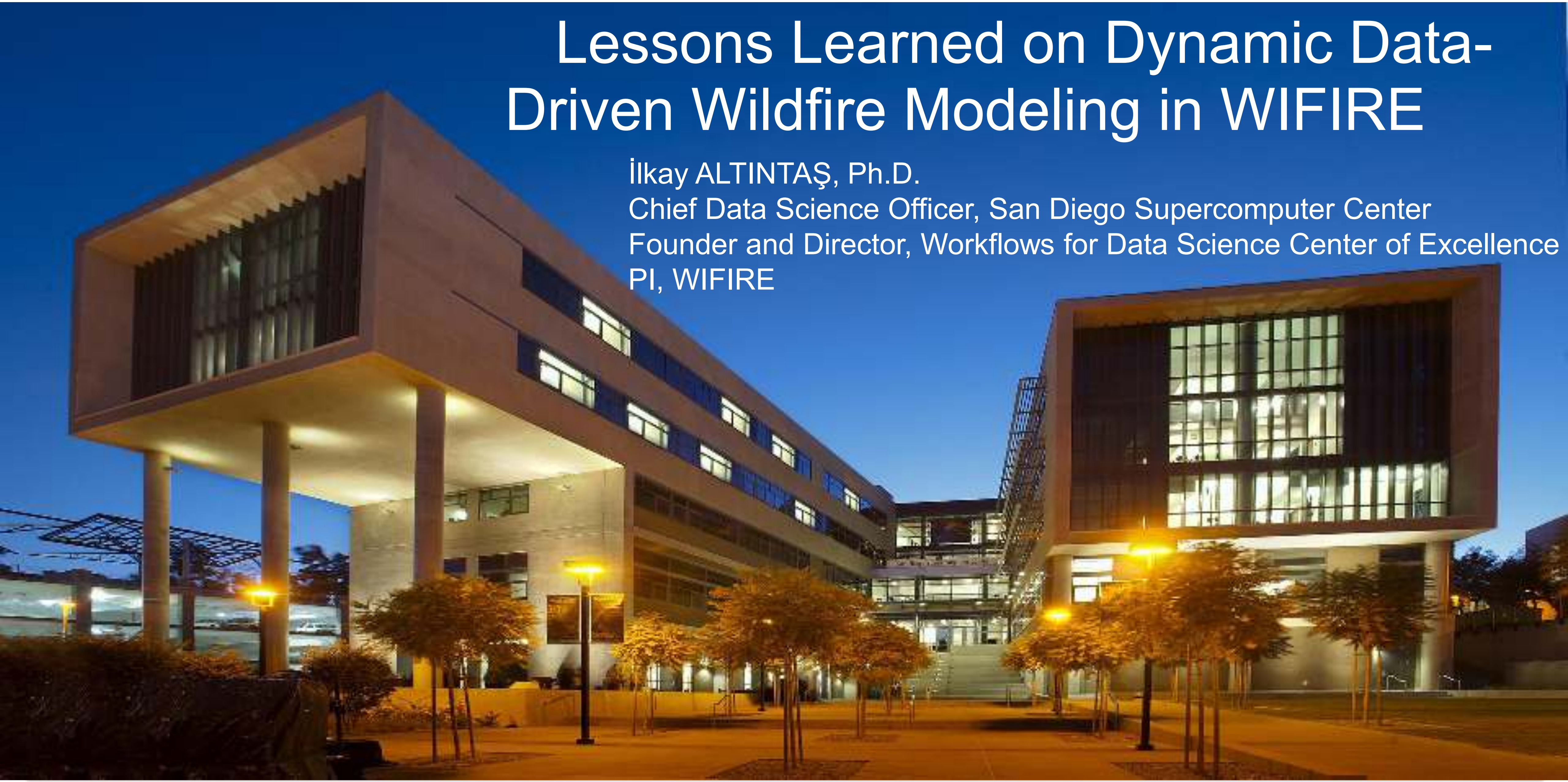
Lessons Learned on Dynamic Data-Driven Wildfire Modeling in WIFIRE

İlkay ALTINTAŞ, Ph.D.
Chief Data Science Officer, San Diego Supercomputer Center
Founder and Director, Workflows for Data Science Center of Excellence
PI, WIFIRE



UC San Diego

WIFIRE

WIFIRE is funded by NSF 1331615

How do we Better Predict Wildfire Behavior?

Fire is Part of the Natural Ecology….

… but requires Monitoring, Prediction and Resilience

- Wildfires are critical for ecology, but volatile
- Fuel load is high due to fire suppression over the last century
- Drought, higher temperatures
- Better prevention, prediction and maintenance of wildfires is needed


Photo of Harris Fire (2007) by former Fire Captain Bill Clayton

Disaster management of (ongoing) wildfires heavily *relies on understanding* their Direction and Rate of Spread (RoS).

What was lacking is…

a dynamic system integration of
real-time sensor networks, satellite imagery, near-real time data management tools, wildfire simulation tools, and connectivity to emergency command centers

…. before, during and after a firestorm.

WIFIRE is funded by NSF 1331615

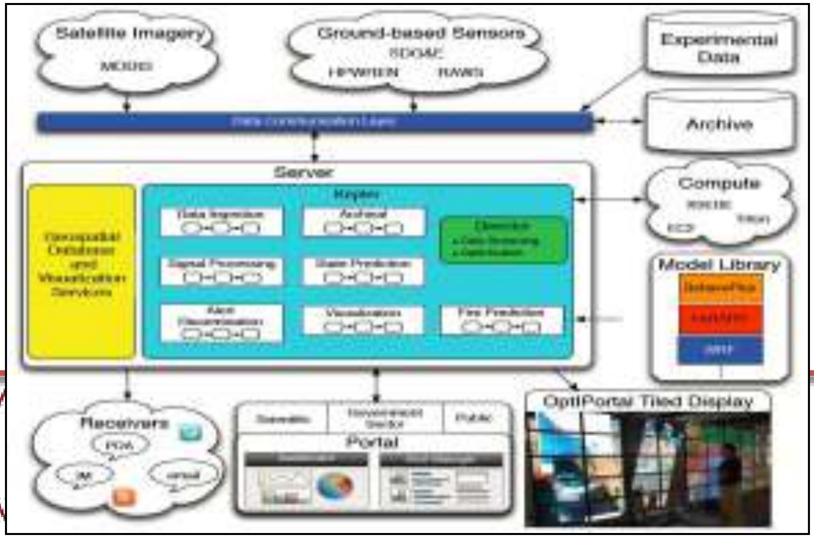
WIFIRE: A Scalable Data-Driven Monitoring, Dynamic Prediction and Resilience Cyberinfrastructure for Wildfires

Big Data

Monitoring

Visualization

Fire Modeling



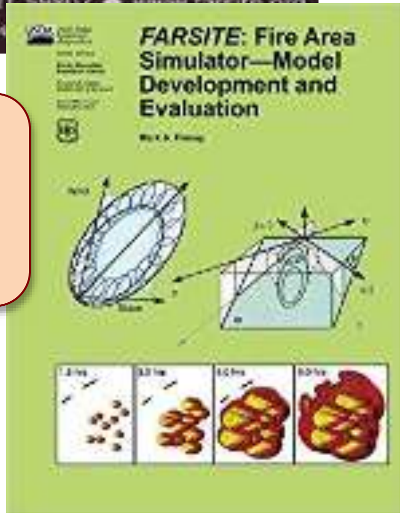
High Performance Wireless Research and Education Network

Major success to bring internet to incident command in the field. Used in over 20 fires over time.

http://hpwren.ucsd.edu/cameras

>160 Meteorological Sensors and Growing

FARSITE

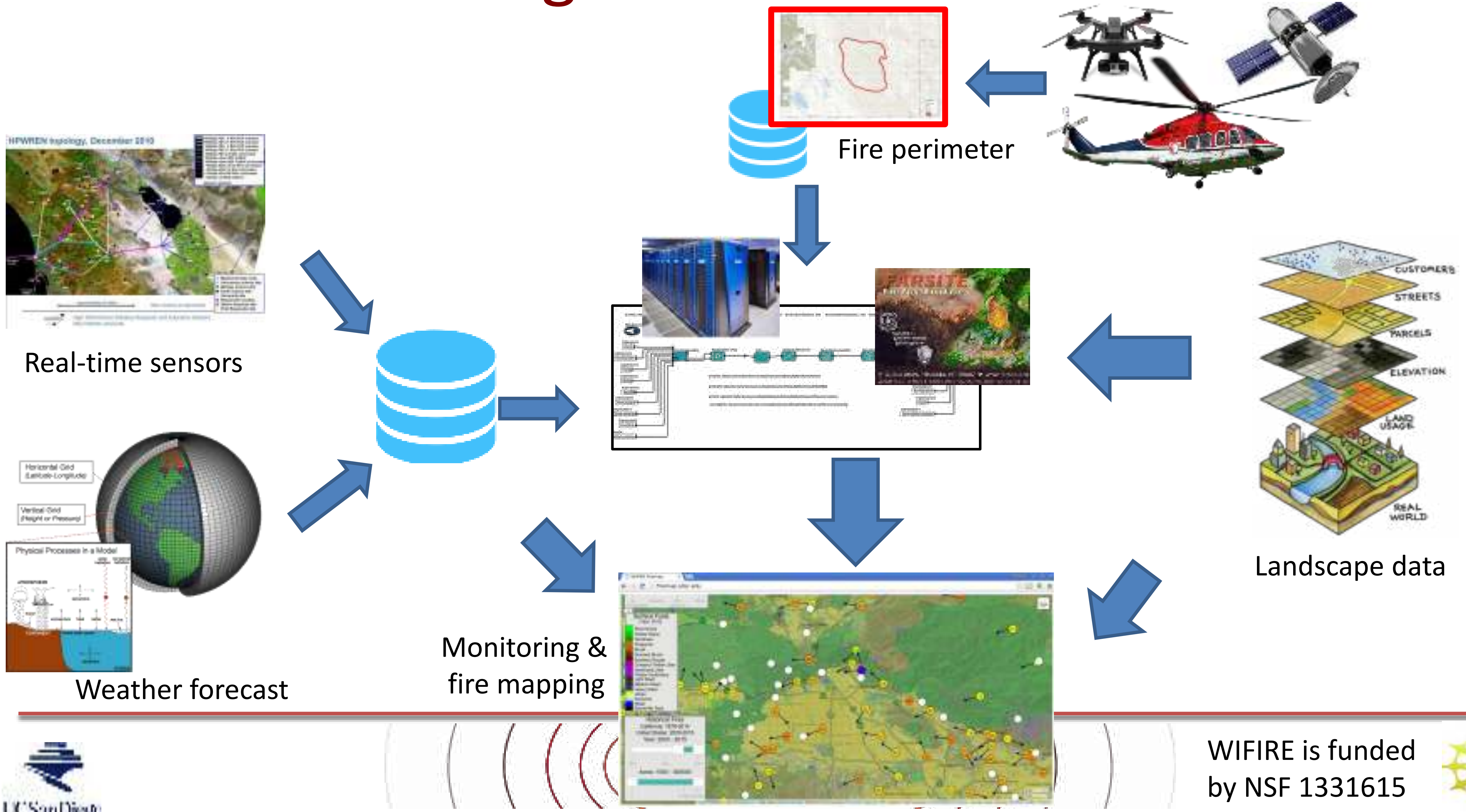


Most popular operational fire behavior modeling system.

