PhD Director
Office of Computational Science
FDA/CDER/OTS

What was asked of the Challengers?



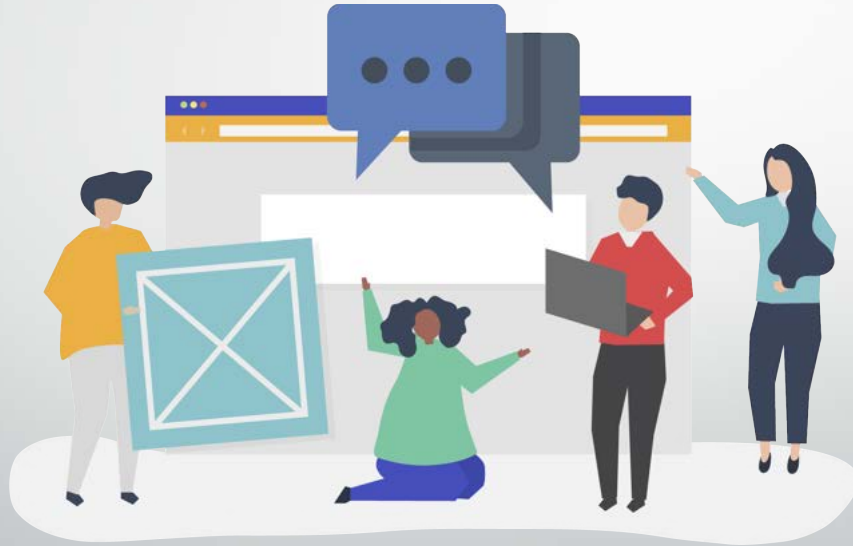
We invited people to join this challenge, to work on prototypes, and to present their proof-of-concepts or solutions at the PHUSE Annual US Connect Event, March 2020 in Orlando.

What was ACTUALLY asked of the Challengers?



We turned the Challenge into an Opportunity and invited participants to present during INNOVATION FRIDAYS!

Welcome to Innovation Fridays



Who responded to the Challenge?



Challenges



OPIOID/SUBSTANCE USE
DISORDER



DRUG SAFETY
SURVEILLANCE USING
SOCIAL MEDIA



APPROACH FOR
PREDICTING DRUG
INTERACTIONS.



OPIOID/SUBSTANCE USE DISORDER

- Co-chairs- James R. Johnson, Mitra Ahadpour and Catherine Li
- Target- Identify patterns in initiation of drug(s), transition to substance use disorder, and reasons for success and failure of treatment and triggers for relapse. Leverage existing algorithms and build on them to develop predictive algorithms for prediction, intervention and treatment.



APPROACH FOR PREDICTING DRUG INTERACTIONS

- Co-chairs- Beverly Hayes and Sanjay Sahoo
- Target- Integration of multiple sources of adverse event data to visualize and analyze the prevalence of drug interactions in the specified populations. Predictive model for identifying potential drug interactions across different patient populations at risk.



POSTMARKET DRUG AND BIOLOGIC PRODUCT SAFETY SURVEILLANCE

- Co-chairs- Patricia Hegarty and Suranjan De
- Target- Adverse event analysis dashboard that provides the ability to view and analyze adverse events across and within specific patient populations over time.

<https://www.phuse.eu/phuse-fda-data-science-innovation-challenge>

OPIOID

- IC03: Opioid/Substance Use Disorder – Transfer Learning for Opioid Usage Monitoring
- IC29: Opioid/Substance Use Disorder – Deep Social Media Mining Tool to Monitor Opioid Use of Female Pregnant Patients
- IC30: Opioid/Substance Use Disorder – Mosaic Portrait of a Persona – Collecting Topic-specific Social Media Data of the Target Population to Piece Together a Persona of Pregnant Women and the Link to Opioid Usage

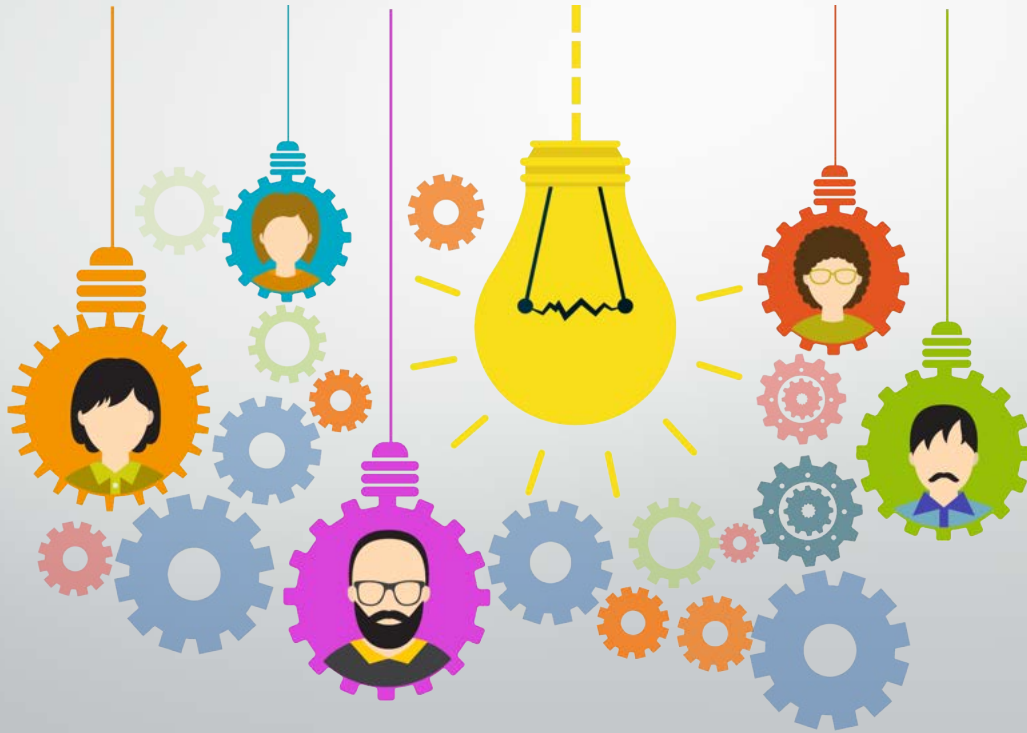
Predicting Drug Interactions

- IC07: Predicting Drug Interactions – Predicting Drug Interactions Using Machine Learning and Data Visualization
- IC12: Predicting Drug Interactions – Prediction of Drug Interactions with Real-world Data
- IC14: Predicting Drug Interactions – Real-world Evidence Drug-Drug Interaction (REDDI) Match
- IC18: Predicting Drug Interactions – Artificial Intelligence for Early Detection of Drug Interactions
- IC26: Predicting Drug Interactions – Approaches in Categorizing and Predicting Drug-Drug Interactions Using Artificial Intelligence and SAS Visual Text Analytics

Drug Safety Surveillance

- IC08: Drug Safety Surveillance – Datacise™ Integrated Safety Explorer
- IC09: Drug Safety Surveillance – PRAView
- IC15: Drug Safety Surveillance – Dash AEs
- IC19: DeepFAERS: AI for Pharmacovigilance by Diving Deep into FAERS
- IC20: Drug Safety Surveillance - Using Media Data and Public Regulatory Database
- IC21: Drug Safety Surveillance – Deep Dissection of Post-market Surveillance Signal
- IC23: Drug Safety Surveillance – A Deep Learning Phylogenomic Approach to Advance Post Market Safety Surveillance of Drugs and Biologic Products
- IC25: Drug Safety Surveillance – Counterfeit Drug Detection
- IC31: Drug Safety Surveillance – Early Safety Signal Detection from Public Internet Sources

THANK YOU and...



