Time-Varying Semantic Representations of Planetary Observations for Discovering Novelties

Srija Chakraborty

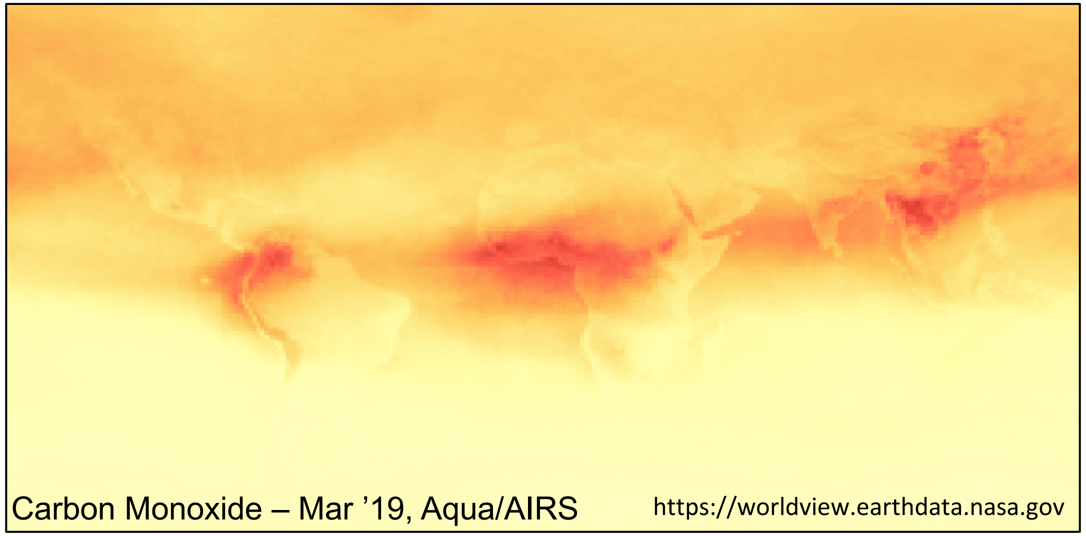
Goddard Space Flight Center

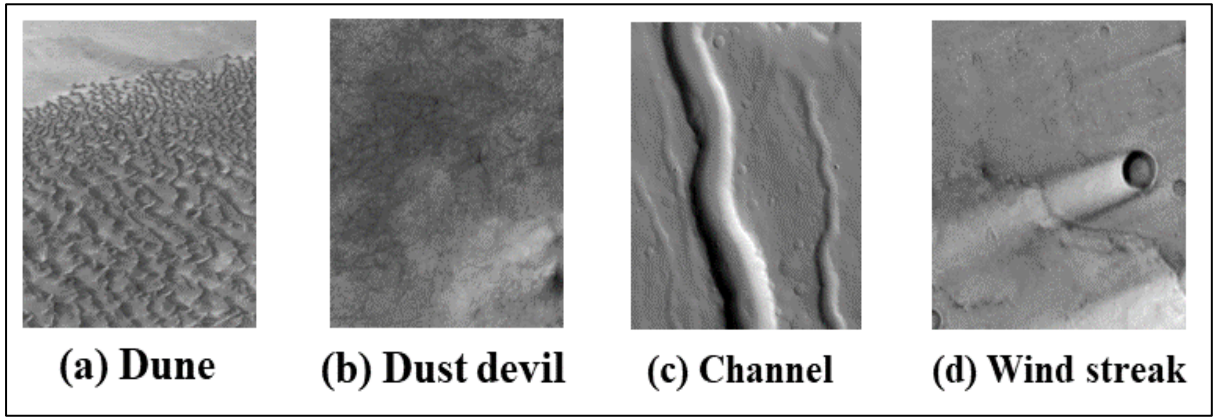
*Arizona State University

Novelty Detection from Remote Sensing Observations

Increasing volume of the data collected by the growing number of Earth and Planetary observation satellites