WIFIRE

WIFIRE is funded by NSF 1331615

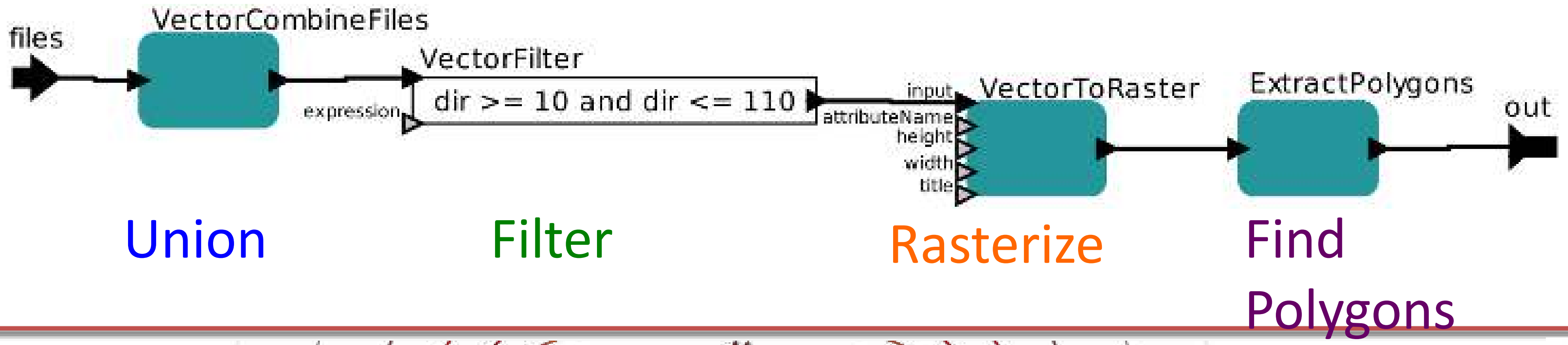
Fire Modeling Workflows in WIFIRE



WIFIRE is funded by NSF 1331615

Many Integrated Workflows and Processing Tools

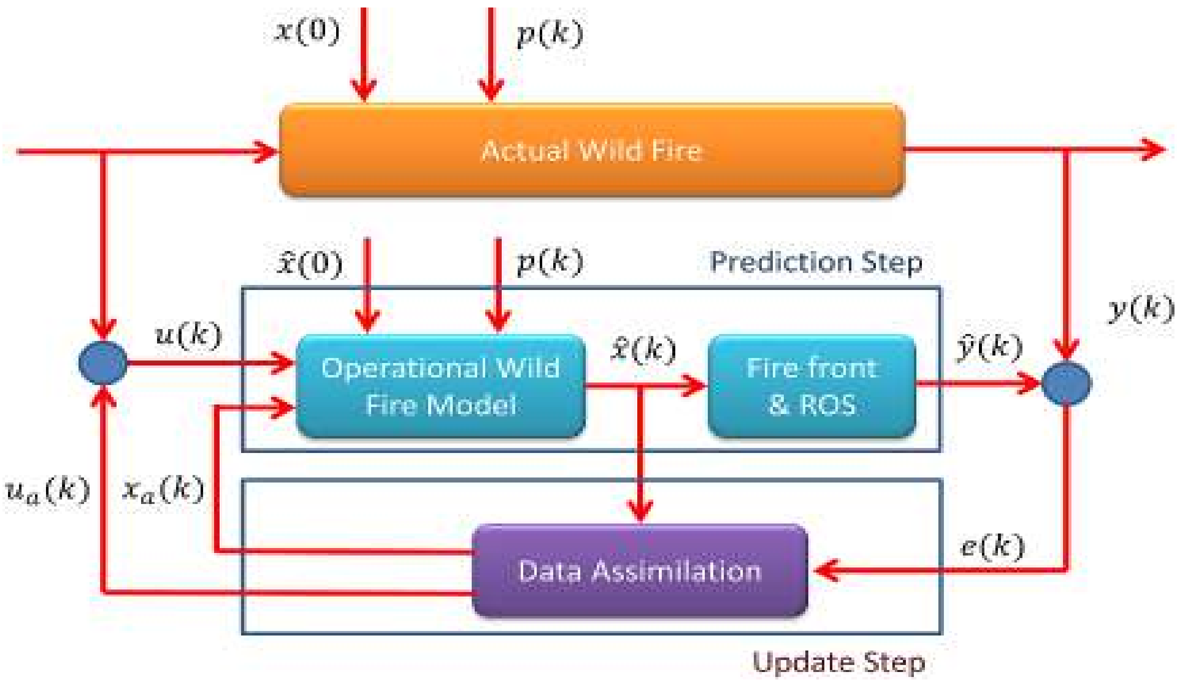
- Components for modeling and data tools
- Understanding of geo data formats
- Transparent interaction with data and computing resources
- Open source and extensible



WIFIRE is funded by NSF 1331615

Closing the Loop using Big Data

-- Wildfire Behavior Modeling and Data Assimilation --

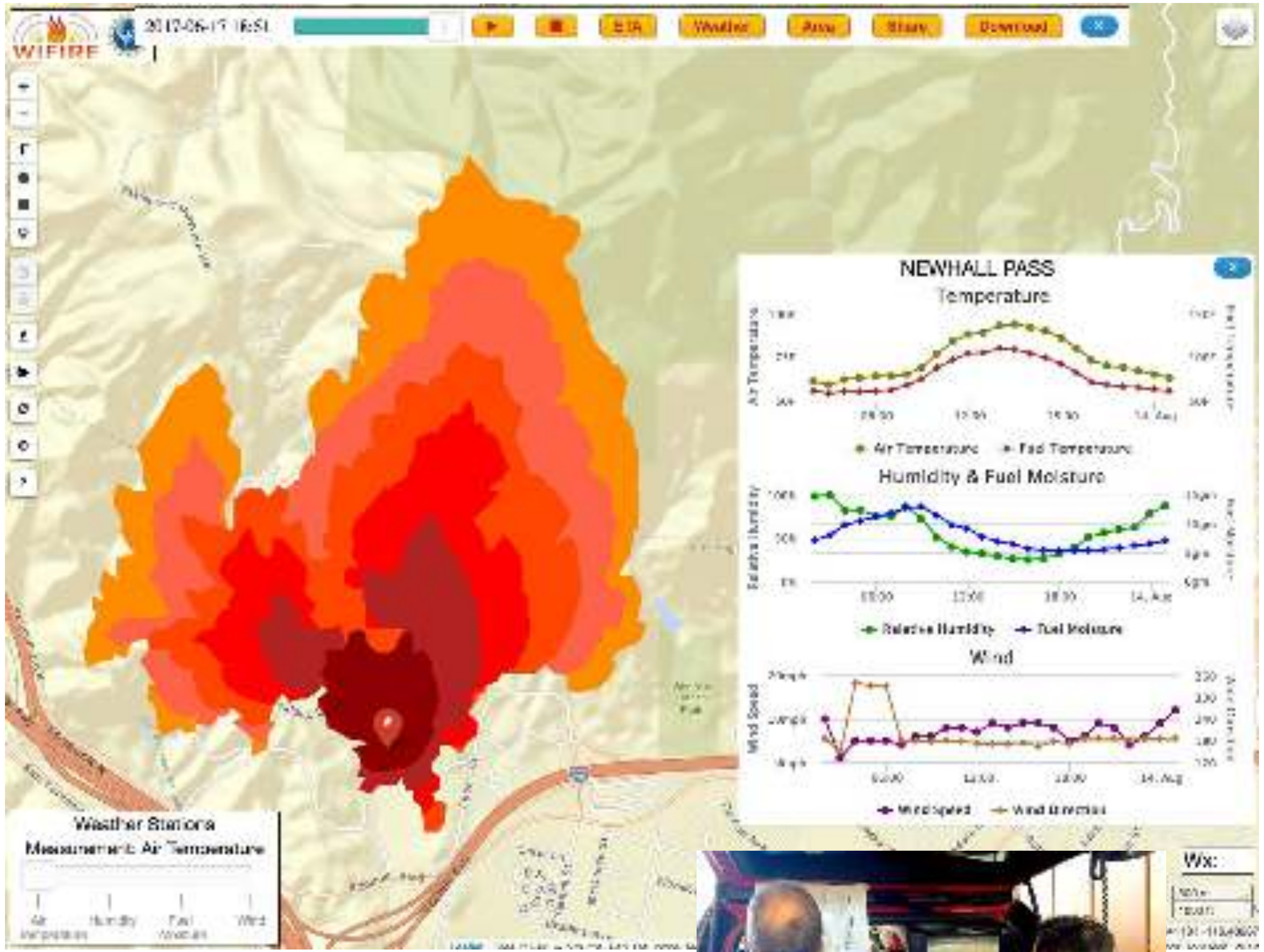


Conceptual Data Assimilation Workflow with Prediction and Update Steps using Sensor Data

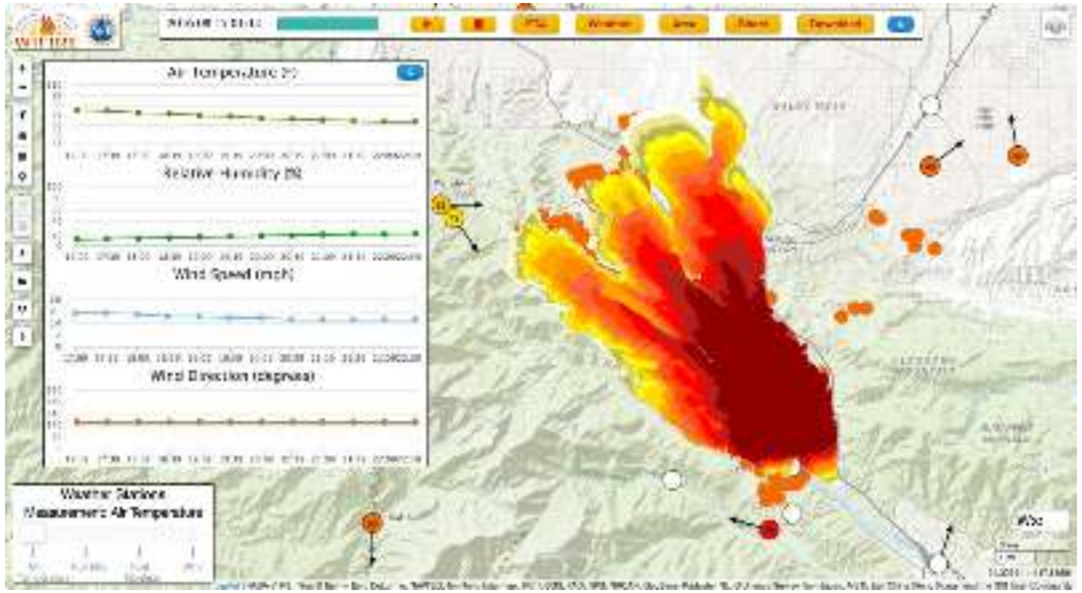
- Computational costs for existing models too high for real-time analysis
- *a priori -> a posteriori*
  - Parameter estimation to make adjustments to the (input) parameters
  - State estimation to adjust the simulated fire front location with an a posteriori update/measurement of the actual fire front location

WIFIRE is funded by NSF 1331615

Data-Driven Fire Progression Prediction Over Three Hours



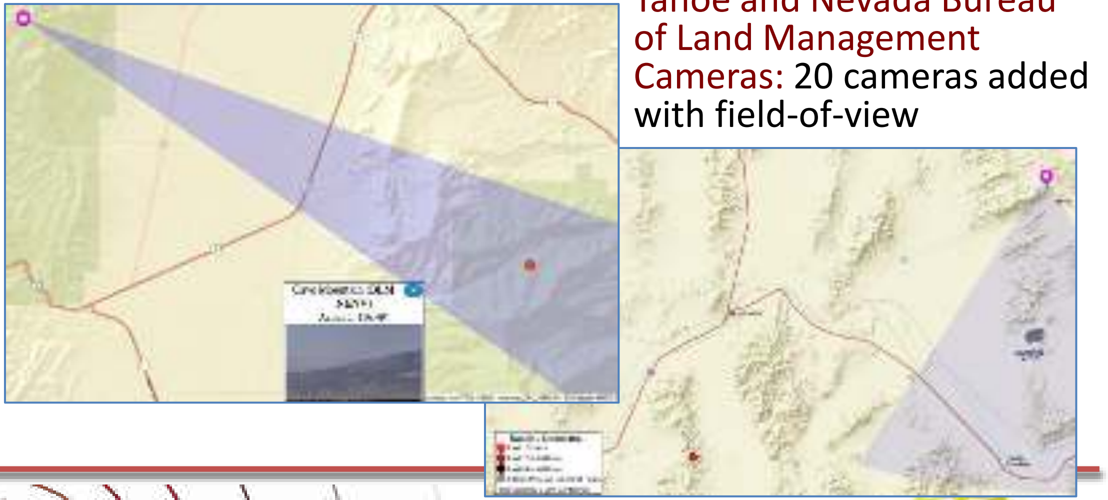
August 2016 – Blue Cut Fire



Collaboration with LA and SD Fire Departments

http://firemap.sdsc.edu

Tahoe and Nevada Bureau of Land Management Cameras: 20 cameras added with field-of-view