OPIOID/SUBSTANCE USE DISORDER

Facilitate effective intervention
for prevention and treatment of opioid/substance use disorder
by harnessing social media data

Catherine Li

The Epidemic

THE OPIOID EPIDEMIC BY THE NUMBERS



130+

People died every day from opioid-related drug overdoses¹ (estimated)



10.3 m

People misused prescription opioids in 2018¹



47,600

People died from overdosing on opioids²



2.0 million

People had an opioid use disorder in 2018¹



81,000

People used heroin for the first time²



808,000

People used heroin in 2018¹



2 million

People misused prescription opioids for the first time²



15,349

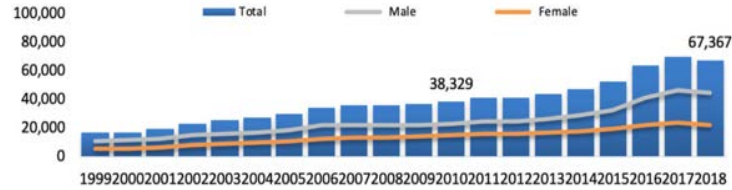
Deaths attributed to overdosing on heroin (in 12-month period ending February 2019)²



32,656

Deaths attributed to overdosing on synthetic opioids other than methadone (in 12-month period ending February 2019)²

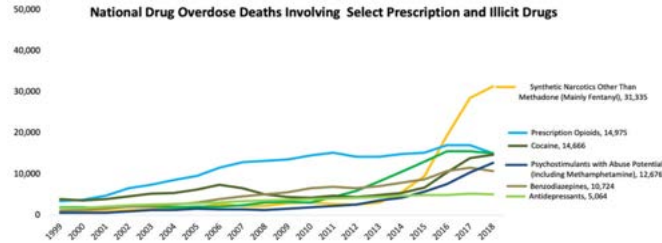
National Drug Overdose Deaths
Number Among All Ages, by Gender, 1999-2018



3 Waves of the Rise in Opioid Overdose Deaths



National Drug Overdose Deaths Involving Select Prescription and Illicit Drugs

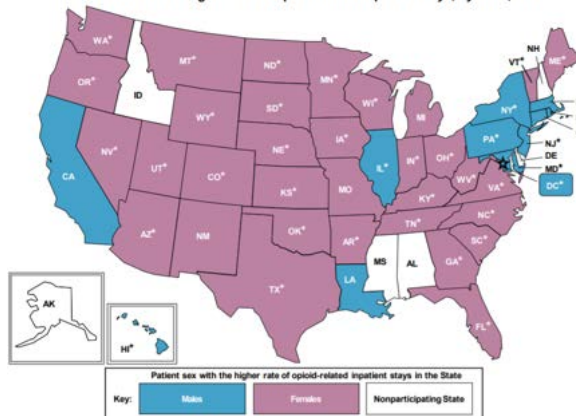


[https://www.drugabuse.gov/drug-topics/trends-statistics/overdose-deaths-rates#:~:text=Drug%20overdose%20deaths%20involving%20prescription,\(Source%3A%20CDC%20WONDER\).](https://www.drugabuse.gov/drug-topics/trends-statistics/overdose-deaths-rates#:~:text=Drug%20overdose%20deaths%20involving%20prescription,(Source%3A%20CDC%20WONDER).)

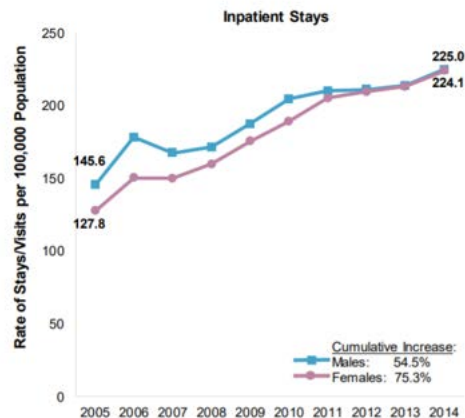
<https://www.cdc.gov/drugoverdose/epidemic/index.html>

Opioid Addition in Women

Patient sex with the higher rate* of opioid-related inpatient stays, by State, 2014



National rate of opioid-related inpatient stays



Reproductive-aged women

- 1/3 filled opioid prescription each year between 2008 and 2012

Every 15 minutes, a baby was born with **NAS**



By the AHRQ:

<https://www.hcup-us.ahrq.gov/reports/statbriefs/sb224-Patient-Characteristics-Opioid-Hospital-Stays-ED-Visits-by-State.pdf>

FDA's Role and Approaches

The Opioid Crisis: FDA's Priorities & Strategies



The Opioid Challenges:

Opioid Abuse and Misuse

Pain in America

Injectable Opioid Shortages

1. Decreasing Exposure & Prevent New Addiction

- Appropriate Dose/Duration Labeling
- Appropriate Packaging, Storage, and Disposal
- Health Care Provider Education

2. Supporting the Treatment of Those With Opioid Use Disorder

- Naloxone
- Medication Assisted Treatment (MAT)

3. Fostering the Development of Novel Pain Treatment Therapies

- Partnerships & Meetings
- Abuse Deterrent Formulations (ADFs)
- Pain Treatment Alternatives

4. Improving Enforcement & Assessing Benefit-Risk

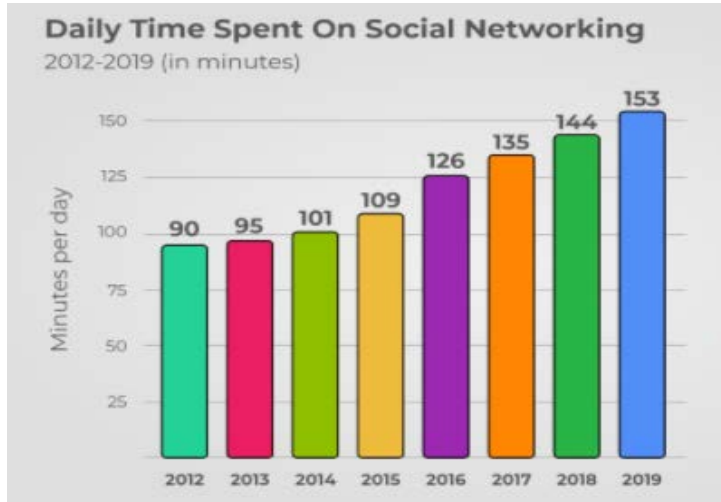
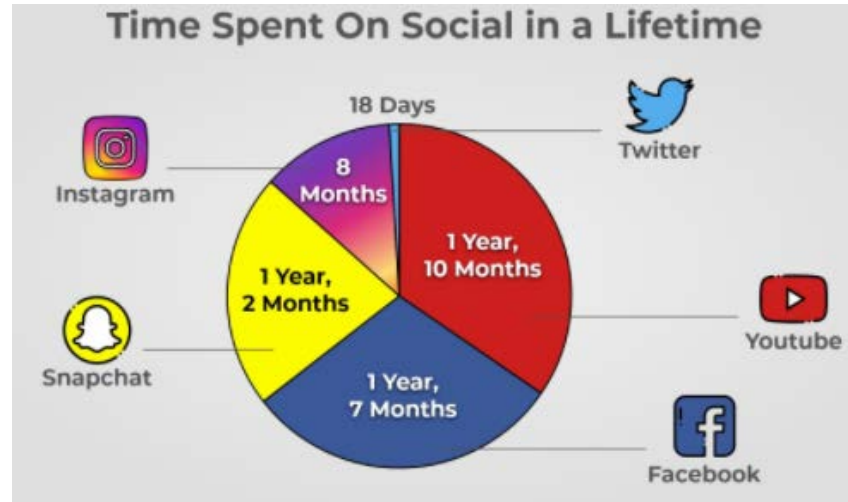
- Improving Enforcement
- Assessing Benefit-Risk