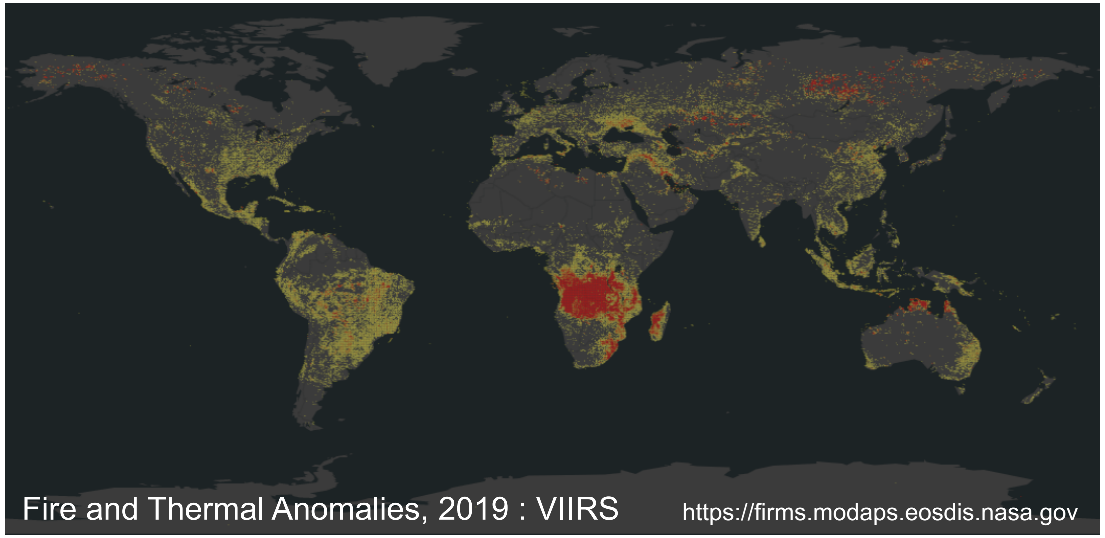
Classes of Interest

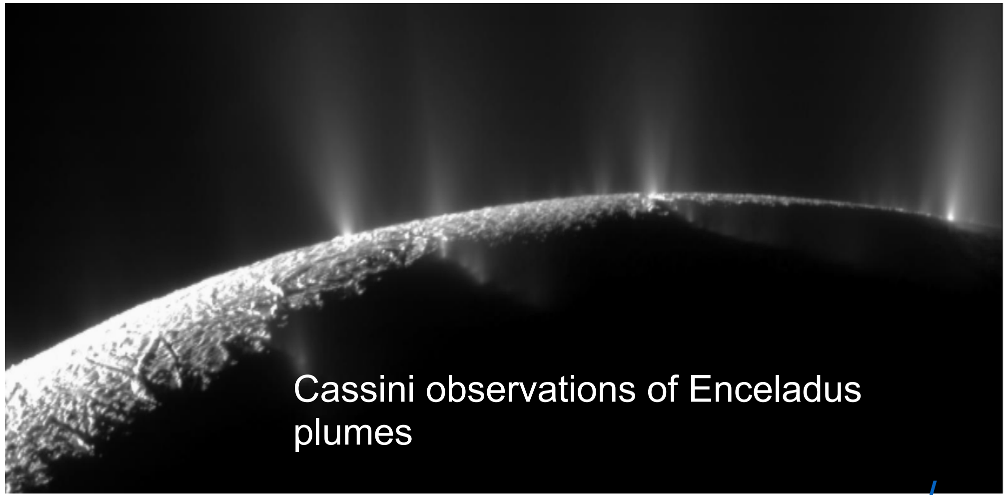

Carbon Monoxide – Mar ’19, Aqua/AIRS https://worldview.earthdata.nasa.gov



http://themis.asu.edu

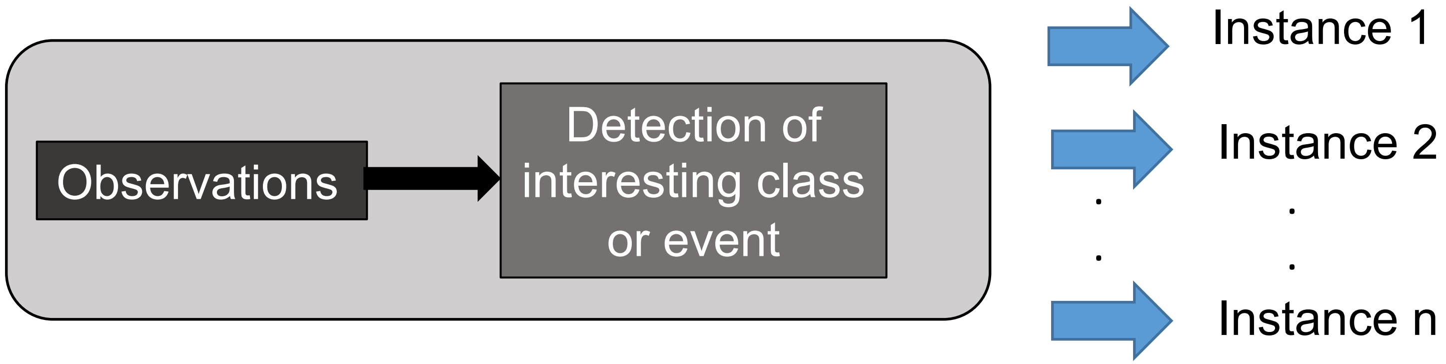
Events


Fire and Thermal Anomalies, 2019 : VIIRS https://firms.modaps.eosdis.nasa.gov