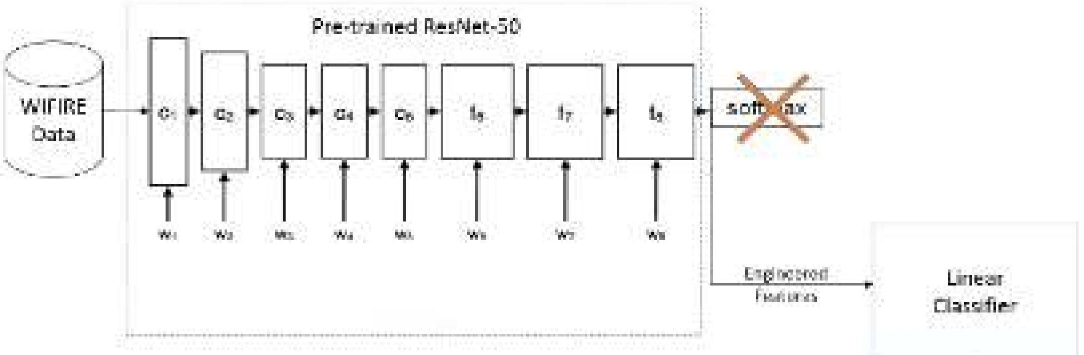


WIFIRE is funded by NSF 1331615

Some Machine Learning Case Studies

- Smoke and fire perimeter detection based on imagery
- Prediction of Santa Ana and fire conditions specific to location
- Prediction of fuel build up based on fire and weather history
- NLP for understanding local conditions based on radio communications
- Deep learning on multi-spectra imagery for high resolution fuel maps
- Classification project to generate more accurate fuel maps (using Planet Labs satellite data)

Classification project to generate more accurate fuel maps



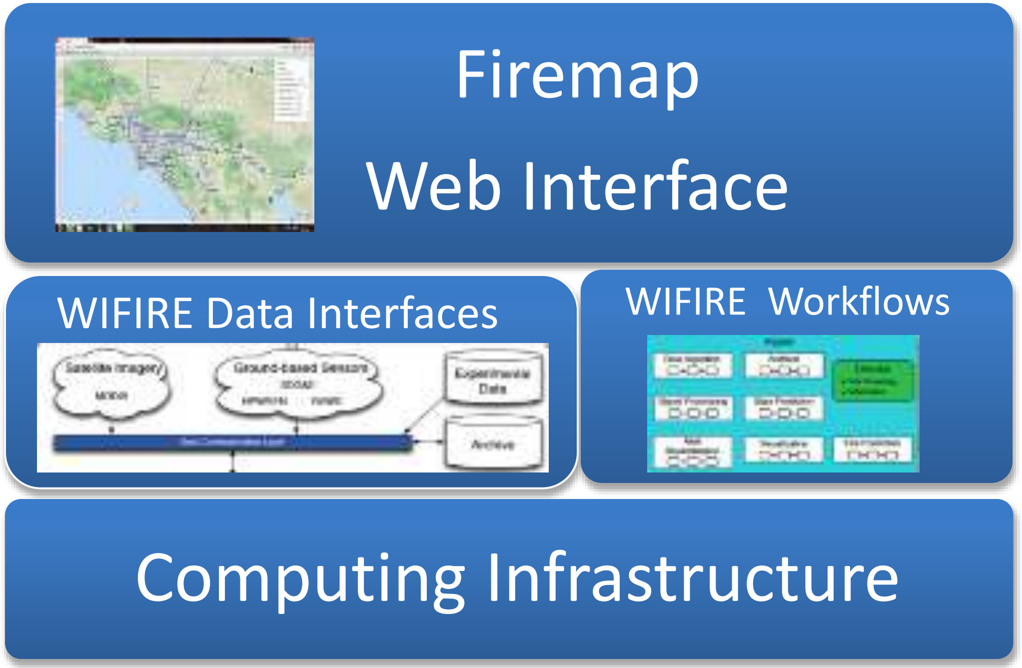
- Accurate and up-to-date fuel maps are critical for modeling wildfire rate of speed and potential burn areas.
- Challenge:
  - USGS Landfire provides the best available fuel maps every two years.
  - The WIFIRE system is limited by these potentially 2-year old inputs. Fuel maps created at a higher temporal frequency is desired.
- Approach:
  - Using high-resolution satellite imagery and deep learning methods, produce surface fuel maps of San Diego County and other regions in Southern California.
  - Use LandFire fuel maps as the target variable, the objective is create a classification model that will provide fuel maps at greater frequency with a measure of uncertainty.

Cluster 1: Short Grass

WIFIRE is funded by NSF 1331615

Firemap Tool

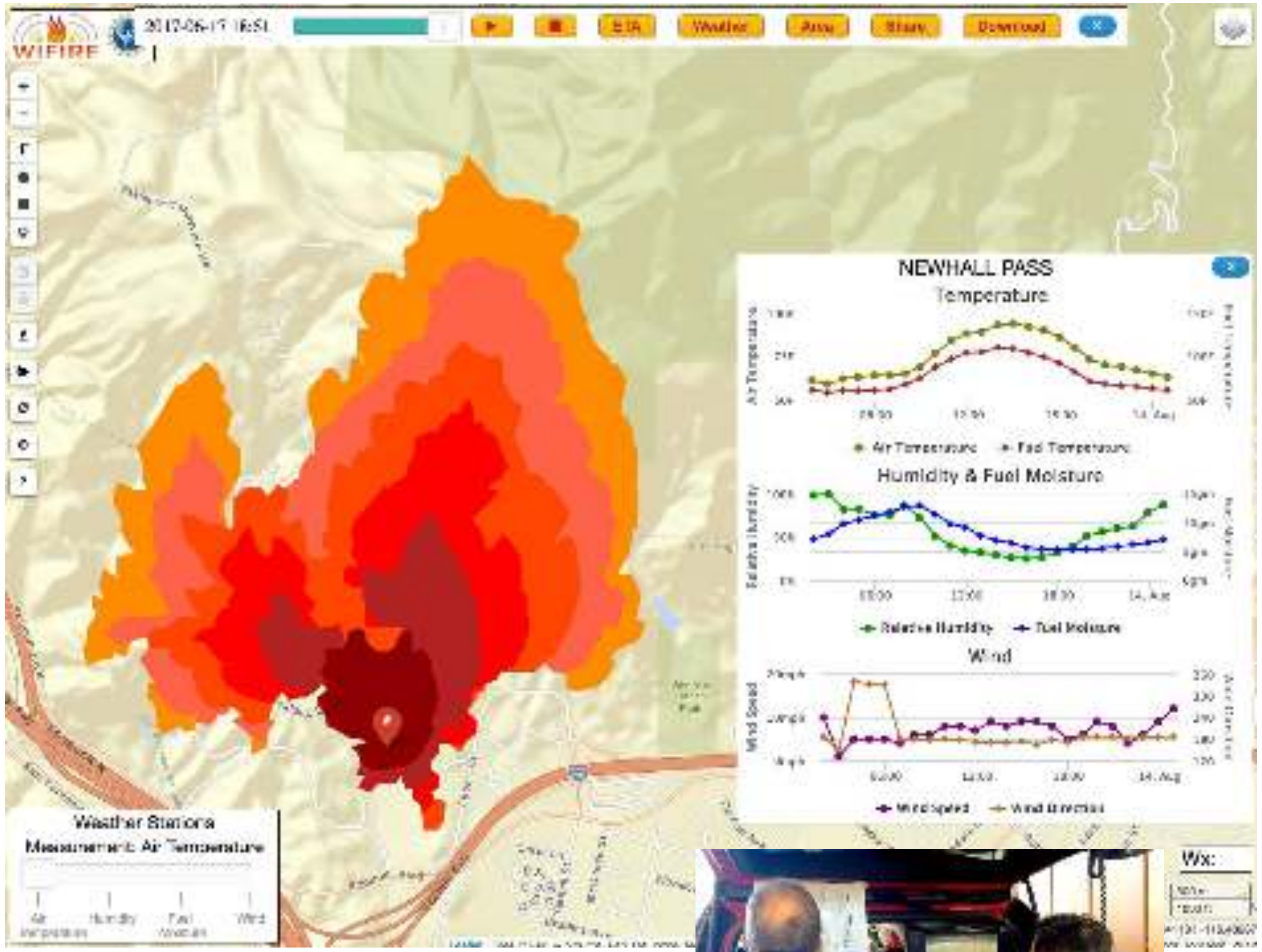
http://firemap.sdsc.edu



- A web-based GIS environment:
  - access information related to fire behavior
  - analyze what-if scenarios
  - model real-time fire behavior
  - generate reports
- Powered by WIFIRE

WIFIRE is funded by NSF 1331615

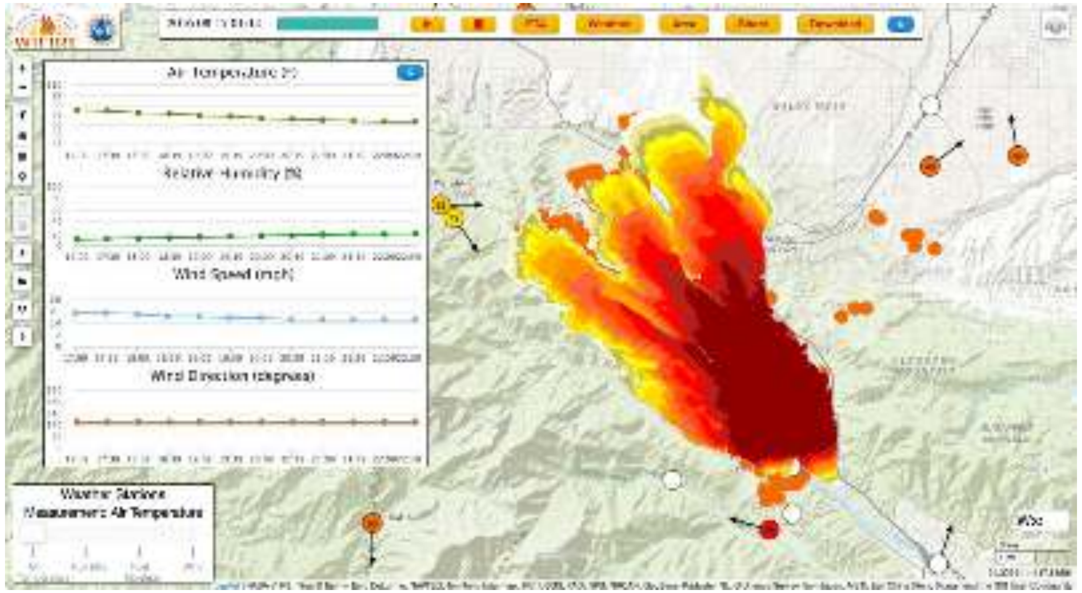
Data-Driven Fire Progression Prediction Over Three Hours



Collaboration with LA and SD Fire Departments

http://firemap.sdsc.edu

August 2016 – Blue Cut Fire



Tahoe and Nevada Bureau of Land Management Cameras: 20 cameras added with field-of-view