<https://www.fda.gov/media/122423/download>

Why - Social Media and Our Lives

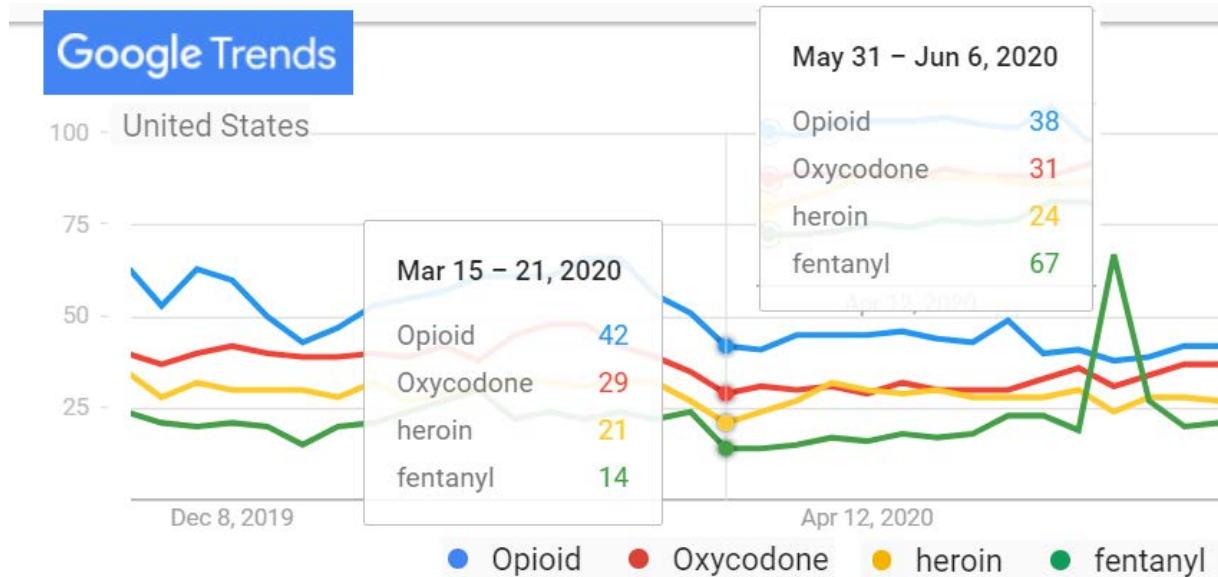
Using social media is an integral part of our daily lives



<https://www.broadbandsearch.net/blog/average-daily-time-on-social-media>

During the COVID-19

Discussions of opioids are increasing during the COVID-19 time



Challenges with Social Media Data

Challenges Presented by Social Media Data Mining

- Huge amount of unstructured data
- Noisy
- Authenticity
- Who is posting?
- Bias
- Reliability
- Verifiability

Opioid/Substance Use Disorder Stream

Applications and Selections

Category	Number
Total Number of Applications	7
Selected	4
Withdrawn	1

Natural Language Analysis of Reddit for Opiates Usage

Mohit Juneja



Phuse | US Connect 2020

Deep social media mining tool to monitor opioid use during Pregnancy with **ThinkTrends AI**



ThinkTrends
make your AI think

ThinkTrends, LLC
171 E Main St, Suite 301
Herndon, VA 20170
Phone: 202-696-9413
www.thinktrends.ai
March 18, 2020
Presented by: Ayelicia Brown, Founder

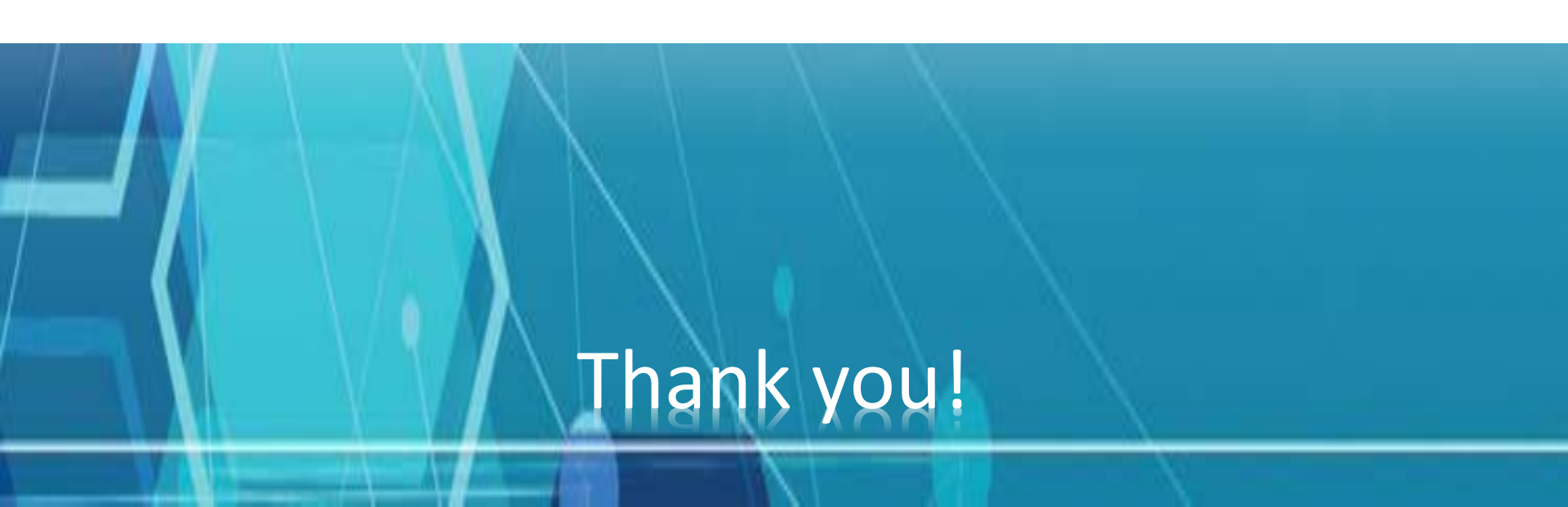
The content of this presentation is proprietary and confidential information of ThinkTrends LLC. It is not intended to be distributed outside the Government without the written consent from ThinkTrends LLC.

Mosaic Portrait of a Persona

Collecting Topic-Specific Social Media Data on Target Population to Piece Together a Persona of Pregnant Women and the Link to Opioid Usage



PhUSE US Connect | Innovation Challenge
Team IBM & FDA



Thank you!