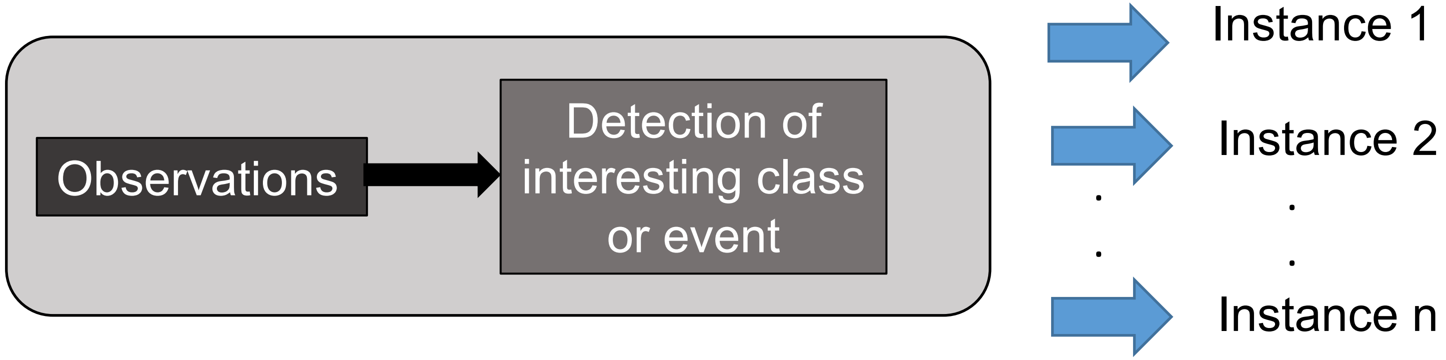

Cassini observations of Enceladus plumes

https://solarsystem.nasa.gov/news/13020/the-moon-with-the-plume/

Expert-like Analysis of Novelty



Expert-like Analysis of Novelty

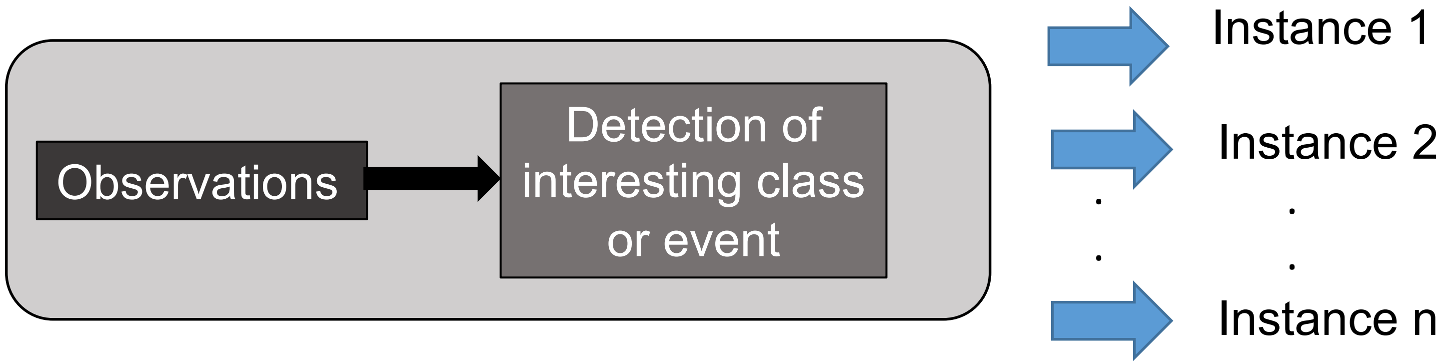


Are all detections equally interesting to an expert?