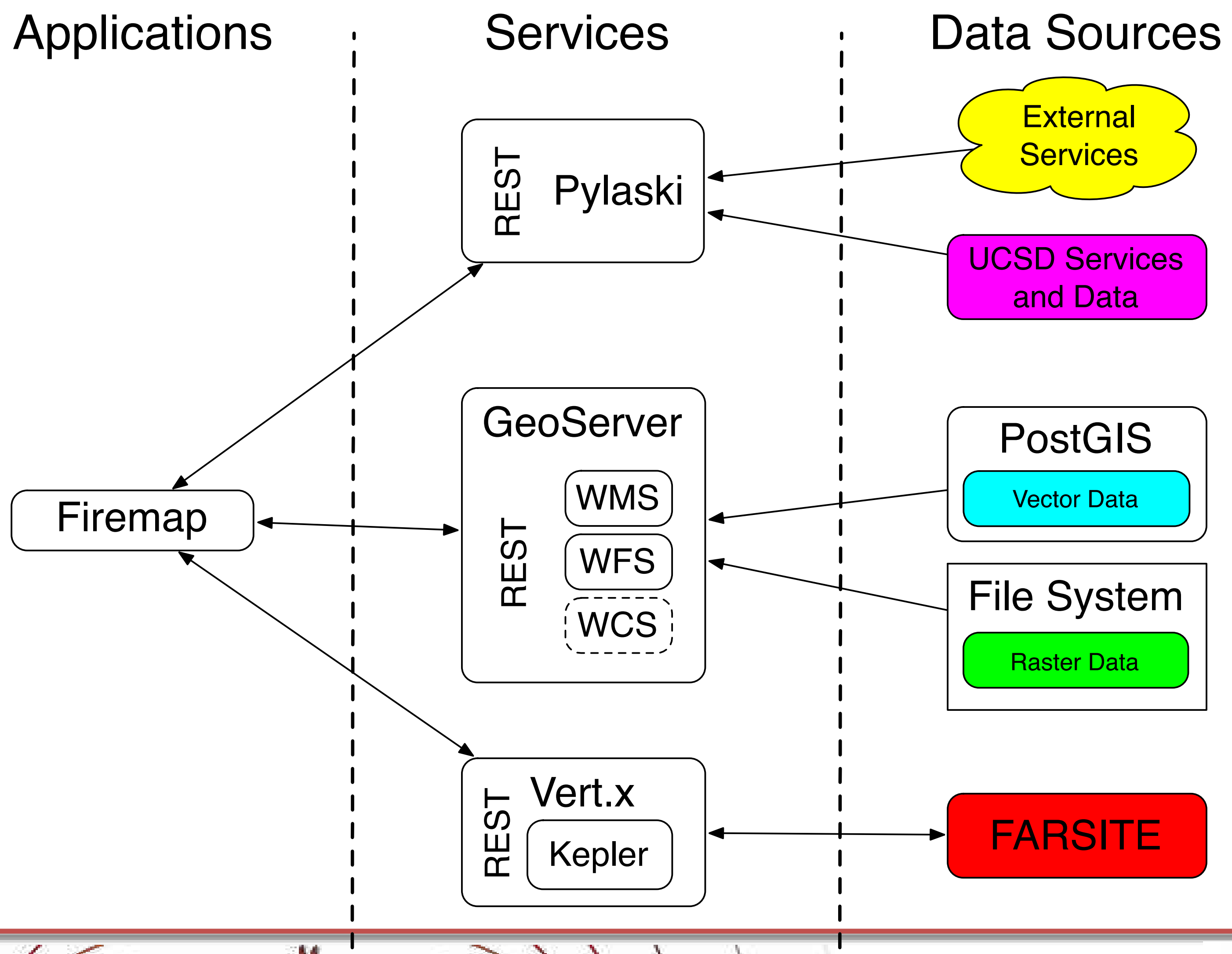
WIFIRE

WIFIRE is funded by NSF 1331615

UC San Diego

Firemap Overview

Architecture



WIFIRE is funded by NSF 1331615

Firemap Layers

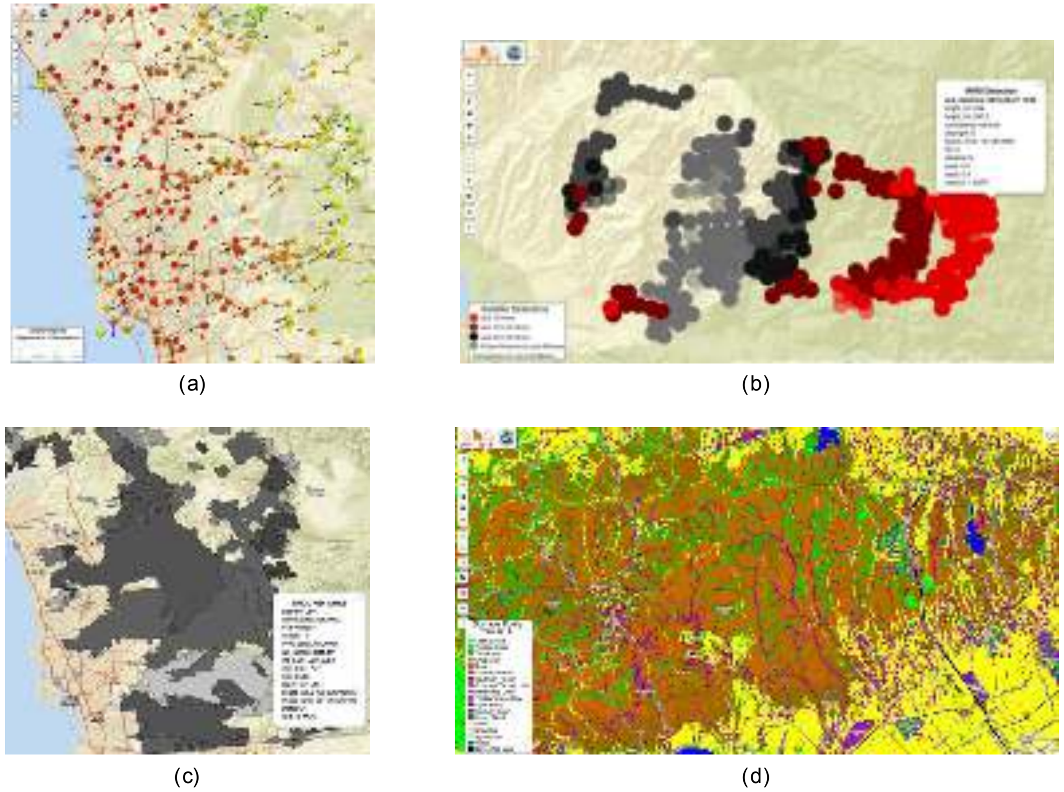
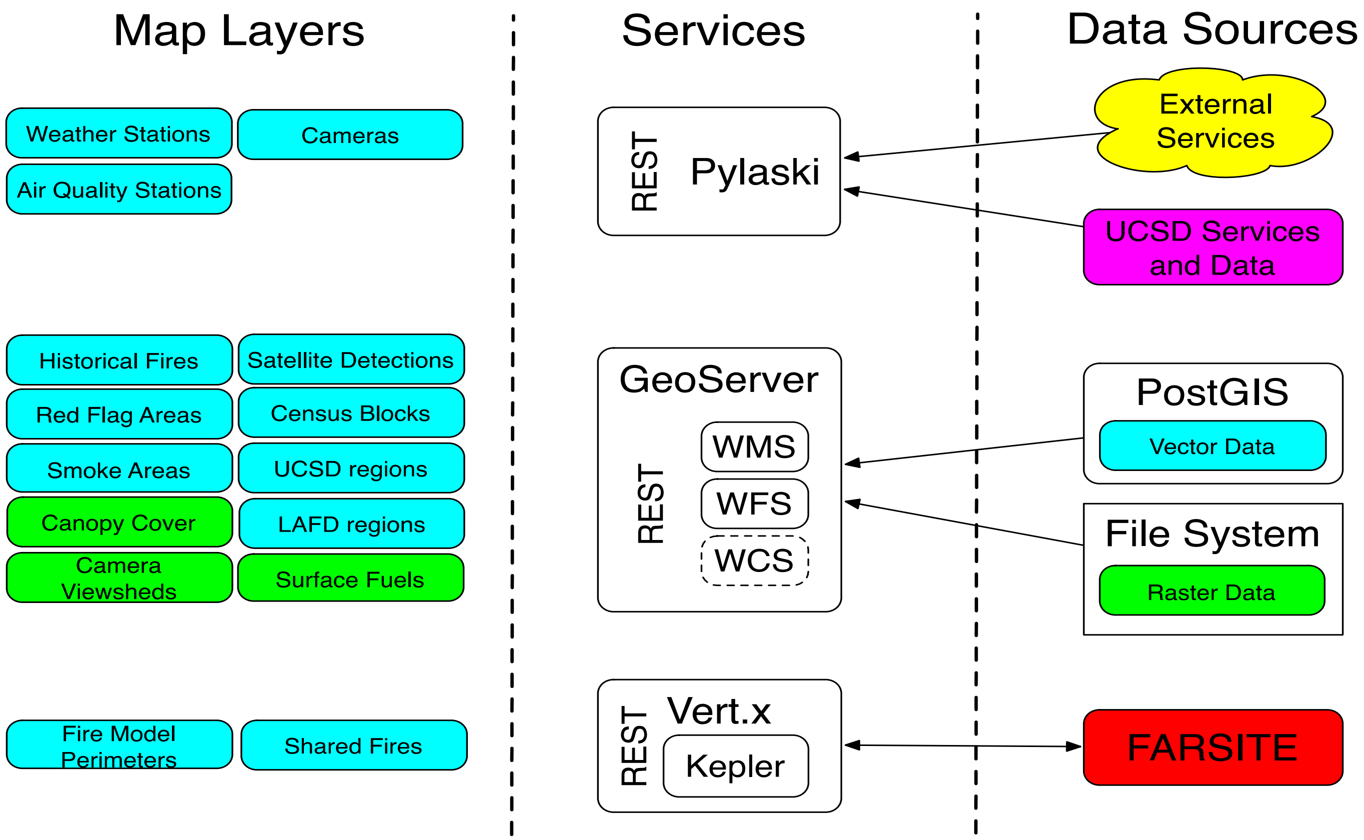


Figure 2: Firemap layers: (a) Weather Stations Layer showing MesoWest stations in San Diego County during Santa Ana conditions on 26 September, 2016; (b) VIIRS Layer showing 7 days of thermal detections of the Rey Fire, north of Santa Barbara, CA, during 14-21 August, 2016; (c) Historical Fires layer showing wildfires in San Diego County during 1970-2015; and (d) Surface Fuels Layer showing the vegetation fuels for the canyons northwest of Los Angeles in 2014.

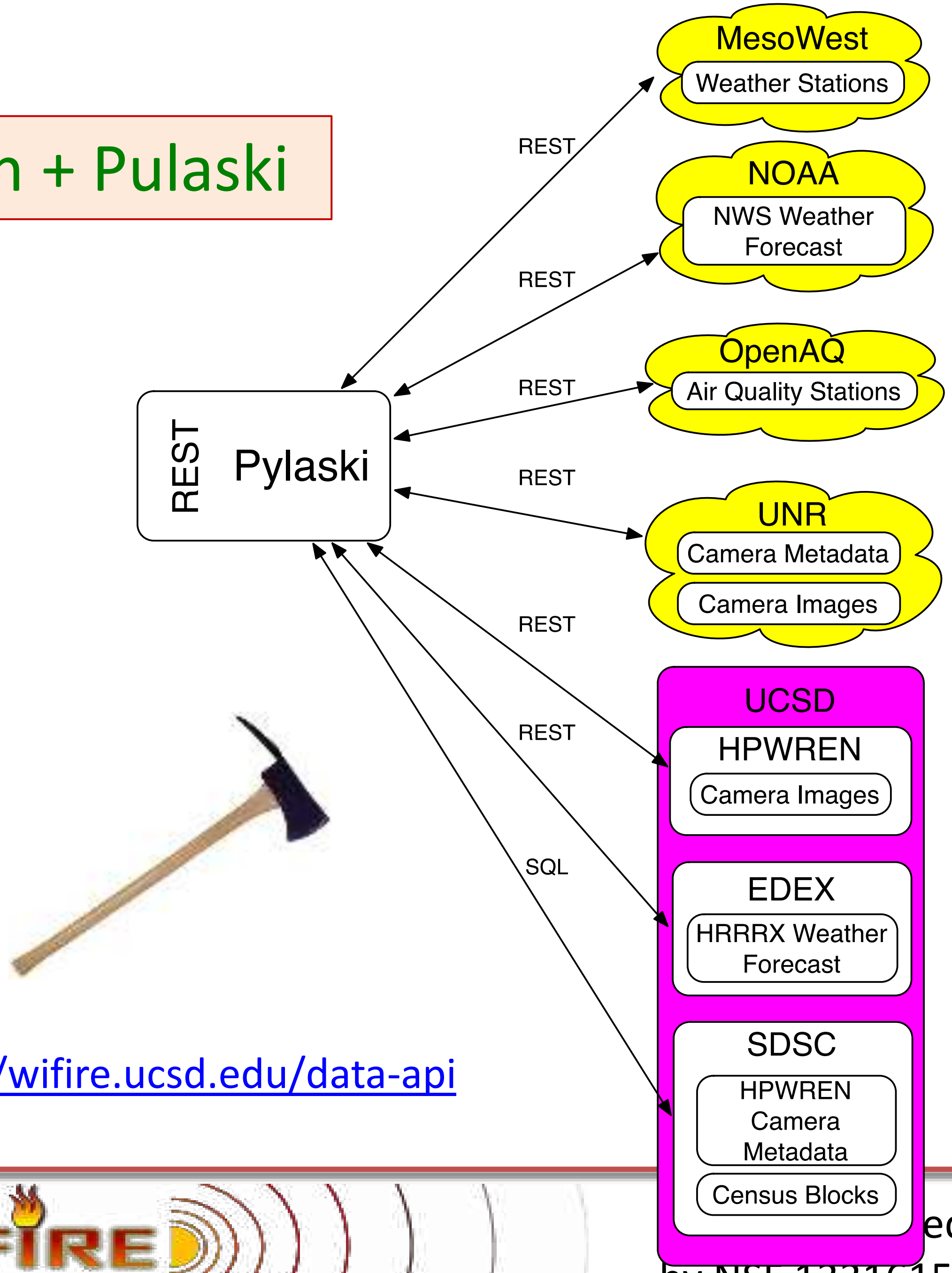


WIFIRE is funded by NSF 1331615

Pylaski

Pylaski = Python + Pulaski

- Uniform access to different data sources and types
  - e.g., OpenAQ (http://openaq.org) provides particulate matter, NO_2, SO_2, ozone, etc.
- Spatial and temporal queries
  - Closest to or within bounding box
  - Latest or within timespan
- Results are GeoJSON