The Computational Science Symposium

September 21st–23rd

Natural Language Analysis of Reddit for Opiates Usage
Mohit Juneja



Analysis Objective and Results

- Develop a solution that provides following capabilities:
 1. Analyze social media posts (Reddit) of Opioid Users to extract insights (demographics, co-occurring opiates, other health conditions/medical drugs linked to opioids usage)
 2. Develop a solution that can scale and support large data volumes

Variable	Class	Precision*	Recall*	F1 score*
Gender (73%)	Male	0.77	0.77	0.77
	Female	0.78	0.78	0.78
Depression	Depression	0.90	0.87	0.88
	Non-Depression	0.89	0.89	0.89
ADHD	ADHD	0.92	0.91	0.91
	Non-ADHD	0.90	0.90	0.90

Variable	Class	Precision*	Recall*	F1 score*
Adverse Events/Anatomy/ Test Treatment	Adverse Event	0.95	0.91	0.93
	Anatomy	0.92	0.97	0.94
	Test Treatment	0.98	0.95	0.96
Pregnancy Status	Pregnant	0.95	0.96	0.95
	Non-Pregnant	0.94	0.92	0.93



Learnings: Modelling Objective

- Criteria for defining appropriate modelling metric:
 1. Objective and measurable based on availability of data
 2. Importance of metric in addressing the underlying objective

Variable	Variable Selection: Reasons	Variable Selection: Challenges/Drawbacks
Depression ADHD	High quality proxy data for classification (Depression vs non depression forums) Less noise Great generalization	Not directly defined as a variable of interest in the problem statement
Pregnancy	Key metric of interest for the analysis perspective	Difficult to measure Noisy data



Learnings: Importance of Human Augmented Workflow in model development

- Human Augmented Workflow:

- For classification problems with rare events, and for results where the prediction accuracy is low,
- human review process can improve accuracy

Human augmented workflow

Manual data labelling
on the *real data*



Model Development



Model Validation and
manual review



Model Deployment

Fully Automated workflow (Proxy Data)

Automated data labelling using *proxy data*



Model Development



Model Validation



Model Deployment

Fully Automated workflow (Auto ML)

Use pre-build and
customized Auto ML
Packages



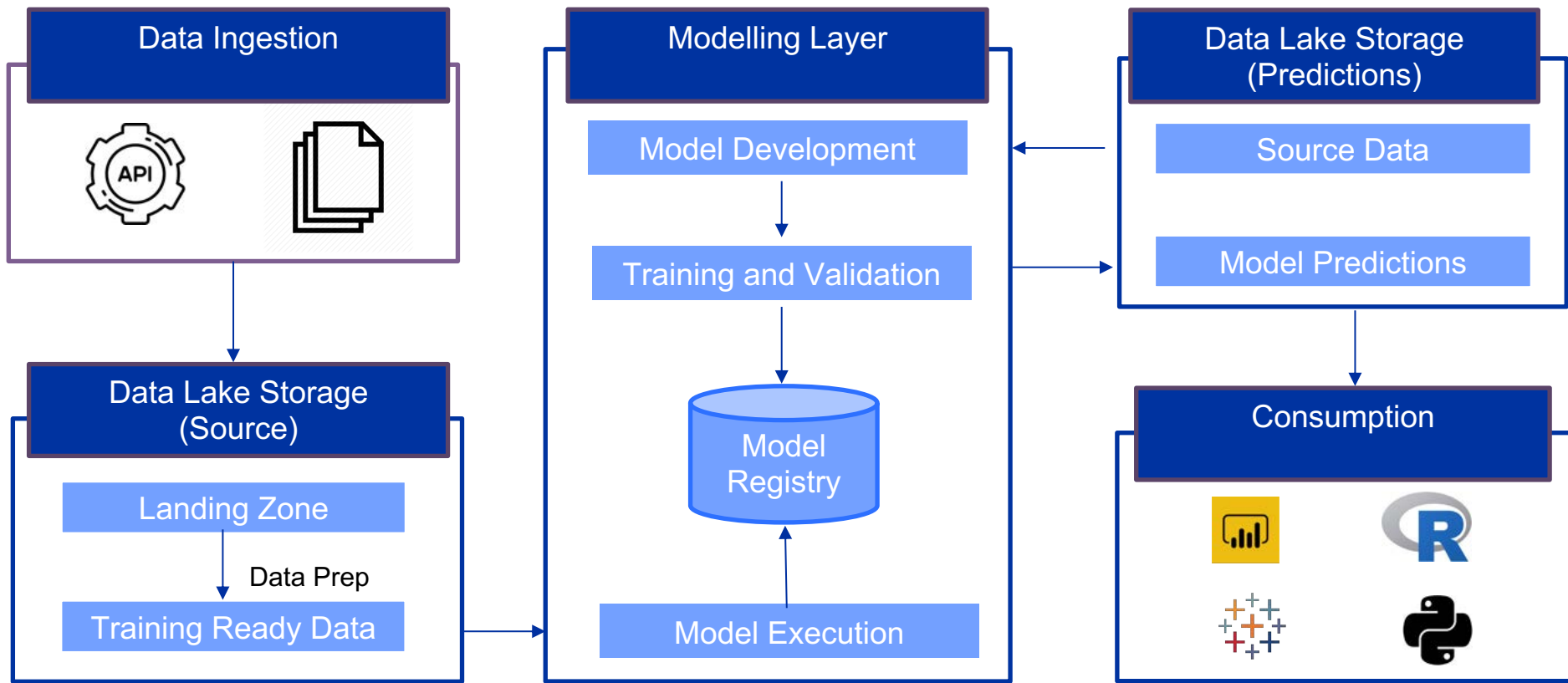
Model Validation



Model Deployment



Learnings: Data Cleansing and Data Engineering





The Computational Science Symposium

September 21st–23rd

Innovation Challenge -
Opioid/Substance Use
Disorder by Pregnant Women

By Qingying (Ally) Lu
IBM/FDA Team

phuse.eu



Panelist Bio



Qingying (Ally) Lu

- IBM Senior Managing Consultant
- Experiences using Machine Learning, Deep Learning, Natural Language Processing, and statistical modeling to gain business insights from data
- Experiences supporting FDA regulatory review process and creating tools based on CDISC standards
- INFORMS Certified Analytical Professional
- PMP Certified Project Manager



A Recap on the Challenge

Problem:

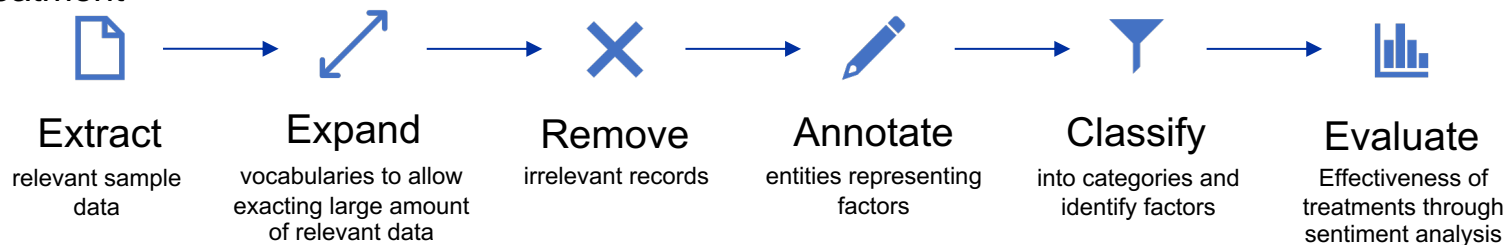
- Opioid/Substance Use Disorder by Pregnant Women