Unusual context of an event

- Spatially associated with other classes; previously unseen
- Observed it in an unusual location or season
- Increases the need for an expert to further analyze the observation

Expert-like Analysis of Novelty



Are all detections equally interesting to an expert?

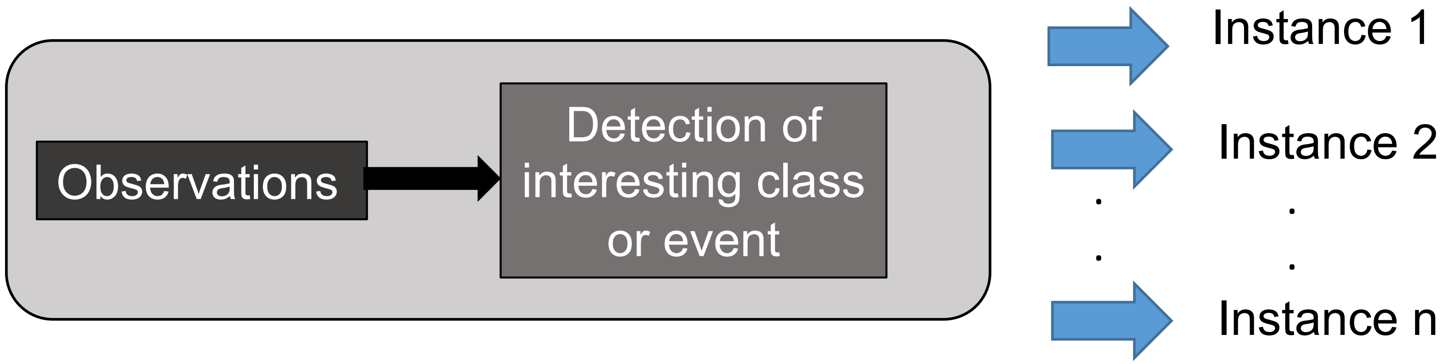
Do these detections change/ improve our understanding of the event/ class?

Unusual context of an event

- Spatially associated with other classes; previously unseen
- Observed it in an unusual location or season
- Increases the need for an expert to further analyze the observation

Refines our knowledge of event/ class
 - better model of where and when to expect a natural hazard or a landform
- Better understanding of driving factors or triggers of the hazard or how these morphologies formed

Expert-like Analysis of Novelty



Are all detections equally interesting to an expert?

Unusual context of an event

- Spatially associated with other classes; previously unseen
- Observed it in an unusual location or season
- Increases the need for an expert to further analyze the observation

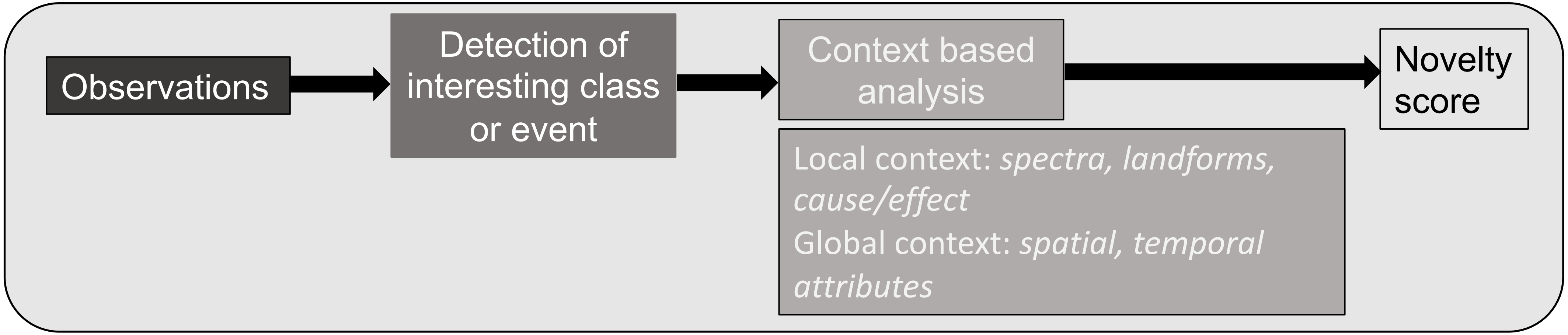
Do these detections change/ improve our understanding of the event/ class?