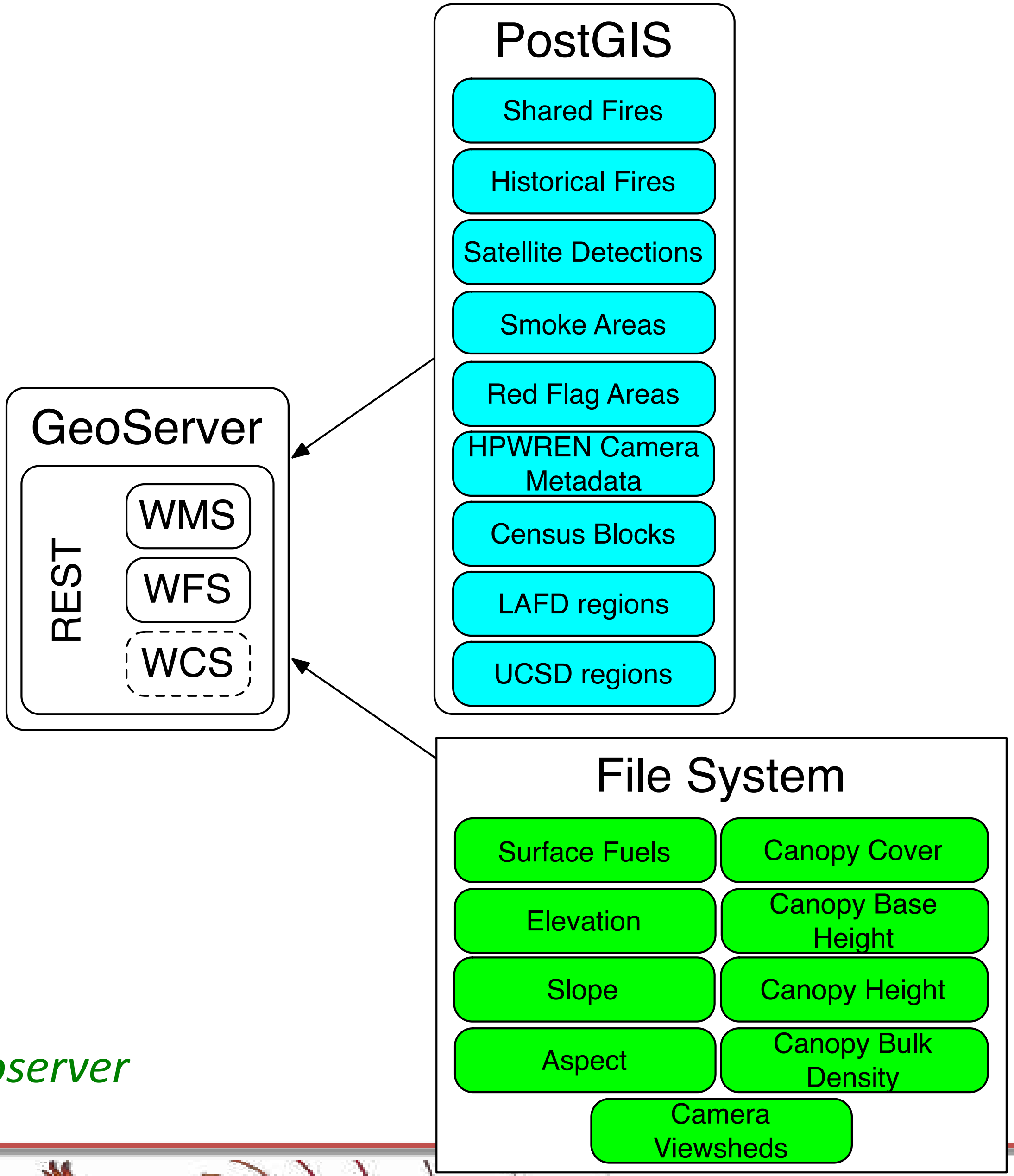


https://wifire.ucsd.edu/data-api

GeoServer

- REST access to GIS datasets
- WMS for image layers
  - Styling specifies color, opacity, etc.
- WFS for GeoJSON
- WCS for array data

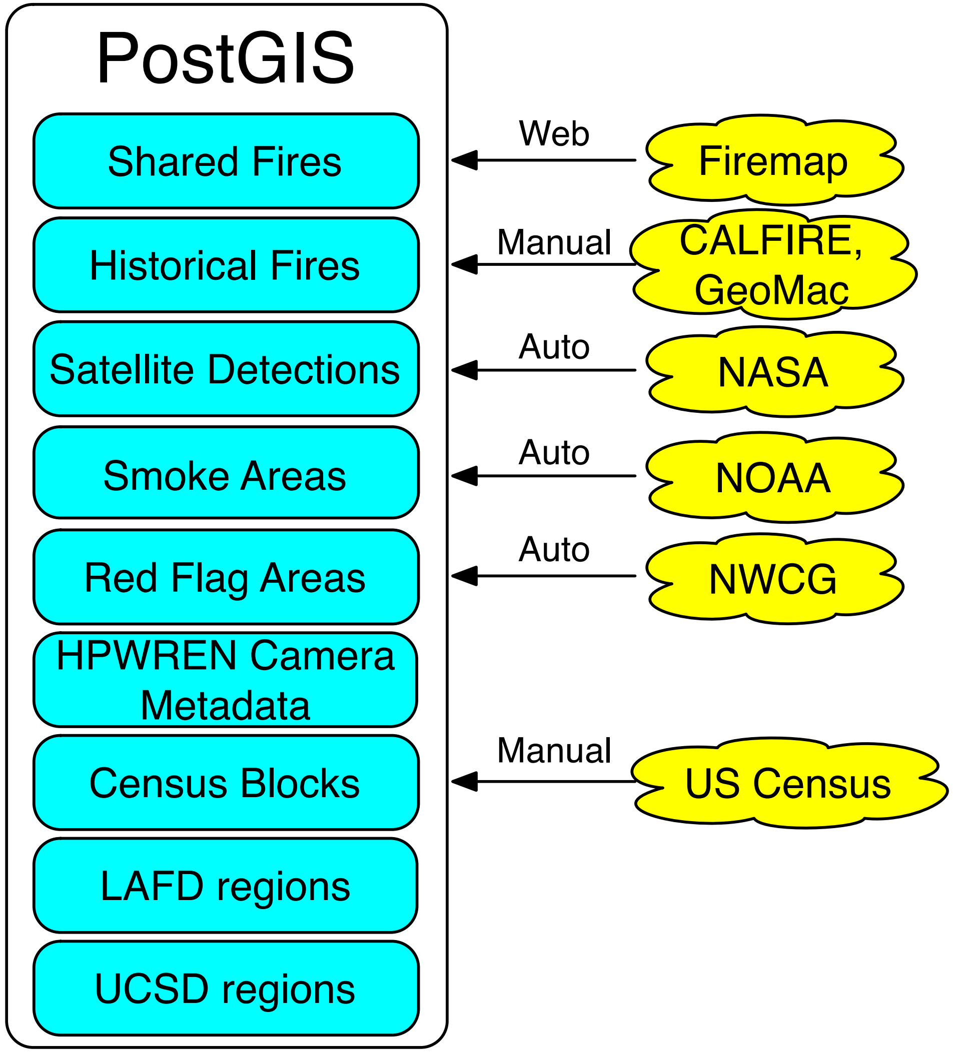
https://firemap.sdsc.edu:8443/geoserver



WIFIRE is funded by NSF 1331615

PostGIS

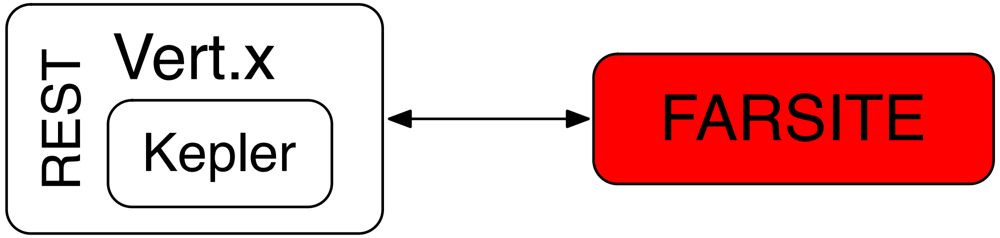
- Storage for vector GIS data
- Several datasets ingested every hour



WIFIRE is funded by NSF 1331615

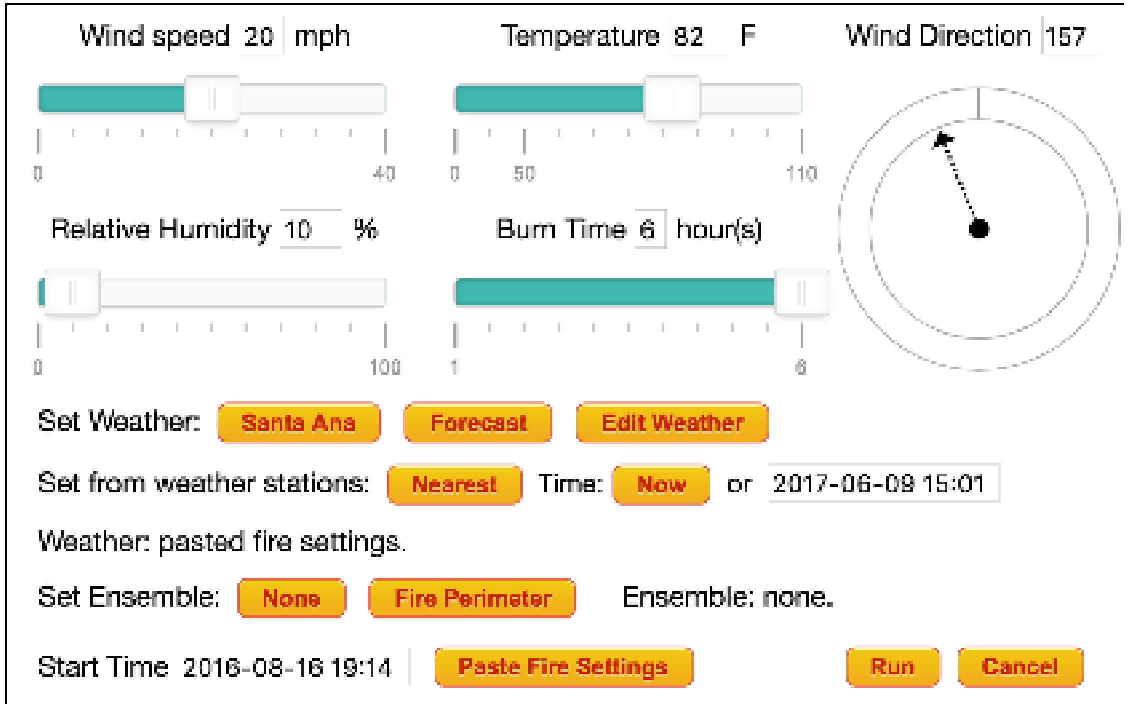
Kepler WebView

- REST service to run Kepler workflows
- FARSITE workflow
  - Inputs: ignition, weather, wind, landscape
  - Outputs: perimeters