Our Team:

- IBM/FDA Joint Team
- 10 members with diverse backgrounds in Data Science, Software Engineering, Public Health, and Business

Our Solution:

- Mine social media data using Natural Language Processing, Deep Learning, and Machine Learning for insights into factors that cause pregnant women to initiate, abuse, and relapse with the use of opioids, and feasibility of intervention and treatment





Our Experiences

- Cross-organization cross-disciplinary collaboration fosters innovation
- Our systematic approach to extracting and mining social media data is built using a combination of Open-Source and proprietary software for highest efficiency
- Our approach enables extracting and mining social network data [on targeted groups or targeted events](#) from [all](#) social network platforms
- Future analyses could be done on geographic distribution and temporal distribution using social network data with geo labels and time stamps
- Potential to provide near real-time analysis results to local governments and organizations for social listening
- Additional insights could be gained by integrating data from social networks with data from other sources, e.g., data/results from this process could be integrated into the FDA Opioid Data Warehouse



The Computational Science Symposium

September 21st–23rd

PHUSE/FDA Innovation Challenge Drug Safety Surveillance Stream

Patty Hegarty, Bayer



PHUSE/FDA Data Science Innovation Challenge



DRUG SAFETY SURVEILLANCE

Define an ideal process for surveillance of multiple data sources that would be useful in the development of a pharmacovigilance tool.

DESCRIPTION

- FDA uses the term pharmacovigilance to mean all scientific and data gathering activities relating to the detection, assessment, and understanding of adverse events following the use of a drug or biologic product.
- Pharmacovigilance improves the safety of medicines by monitoring a product once it has been launched into the market.
- **Intended populations include but are not limited to:** Adults, Pediatrics, Geriatrics, Patients receiving newly approved products
- **Collected data could include but are not limited to:** Demographic, Geolocation, Product in use, Concomitant medications, Adverse event date of onset, intervention, and resolution

DATA SOURCES

Sources of data could include but are not limited to: Twitter, Facebook, openFDA (e.g. FDA Adverse Event Reporting System data), National Library of Medicine, Publicly available datasets (e.g. MIT MIMIC), DrugBank, KEGG, NDF-RT

DELIVERABLES

- Adverse event analysis dashboard that provides the ability to view and analyse adverse events across and within specific patient populations over time.
- **Use your imagination!**
- **Do something unexpected.**

Drug Safety Surveillance Stream Metrics

Stream topic elicited a tremendous response!!

- **15** submissions to the stream
- **12** accepted, **3** declined
- **5** accepted teams dropped out during preparation phase
 - **2** due to time commitment required
 - **1** due to lack of company approval of solution
 - **1** after pivot to virtual approach
 - **1** unspecified
- **7** presentations made during the “Innovation Friday” series



Drug Safety Surveillance Stream

All about DIVERSITY

PROBLEMS TACKLED

- ❖ signal detection
- ❖ AE reporting
- ❖ post market safety surveillance
- ❖ counterfeit drug detection

METHODOLOGY

- ❖ AI & Machine Learning
- ❖ data mining
- ❖ statistical modeling
- ❖ extensive use of social media & public databases

PARTICIPANTS

Organizations:

- ❖ FDA
- ❖ Academia
- ❖ BioPharma / CRO
- ❖ Technology Company

Roles:

- ❖ Data Scientists
- ❖ Medical Professionals
- ❖ Developers / Programmers
- ❖ Statisticians

SOLUTIONS PRESENTED

- ❖ dashboards
- ❖ interactive data exploration portal
- ❖ active surveillance system
- ❖ counterfeit drug detection technology