Refines our knowledge of event/ class

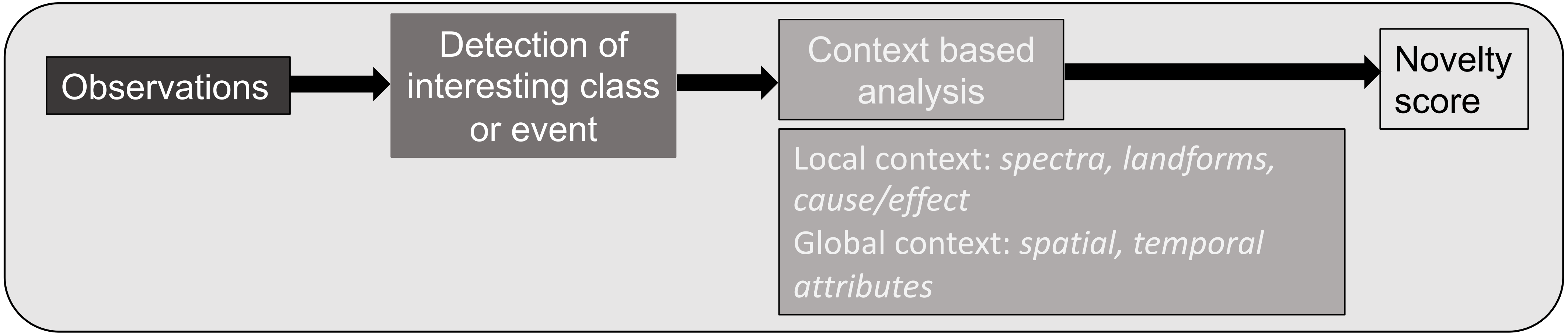
- better model of where and when to expect a natural hazard or a landform
- Better understanding of driving factors or triggers of the hazard or how these morphologies formed

Ranking anomalies to improve our ability to interpret extremes, targeting future observations, support scientific discovery and planetary exploration

Context-based Representations of Events and Classes of Interest



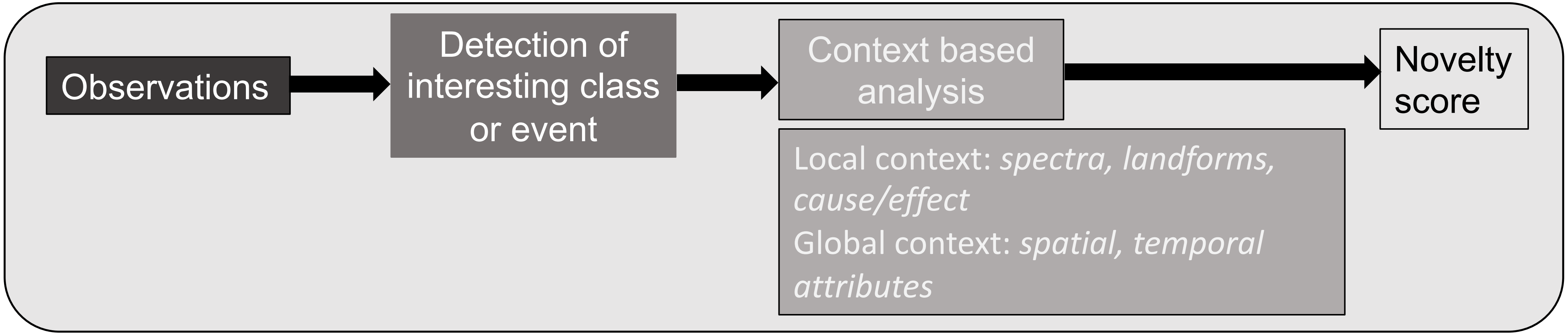
Context-based Representations of Events and Classes of Interest



Towards expert-like interpretation of observations:

- Representations that learn local and global context
- Update this representation to indicate the shift in the understanding of the process through discovered anomalies
- Explored on multispectral Mars (THEMIS) VIS/IR images and currently on (MODIS, VIIRS) wildfire

Context-based Representations of Events and Classes of Interest



Towards expert-like interpretation of observations:

- Representations that learn local and global context
- Update this representation to indicate the shift in the understanding of the process through discovered anomalies
- Explored on multispectral Mars (THEMIS) VIS/IR images and currently on (MODIS, VIIRS) wildfire