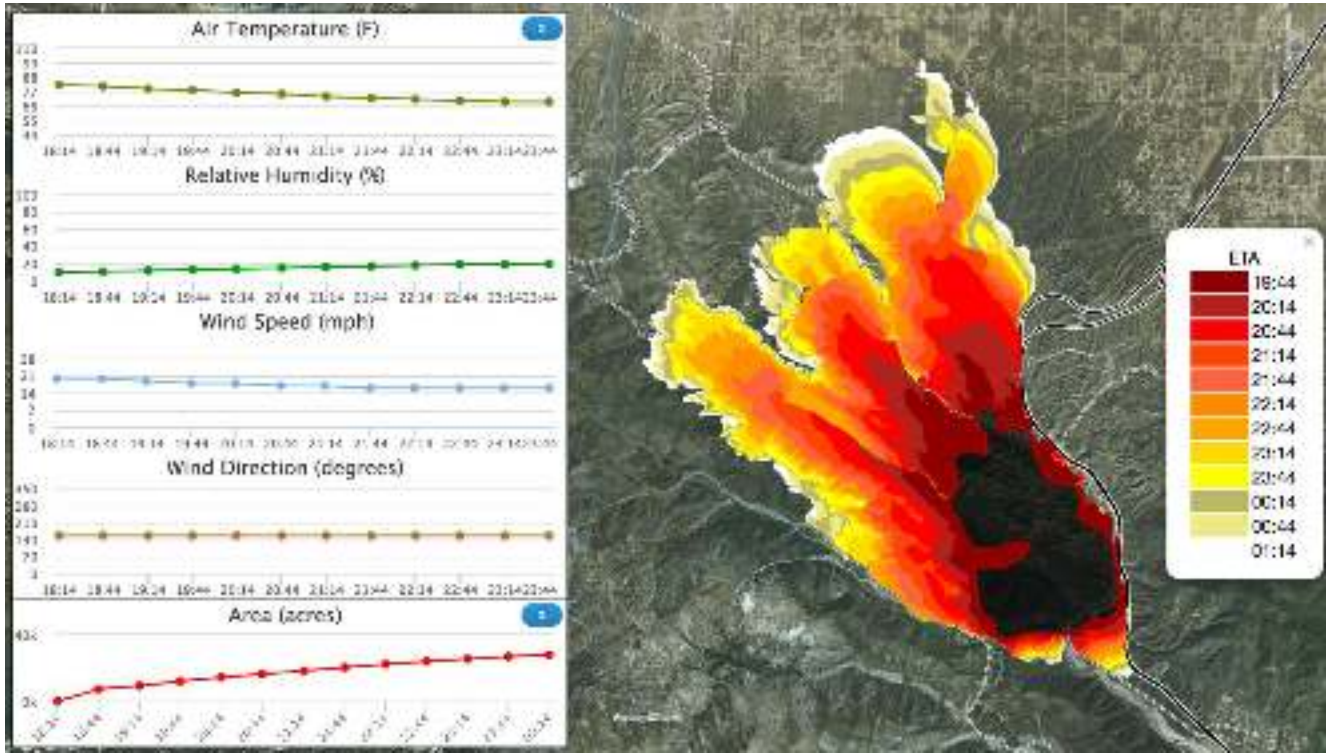


WIFIRE is funded by NSF 1331615

Firemap Predictive Modeling



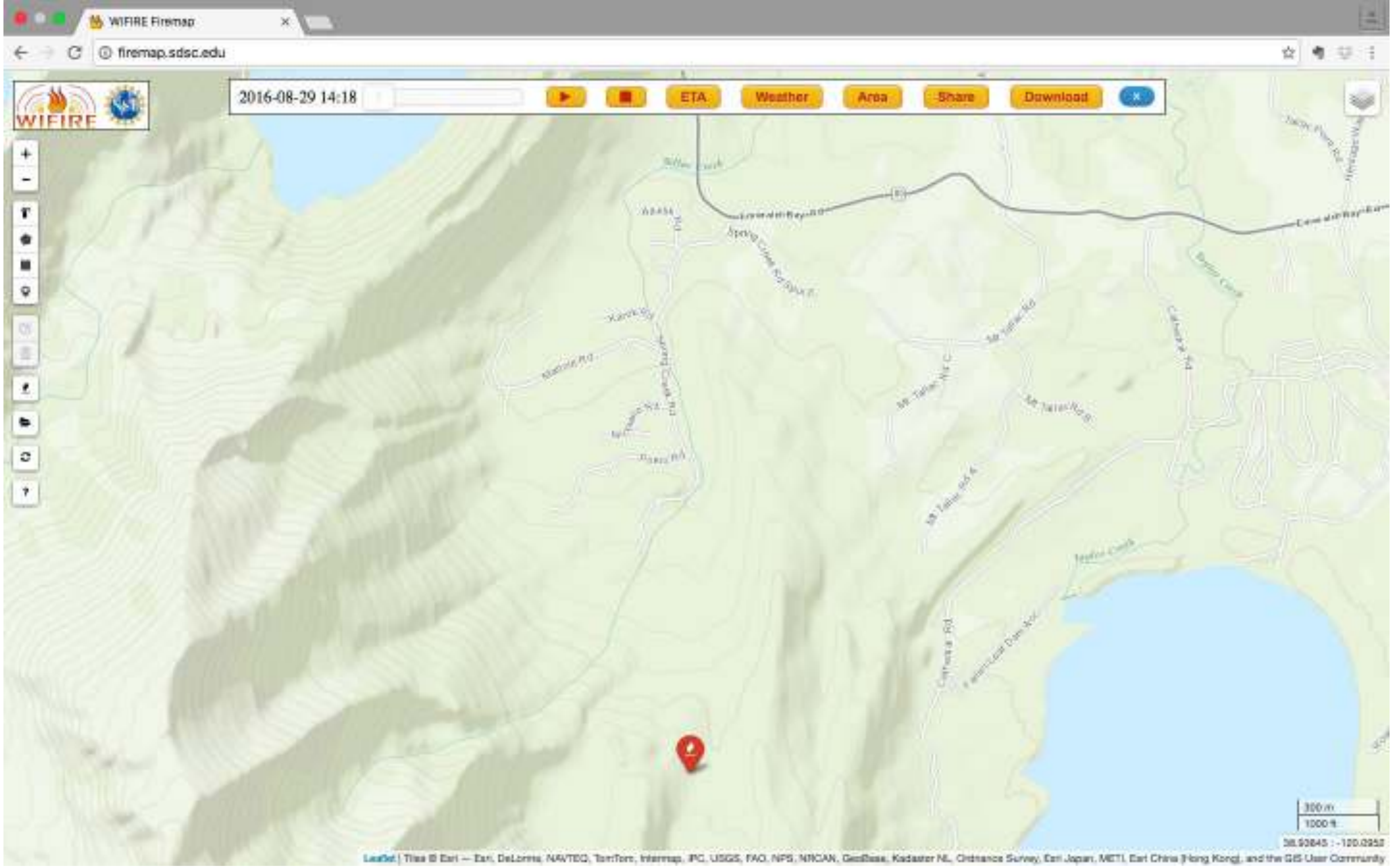
(a)



(b)

Figure 3: Firemap predictive modeling: (a) the configuration dialog specifying the simulation time, and weather conditions; (b) predicted fire spread for the Blue Cut Fire showing the weather conditions, fire perimeters, area, and arrival times.

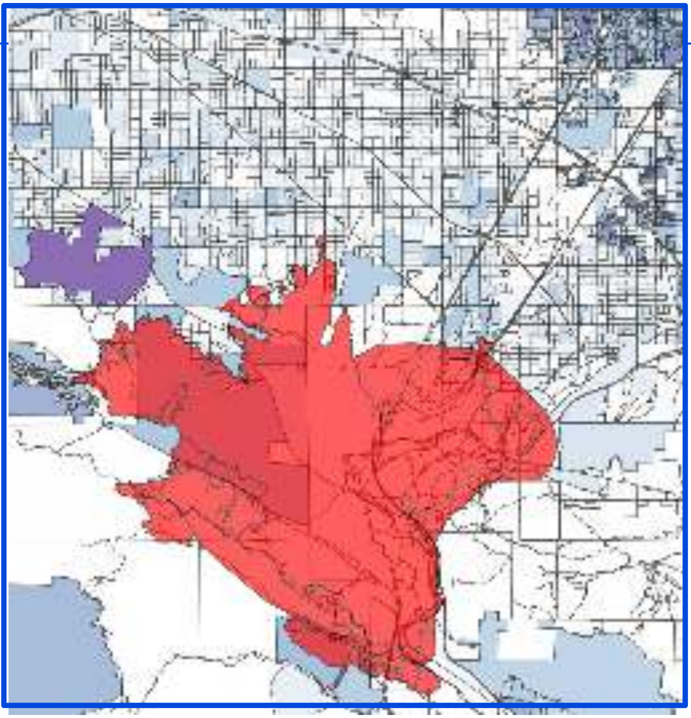
WIFIRE is funded by NSF 1331615



WIFIRE is funded by NSF 1331615

Some Related Experimental Features

- 2016 Blue Cut Fire in San Bernardino County
- 105 homes and 213 other buildings destroyed*
- Population estimates from US Census Bureau

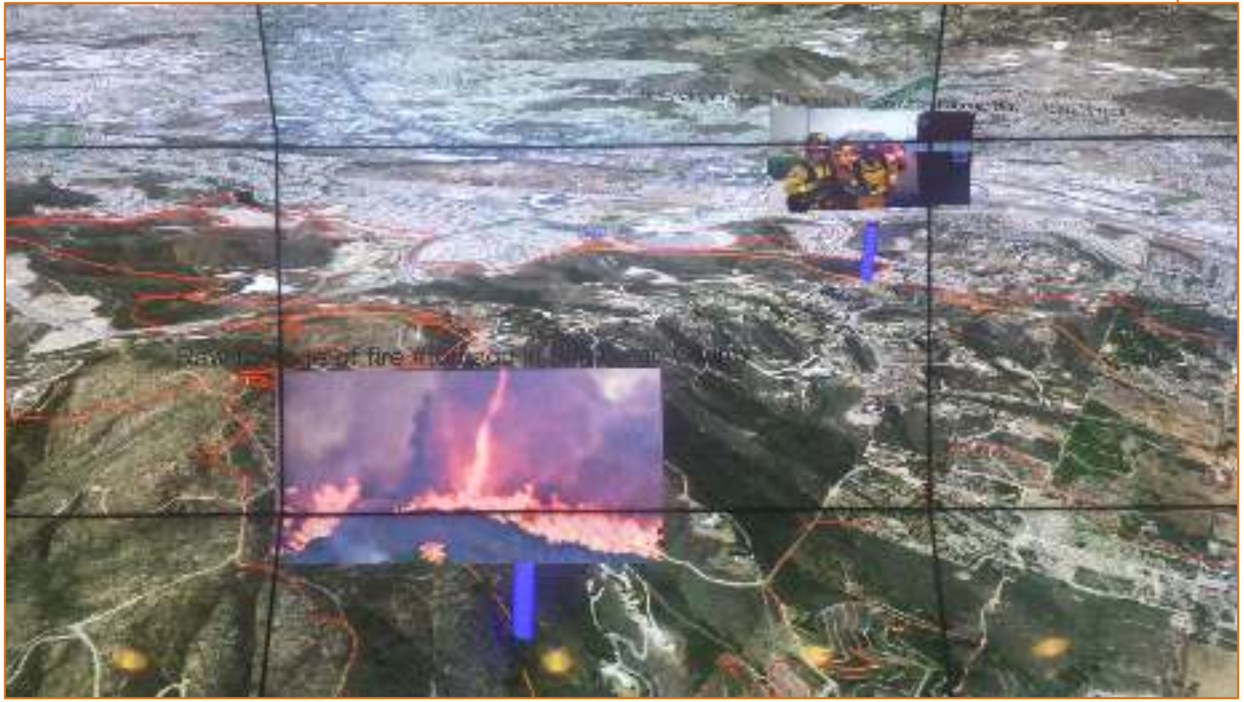


Population Density Near Wildfires

*http://www.latimes.com/local/lanow/la-me-ln-blue-cut-fire-20160819-snap-story.html

- Terrain imagery from 2009 and 2012
- Fire perimeters from the San Diego May 2014 fires
- Images were collected via Twitter
  - Captioned with tweet tag on top