Drug Safety Surveillance Stream

THANK YOU to the Presentation Teams



Deep Dissection of Postmarket Surveillance Signal



POSTMARKET DRUG AND BIOLOGIC
PRODUCT SAFETY SURVEILLANCE

Counterfeit Drug Detection



Early Safety Signal Detection
from Public Internet Sources

IBM Watson Health

Postmarket Product Safety Surveillance
Datacise™ Integrated Safety Explorer



DeepFAERS: AI for
Pharmacovigilance
by Diving Deep
into FAERS



#PRAView
Drug Safety Surveillance



The Computational Science Symposium

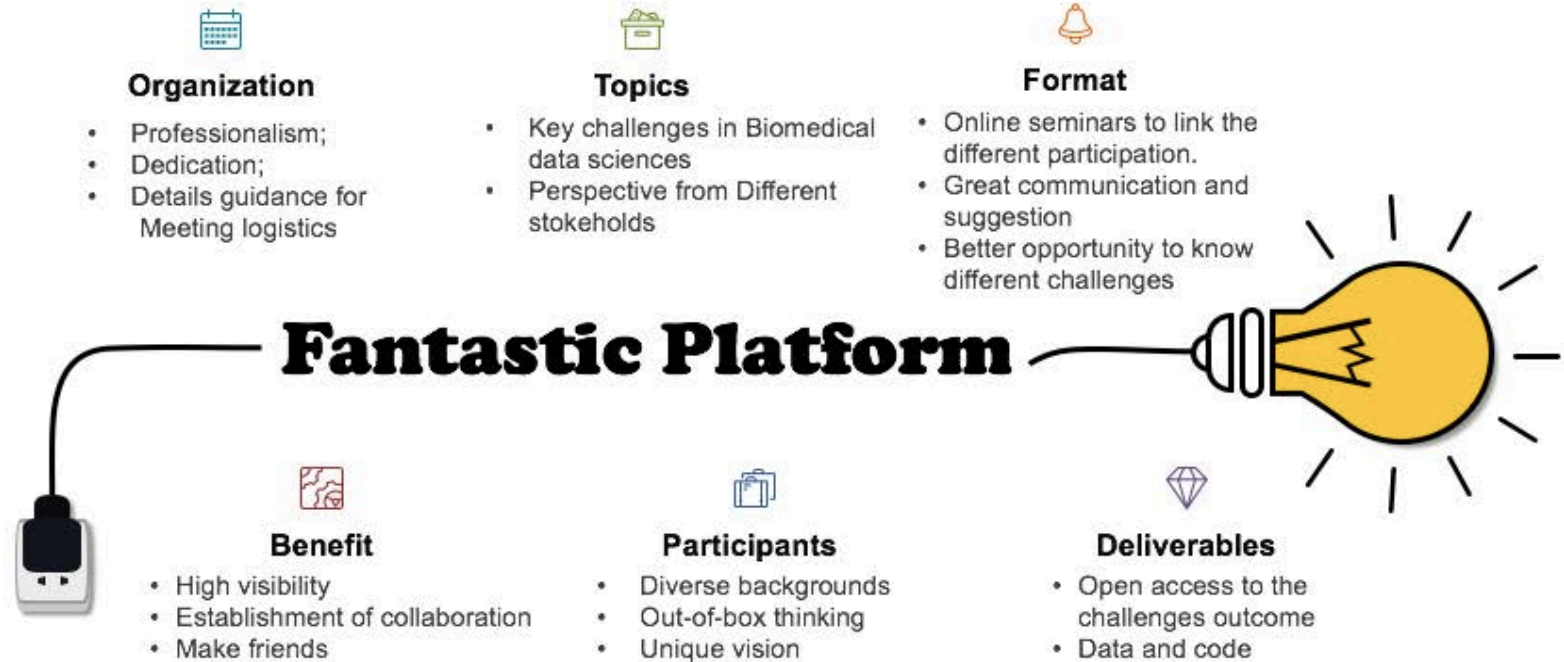
September 21st–23rd

Drug Safety Surveillance Using Social Media

Zhichao Liu, Ph.D.
DBB, NCTR, FDA
2020-09-04



Great Experience - Our Story



phuse·go·virtual
US 2020

The Computational Science Symposium

September 21st–23rd

PHUSE/FDA Data Science Innovation Challenge
Predicting Drug Interactions Stream

Bev Hayes, Janssen Research & Development



PHUSE/FDA Data Science Innovation Challenge



APPROACH FOR PREDICTING DRUG INTERACTIONS

Overcome the limitations of traditional methods for mining post-market data sources or observational healthcare data to better predict drug interactions.

DESCRIPTION

We are seeking **novel approaches to collect and analyse post-market and observational health data to predict potential drug-drug interactions.**

Intended populations include: Geriatrics, Immunosuppressed patients, Patients with a high number of prescribers

Collected data could include but are not limited to: Demographic, Number of prescribers per patient, Currently prescribed medications, Diseases and conditions

DATA SOURCES

Sources of data could include but are not limited to:

openFDA, DrugBank, KEGG (Kyoto Encyclopedia of Genes and Genomes), NDF-RT (National Drug File - Reference Terminology)

DELIVERABLES

Integration of multiple sources of adverse event data to visualise and analyse the prevalence of drug interactions in the specified populations.

Predictive model for identifying potential drug interactions across different patient populations at risk.



Innovation Fridays



A predictive model using machine learning with visualization capability in Tableau



Prediction of Drug Interactions with Real-World Data



Real-world Evidence Drug-Drug Interaction (REDDI) Match

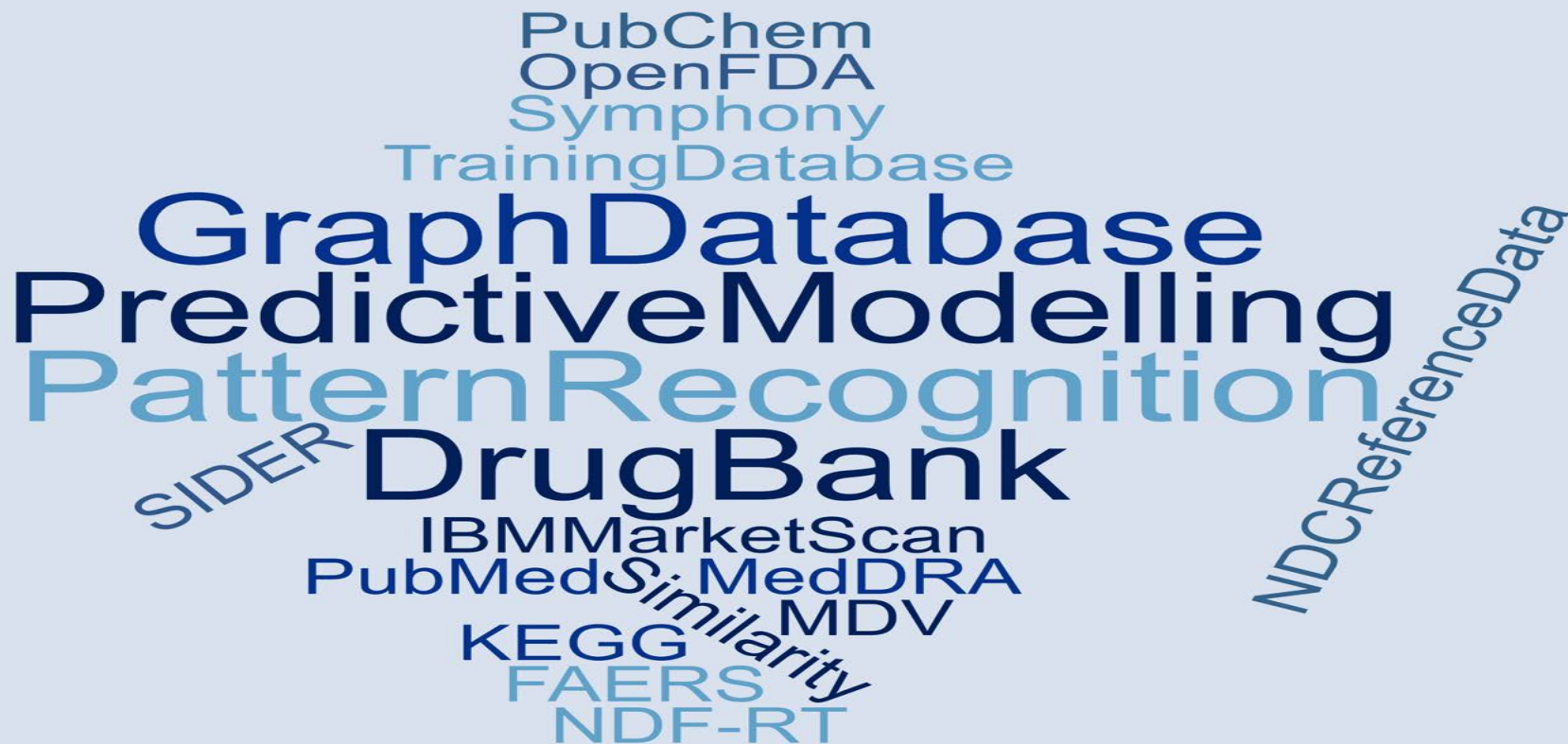


Prediction of Drug-Drug Interactions (DDI) with Artificial Intelligence



Approaches in Categorizing and Predicting Drug-Drug Interactions using Artificial Intelligence and SAS Visual Text Analytics

Using Data Science to Develop an Approach to Predict Drug Interactions



Thank you ...

Prediction of Drug-Drug Interactions (DDI) with Artificial Intelligence

Zhichao Liu, Ph.D. and Weida Tong, Ph.D.
Division of Bioinformatics & Biostatistics (DBB),
National Center for Toxicological Research (NCTR),
U.S. Food and Drug Administration (FDA)
03-09-2020



Breakthroughs that
change patients' lives

Prediction of Drug Interactions with Real-World Data

James J. Kim

PHUSE/FDA Data Science Innovation Challenge
Approach for Predicting Drug Interactions



Real-world Evidence Drug-Drug Interaction (REDDI) Match

Kathleen Mandziuk, MPH, RN, Ashwin Kumar Rai, Kerry Deem
9 March 2020

PHUSE/FDA: Approach for Predicting Drug Interactions

Olajide Ajayi, Thuy Luu and Jeff Allen
Data Science Innovation Challenge
Orlando, FL
March 9th, 2020



Approaches in Categorizing and Predicting Drug-Drug Interactions using Artificial Intelligence and SAS® Visual Text Analytics

Soundarya Palanisamy (Msc. - Data Scientist)
Haider Ali (Msc. - Data Scientist)
Muhammad Hamiz Ahmed (Msc. - Data Scientist)
Luiza Regenold (Msc. - Clinical Pharmacist)
Rhim Shu (MD, MPH, MSc. - Endocrinologist, Public Health & HTA Expert)



The background of the slide features a complex network of white lines connecting various nodes, some of which are represented by blue and white circles. Overlaid on this network are several hexagonal shapes in various shades of blue, some with white outlines and others filled with solid color. The overall aesthetic is technical and digital, suggesting a focus on computational science.