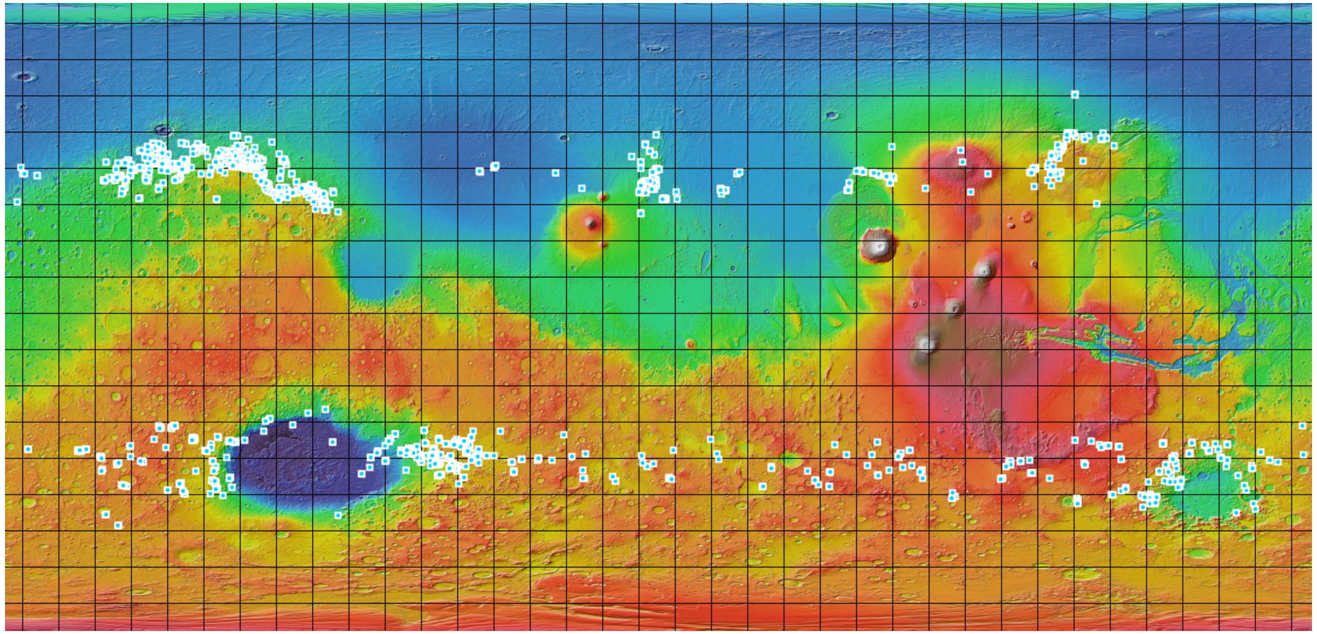
$\mathrm{O} = \{\mathrm{o}_1, \mathrm{o}_2, \ldots, \mathrm{o}_\mathrm{N}\}$ Data Repository of all N past observations

$D = \{o_j | C_j \in C, D \subseteq O\}$ Total observations N_C with C classes of interest

Representations: o_j: $< \boldsymbol{detected\ classes, (lat, long), seasons, surface\ properties} >$

Identify Spatial Locations

Based on our past observations, where do we expect to find the given landform?



Spatial Location of glacier-like forms

R_i: < expected interclass association, **expected location**, expected season, surface properties >

For every instance θ_i in D belonging to class i:

Expected Spatial locations: $\Omega_i = \{(X_{\omega_i} + \sigma_{lat_{\omega_i}}, Y_{\omega_i} + \sigma_{long_{\omega_i}})\}$,
where $(X_{\omega_i}, Y_{\omega_i})$ are centers of ω clusters (by grouping observations within a spatial bound)
$\sigma_{lat_{\omega_i}}, \sigma_{long_{\omega_i}}$ *after assigning all* θ_i of class i to the closest center

Souness, C. et al., 2012. An inventory and population-scale analysis of martian glacier-like forms. Icarus 217, 243–255. http://dx.doi.org/10.1016/j.icarus.2011.10.020.

Learning Most Prevalent Season

Based on our past observations, when do we expect to observe the given landform?