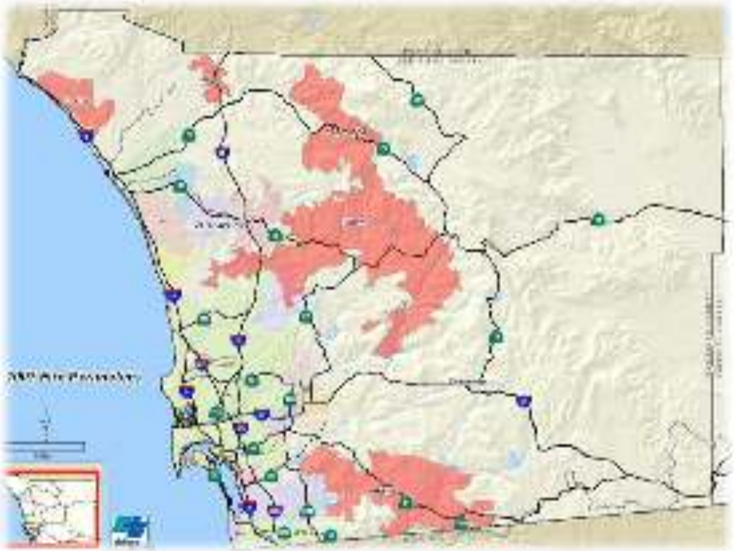
3D Terrain with Photos and Perimeters

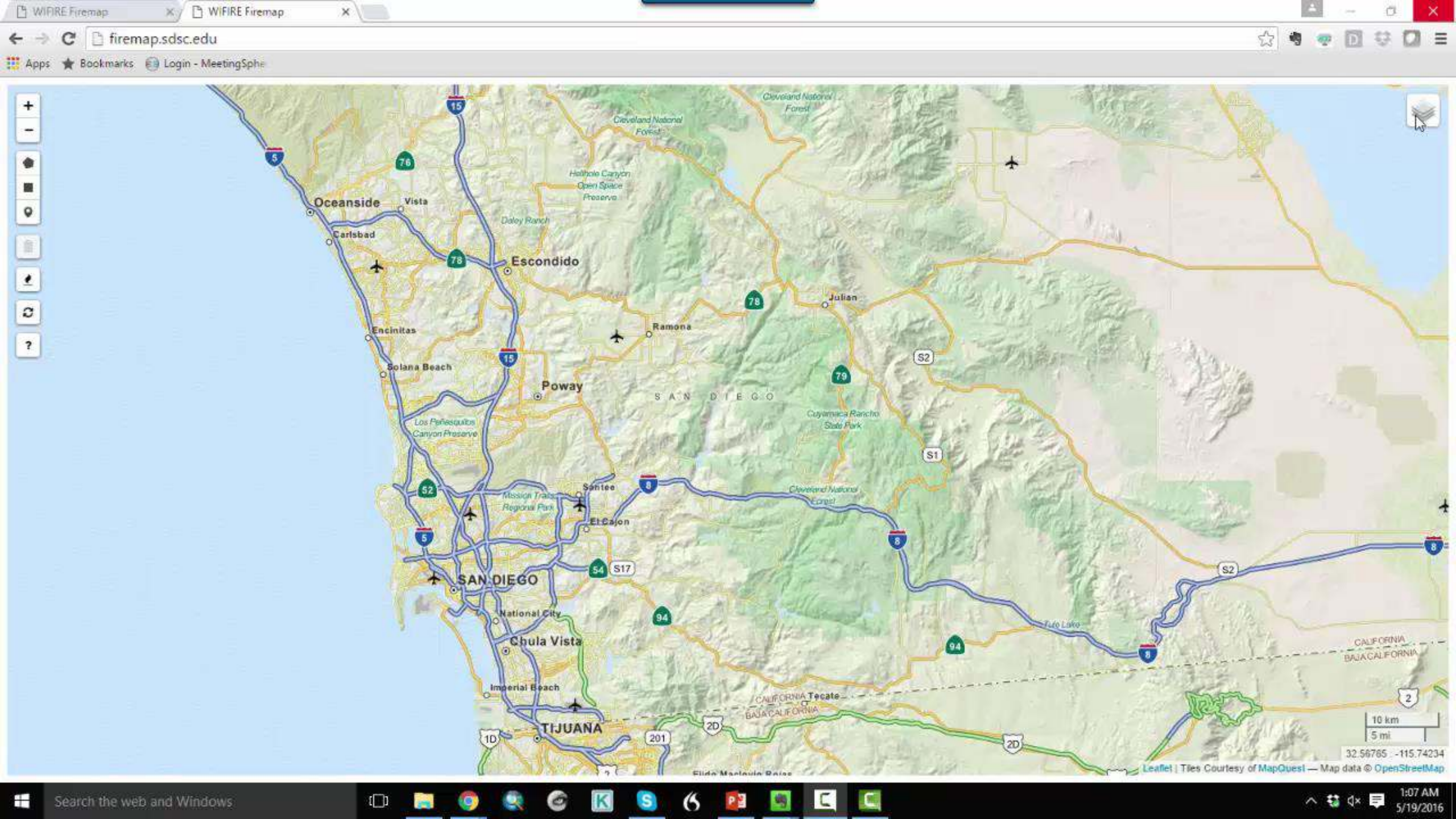


WIFIRE is funded by NSF 1331615

Crowd-Sourced Fire Perimeter Calculation

- Photos of fire locations from Twitter and Instagram
- Custom app allows recording additional data such as phone's location and orientation, focal length, etc.
- Improve camera location and orientation calculation through terrain map matching
- Image processing for fire/smoke localization within photos
- Goal: map of fire perimeter over time



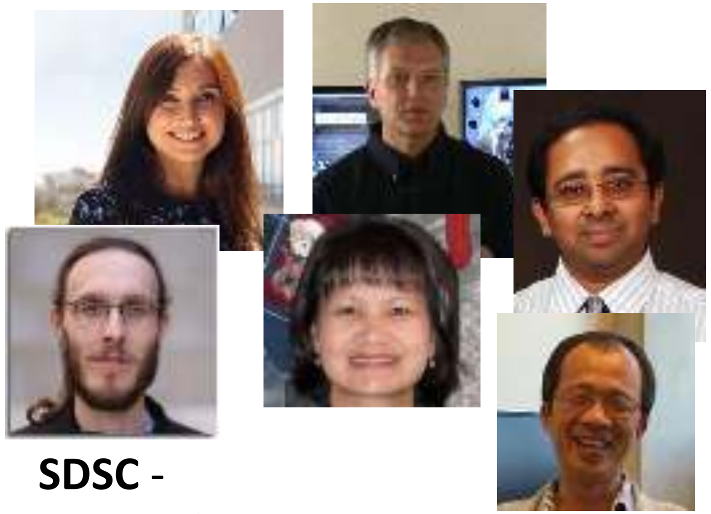

WiFIRE Firemap
WiFIRE Firemap
firemap.sdsc.edu
Apps
Bookmarks
Login - MeetingSphe
Oceanside
Vista
Carlsbad
Escondido
Encinitas
Solana Beach
Poway
Ramona
Julian
SAN DIEGO
Santee
El Cajon
SAN DIEGO
National City
Chula Vista
Imperial Beach
TIJUANA
Tecate
Cleveland National Forest
Hellhole Canyon Open Space Preserve
Daley Ranch
Los Peñasquitos Canyon Preserve
Mission Trails Regional Park
Cuyamaca Rancho State Park
CALIFORNIA
BAJA CALIFORNIA
10 km
5 mi
32.56765 , -115.74234
Leaflet | Tiles Courtesy of MapQuest — Map data © OpenStreetMap
Search the web and Windows
1:07 AM
5/19/2016

Some Lessons Learned

- User interface needs to be simple and not overloaded
- New fuel maps are vital to future success of fire modeling
- Data assimilation is needed to capture dynamics of fire
- Operationalization has production and support sustainability challenges unlike research sustainability
- On-demand large computing needs workflow automation
- Social media data is potentially useful, but not reliable

WIFIRE is funded by NSF 1331615

WIFIRE Team: It takes a village!



SDSC -
Cyberinfrastructure,
Workflows,
Data engineering,
Machine Learning,
Information
Visualization,
HPWREN

Calit2/QI-
Cyberinfrastructure, GIS,
Advanced Visualization,
Machine Learning,
Urban Sustainability,
HPWREN

UCSD MAE - Data assimilation

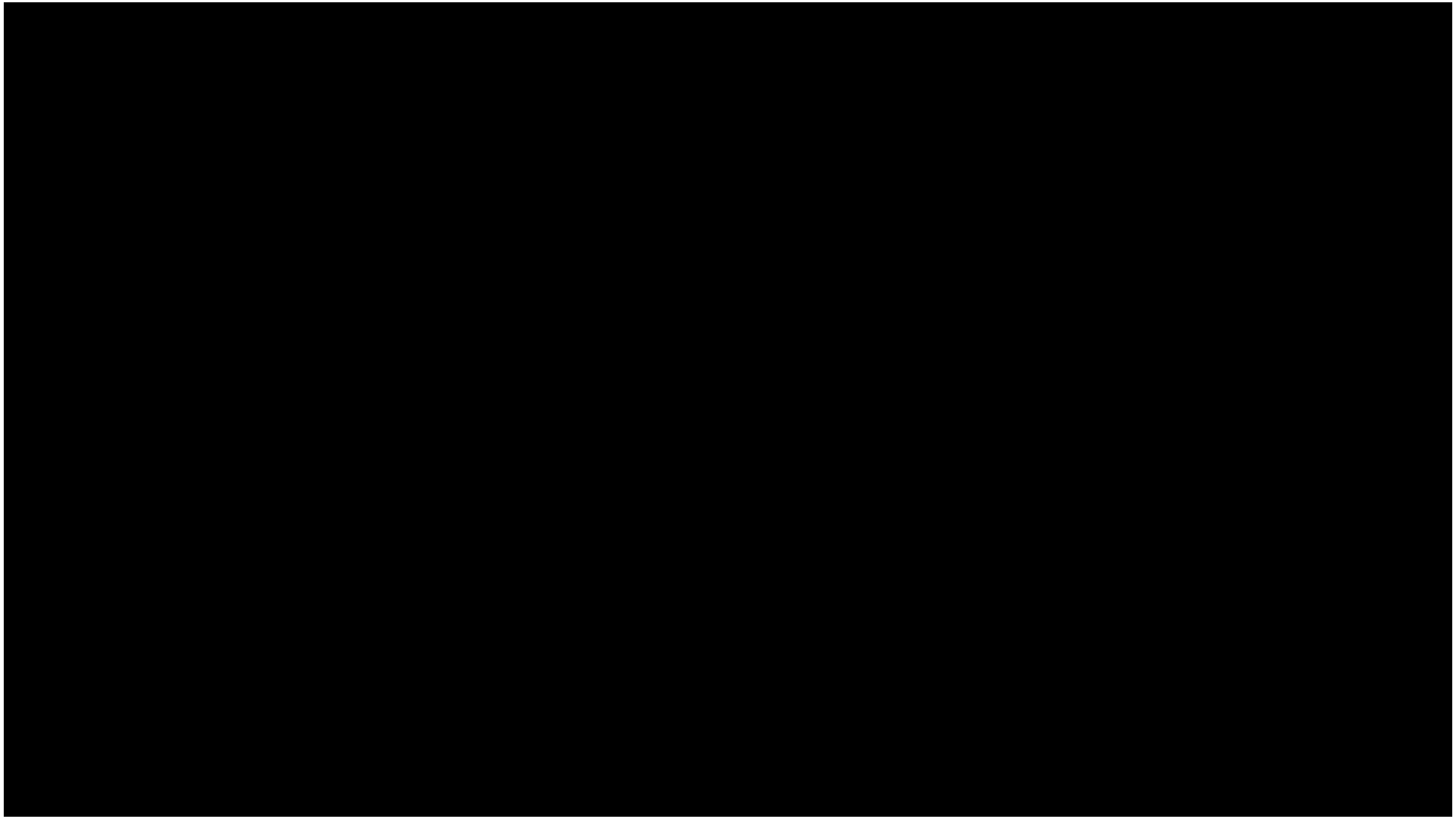
SIO - HPWREN

UMD - Fire modeling

- PhD level researchers
- Professional software developers
- 24 undergraduate students
  - UC San Diego
  - UC Merced
  - MURPA University
  - University of Queensland
- 1 high school student
- 4 MSc and 5 MAS students
- 2 PhD students (UMD)
- 1 postdoctoral researcher

http://wifire.ucsd.edu

WIFIRE is funded by NSF 1331615



WIFIRE is funded by NSF 1331615

Questions?

CONTACT

Ilkay Altintas, Ph.D.
Email: ialtintas@ucsd.edu

Parts of the presented work is funded by US NSF, US DOE, US NIH,
UC San Diego and various industry partners.