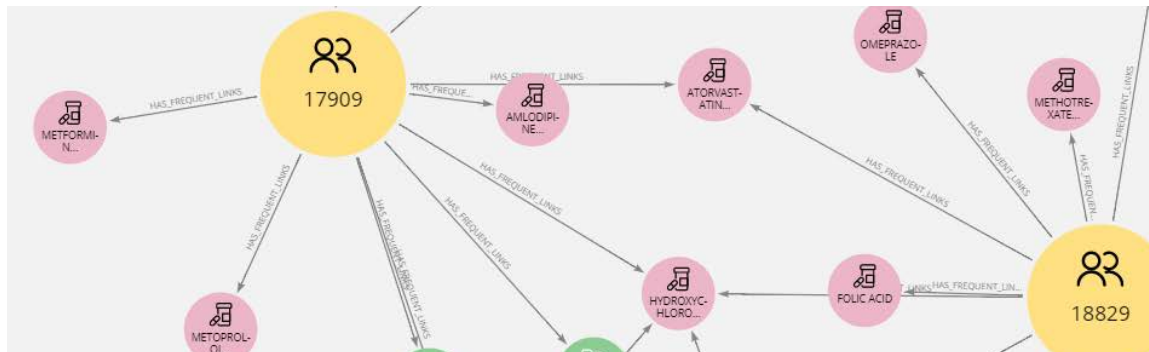
phuse·go·virtual
US 2020

The Computational Science Symposium

September 21st–23rd

Kathleen Mandziuk, *PRA Health Sciences*

- Fill gap of post-marketing polypharmacy safety
- Dynamic detection - flexible
- Real-time information
- User-friendly
- Depth & Breadth of data and insights





PHUSE Challenge Take-aways

- New solution development
- Multi-disciplinary & partnership team
- Delay and virtual – solution expansion
- FDA & Industry feedback
- Launching for real world use – internal & external



The Computational Science Symposium

September 21st–23rd

**PHUSE/FDA Data Science Innovation Challenge
Feedback Survey**

Bev Hayes
Janssen Research & Development



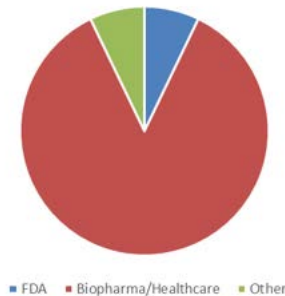
Survey Population

Audience:

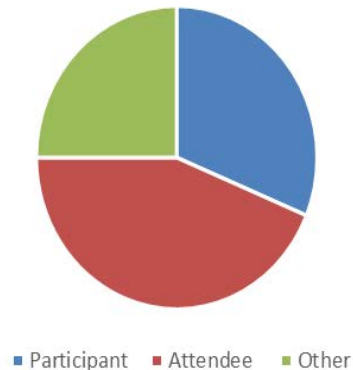
PHUSE members and innovation submissions

- 14 Responders: **largely from BioPharma/Healthcare** (includes CROs)
- **5 of 14** respondents **participated** in the challenge
- **Well socialized** (6 of 14 respondents heard about the challenge *outside of PHUSE*)

Business affiliation



Association with the Innovation Challenge

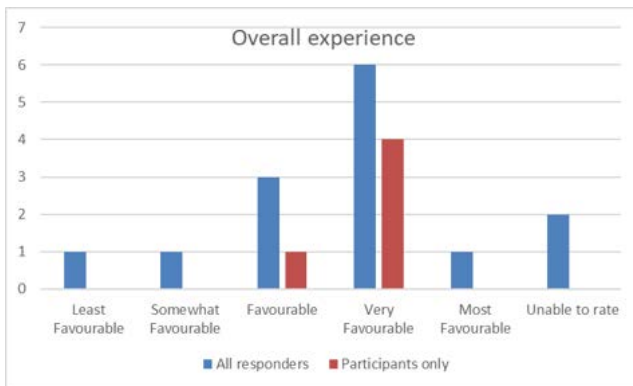
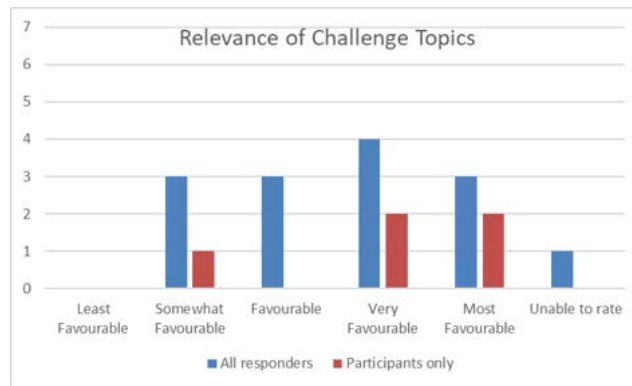
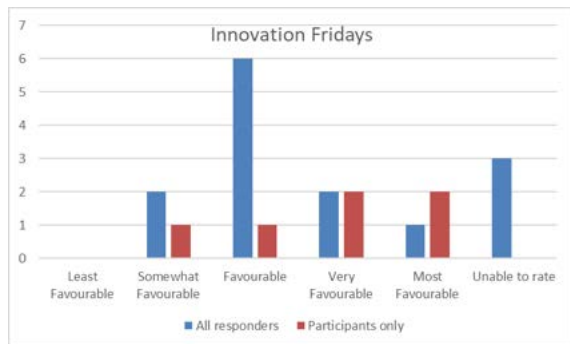
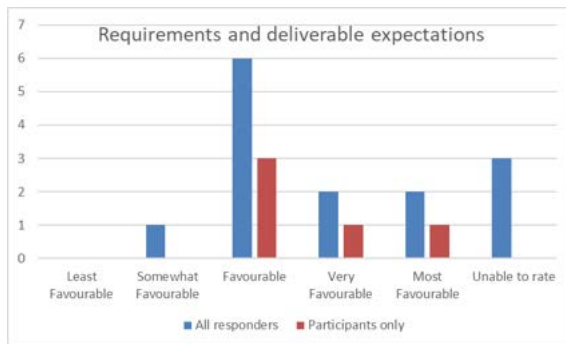
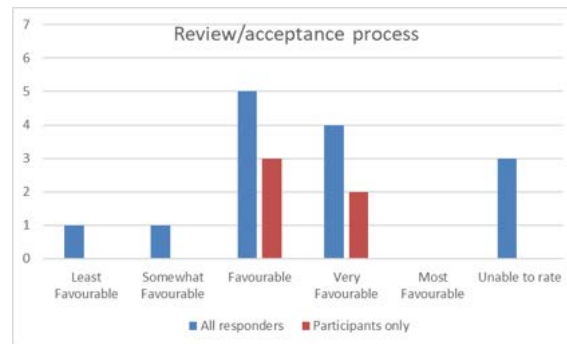


How did you hear about the PHUSE/FDA Data Science Innovation Challenge?