Social Science Data

Example Data

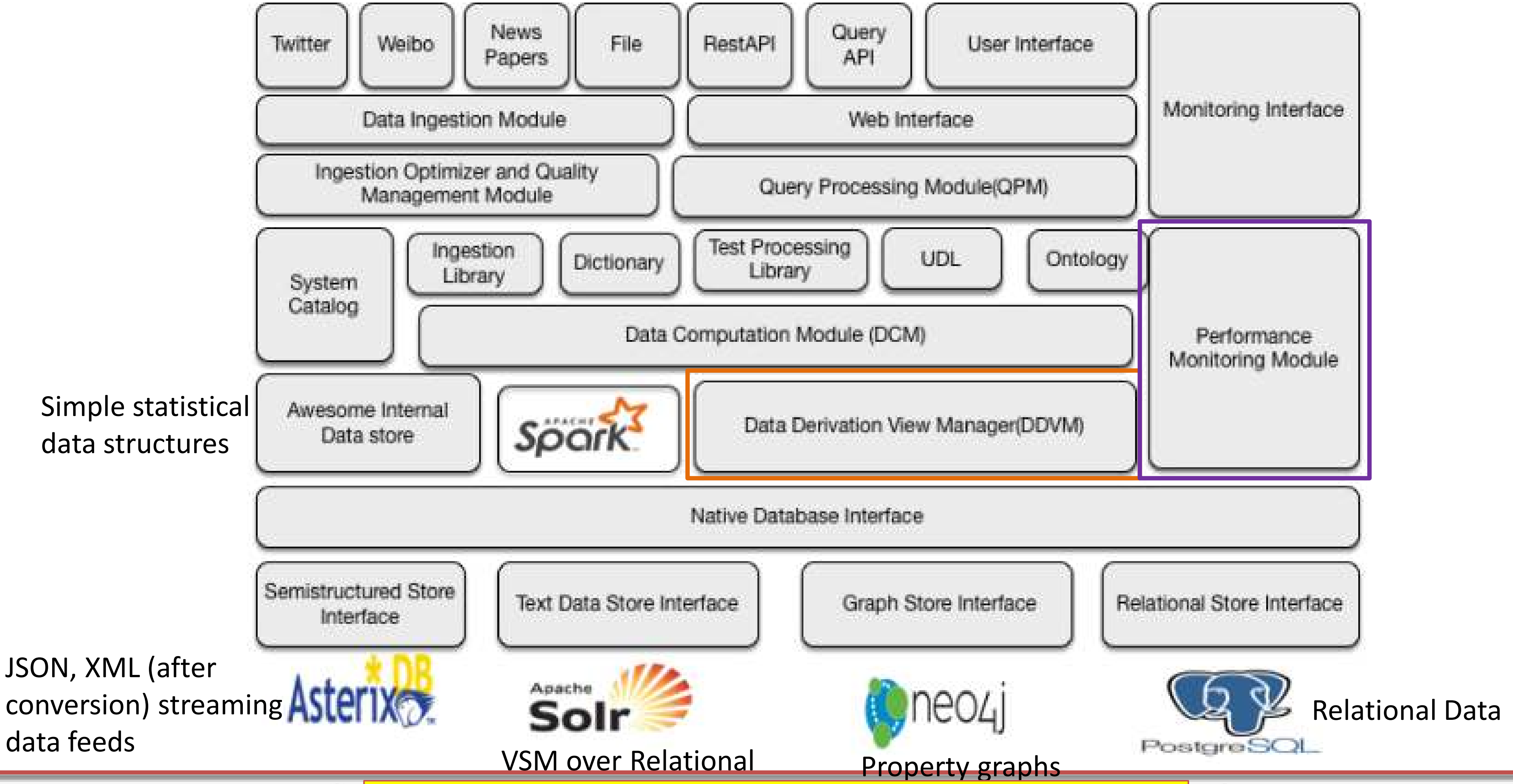
- Microblogs
- News articles/blogs/comments
- Political/geographic organization structures
- Biographies
- Social Media Data (e.g., Instagram)
- Message networks (e.g., emails)
- Legal documents
- Economic surveys
- Demographic data

Example Analyses

- Political Science: Can we **detect situations** where the Chinese news media is trying to suppress information about local protests while people on social media are actively discussing it? Can we detect them in near-real time?
- Sociology: Can we **construct a statistical model** to detect and predict political unrest by observing user behavior and actions on social media? Can we trace how political movements form by watching important players in a situation?
- Social Medicine: Can we **identify individuals and communities** who are at high-risk for HIV/AIDS? Can we use a combination of social media and transmission networks to make informed medical intervention (e.g., proactive vaccination)?
- Social Media Analytics: Can we **develop a prediction model** for identifying potential violence by watching social media? Can we identify and correlate disparate events that may lead to such violence?
- Computational Jurisprudence: Does the government **ever provide compensation** for cases against it? Are there **any networks** of judges and lawyers who never penalize the government?

AWESOME Polystore Architecture

(Analytical Workbench for Exploration of SOcial MEdia)



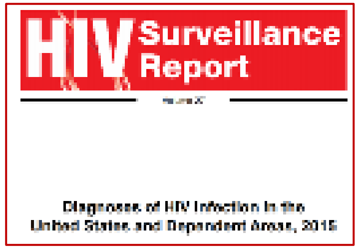
All data objects are temporal and may have annotations

WIFIRE is funded by NSF 1331615

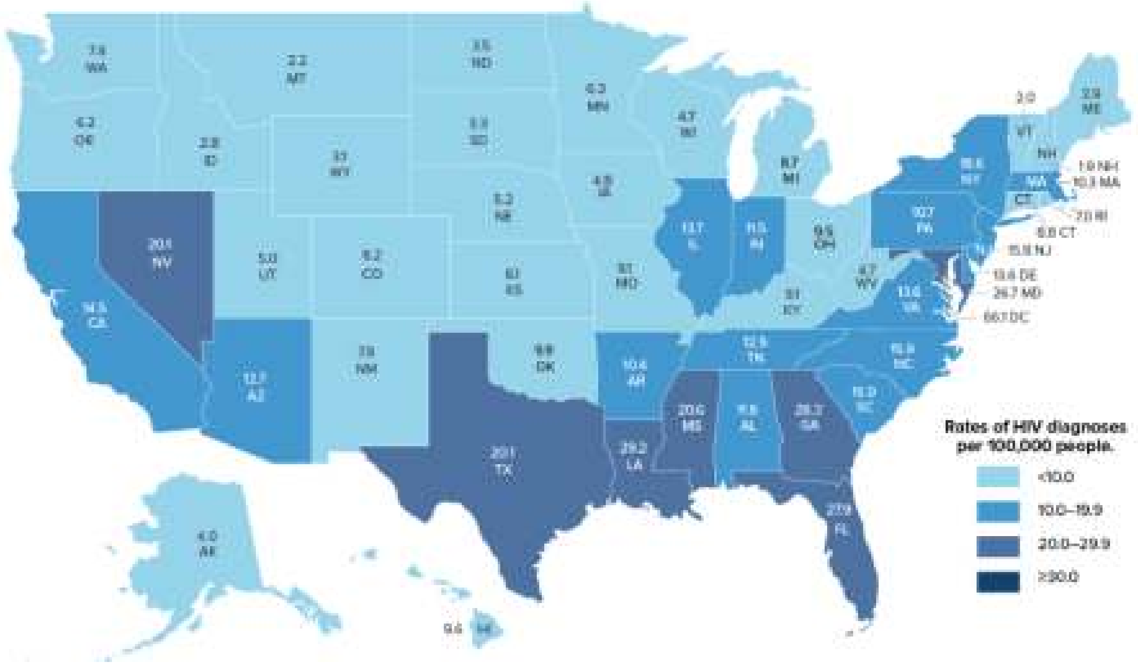
A Public Health Example (UCLA-led)

- CDC reports are usually delayed by 2+ years
- We need current, continual estimates of HIV counts and rates
  - Deploy resources according to urgency
    - Mobile clinics
  - Determine new relevant events
    - Burst in meth use activity
  - Determine efficacy of measures
    - Clean need exchange programs

Numbers and rates of diagnoses of HIV infection during 2010–2014 and preliminary numbers for 2015 are based on data from all 50 states, the District of Columbia, and 6 U.S. dependent areas (American Samoa, Guam, the Northern Mariana Islands, Puerto Rico, the Republic of Palau, and the U.S. Virgin Islands).



Rates of HIV Diagnoses Among Adults and Adolescents in the US in 2015, by State

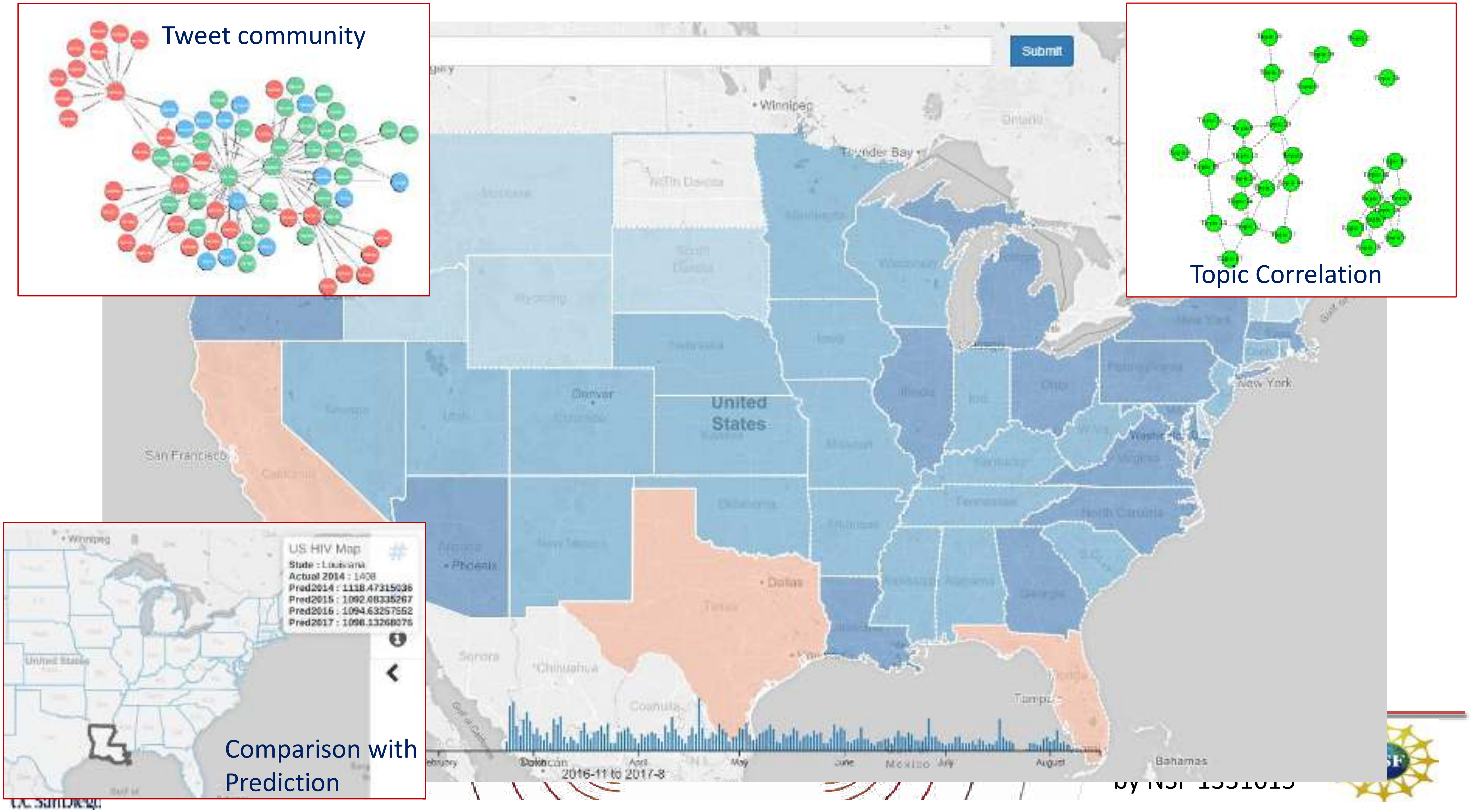


Can we use Social Media more timely estimates?

WIFIRE is funded by NSF 1331615

Data Sources and Derived Data

- Current sources
  - Twitter – real-time JSON data (seed keywords from HIV experts)
    - Has geocodes, text content and topic markers (hashtags)
  - Prior statistics from CDC – mostly tables
  - Geocoded locations of interest
- Derived data
  - What are the user communities and what are they talking about?
    - Derive communities (graphs of closely communicating users) and perform theme extraction with topic model like techniques
  - Is there a shift in topics like "social stigma" after a policy has been implemented?
    - Perform time-series analysis based on drifts in topics created from the above


Tweet community
Topic Correlation
Submit
Winnipeg
Thunder Bay
San Francisco
Denver
United States
Phoenix
Dallas
New York
Washington
Tampa
Bahamas
2016-11 to 2017-8
US HIV Map
State : Louisiana
Actual 2014 : 1408
Pred2014 : 1118.47315036
Pred2015 : 1092.08335267
Pred2016 : 1094.63257552
Pred2017 : 1098.13268076
Comparison with Prediction

Please direct all questions on AWESOME to:

Amarnath Gupta (a1gupta@ucsd.edu)