Favorability Ratings





Feedback

12/14 Respondents indicate they would be interested in participating in a **future innovation challenge**

What we did well

- ✓ Virtual organization/management worked well
- ✓ Publicity (within PHUSE) was good
- ✓ Interesting concept to inspire people to try out innovation

How we can improve

- ❑ Innovation topics related to data submission (e.g. xpt vs. XML, JSON, RDF based, etc...)
- ❑ Innovation topics related FDA requirements translated into CDISC implementation
- ❑ Allow plenty of lead time with clear requirements

Most respondents felt that **slide presentations** are the desired means to present solution

- **some** recommend slides with a **demo/prototype**



Conclusion



Although small sample size, generally a positive experience



A step forward in the industry, sparking innovation



An appetite for more, broaden the topics



Time commitment may be a challenge

Consider this when defining scope, requirements and deliverables

To view the recorded sessions of
Innovation Fridays 2020 go to
<https://www.phuse.eu/phuse-fda-data-science-innovation-challenge>



**The premier community for people
working in the biometric area**

Contact Channels

📞 UK +44 1843 609600
📞 USA +1 609 514 5105
✉ office@phuse.eu
🌐 [phuse.eu](https://www.phuse.eu)

Social Media

🐦 [@phusetwitta](https://twitter.com/phusetwitta)
📘 [/phusebook](https://www.facebook.com/phusebook)
▶ [/phusetube](https://www.youtube.com/channel/UCphusetube)
🌐 [/company/phuse](https://www.linkedin.com/company/phuse)

phuse·go·virtual
US 2020

The Computational Science Symposium

September 21st–23rd

phuse.eu



Rome wasn't built in a day
But they were laying bricks every hour



PHUSE/FDA Data Science Innovation Challenge



Opioid/Substance Use
Disorder



Drug Safety
Surveillance



An Approach for
Predicting Drug
Interactions



Listen; learn
Redesign; deliver





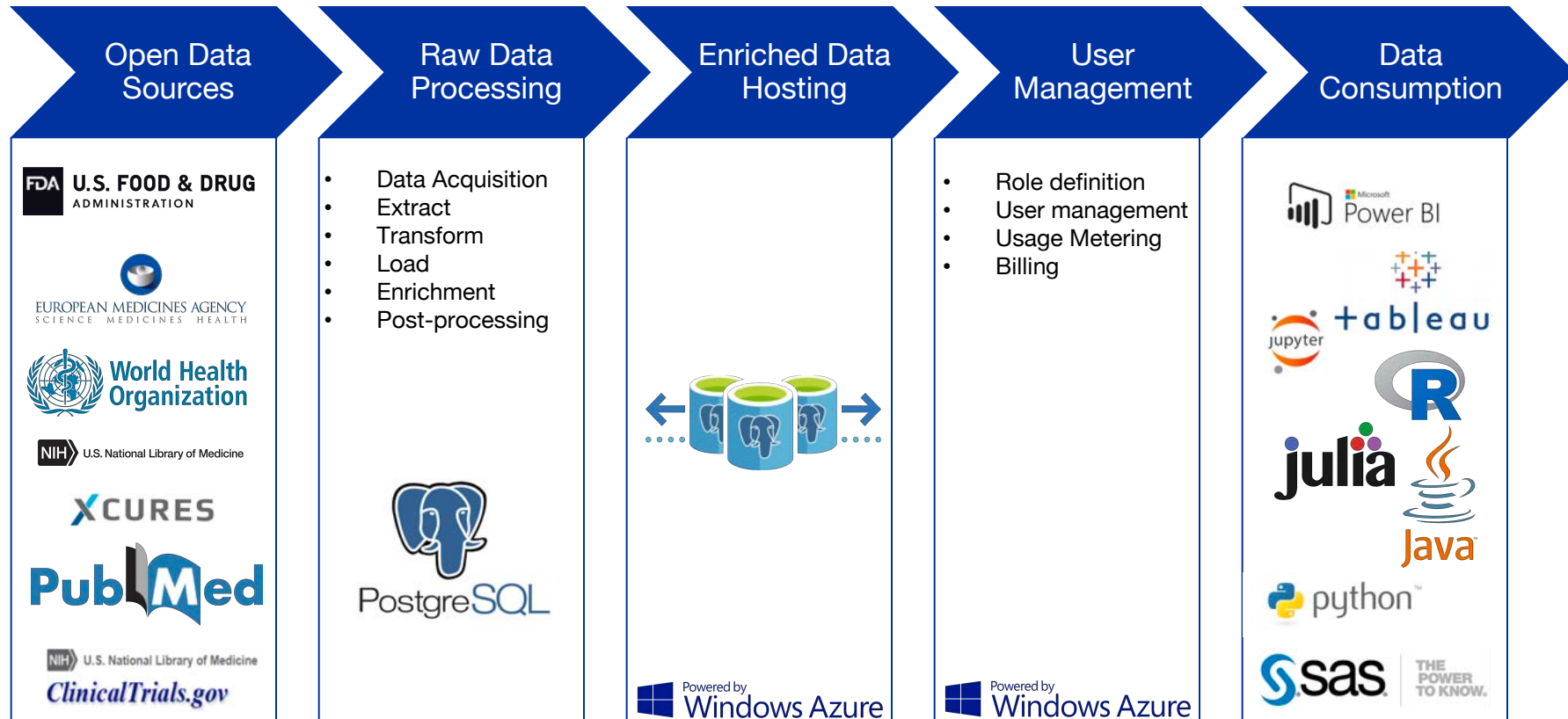


Shoot for the stars



- Increase member engagement in WGs, events
- “From Zero to Analysis” in under five minutes
- Accelerate discovery of solutions to challenges
- Learn new tools (Jupyter, Python); Best Practices
- Members quickly launch collaborative research
- Make data analytics accessible to all user levels
- Minimize any IT involvement

System Process Flow



Next Steps for PODR



First beta release to
interested PHUSE
members



Build more use cases



Working with software
companies to add in
more capabilities



Looking for other data
sources to load



Designing the second
PHUSE/FDA Data
Science Innovation
Challenge



First Announcement



**The second PHUSE/FDA Data Science
Innovation Challenge will run in 2021**



4Q20 is the design phase – all ideas welcome: workinggroups@phuse.eu
Dedicated discussion/session At the PHUSE U.S. SDE 19th Nov 2020
Challenge launches January 2021



Presentation / award event September 16th 2021
(day after CSS 2021)

A wooden signpost with two directional signs. The top sign is a light-colored wooden arrow pointing right, with the word "QUESTIONS" in bold black capital letters. The bottom sign is a light-colored wooden arrow pointing left, with the word "ANSWERS" in bold black capital letters. The signpost is made of a dark wooden pole. The background is a bright blue sky with scattered white clouds.

QUESTIONS

ANSWERS

phuse·go·virtual
US 2020

The Computational Science Symposium

September 21st–23rd

Thank for attending the PHUSE/FDA Virtual CSS 2020

phuse.eu



EU Connect 2020

The Clinical Data Science Conference

9th–13th November

phuse.eu



US 2021

The Clinical Data Science Conference

June 13th-16th
Renaissance Orlando at
SeaWorld Florida

phuse.eu



US 2021

The Computational Science Symposium

September
Silver Spring Civic Center
Silver Spring, Maryland, USA

phuse.eu