Mar year[a]	Spring equinox ($L_s = 0°$)[b]	Summer solstice ($L_s = 90°$)[b]	Autumnal equinox ($L_s = 180°$)[b]	Winter solstice ($L_s = 270°$)[b]
25	05-31-2000	12-16-2000	06-17-2001	11-11-2001
26	04-18-2002	11-03-2002	05-05-2003	09-29-2003
27	03-05-2004	09-20-2004	03-22-2005	08-16-2005
28	01-21-2006	08-08-2006	02-07-2007	07-04-2007
29	12-09-2007	06-25-2008	12-25-2008	05-21-2009
30	10-26-2009	05-13-2010	11-12-2010	04-08-2011

R_i: < expected interclass association, expected location, **expected season**, surface properties >

For every instance θ_i in D belonging to class i:

Expected season of prevalence, μ_i at each spatial cluster:
probability mass function from binary seasonal vector $\mathbf{t}$, $|\mathbf{t}| = \Phi$ (number of seasons on the surface)

$$t_{\theta_i}[k] = \begin{cases} 1, & k = \varphi, \\ 0, & \text{otherwise} \end{cases}$$

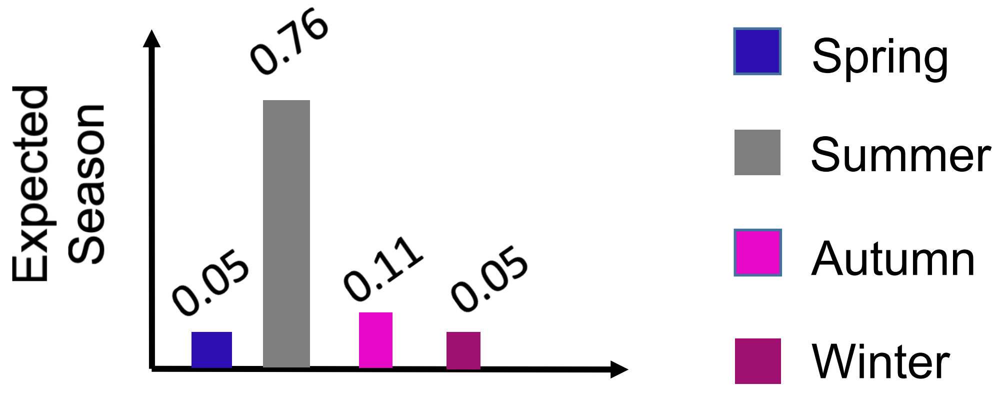
$$T_i = E[k_i] = \sum_{k=1}^{\phi} k.\, f(k)_i, \quad \sigma(k_i) = \sum_{k=1}^{\phi} (k - E[k_i])^2.\, f(k)_i, \qquad \text{where } f(k)_i = \frac{1}{N_{c_i}} \sum_{\theta_i \in i} t_{\theta_i}[k]$$

Cantor, B.A. et al., 2006. Mars Orbiter Camera observations of Martian dust devils and their tracks (September 1997 to January 2006) and evaluation of theoretical vortex models, Journal of Geophysical Research, vol. 111(E12)

Learning Most Prevalent Season

Based on our past observations, when do we expect to observe the given landform?

Mar year[a]	Spring equinox ($L_s = 0°$)[b]	Summer solstice ($L_s = 90°$)[b]	Autumnal equinox ($L_s = 180°$)[b]	Winter solstice ($L_s = 270°$)[b]
25	05-31-2000	12-16-2000	06-17-2001	11-11-2001
26	04-18-2002	11-03-2002	05-05-2003	09-29-2003
27	03-05-2004	09-20-2004	03-22-2005	08-16-2005
28	01-21-2006	08-08-2006	02-07-2007	07-04-2007
29	12-09-2007	06-25-2008	12-25-2008	05-21-2009
30	10-26-2009	05-13-2010	11-12-2010	04-08-2011



R_i: $<$ expected interclass association, expected location, **expected season**, surface properties $>$

For every instance θ_i in D belonging to class i:

Expected season of prevalence, μ_i at each spatial cluster:
probability mass function from binary seasonal vector $\mathbf{t}$, $|\mathbf{t}| = \Phi$ (number of seasons on the surface)

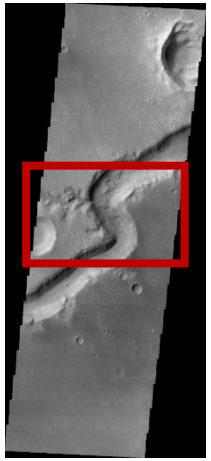
$$t_{\theta_i}[k] = \begin{cases} 1, & k = \varphi, \\ 0, & \text{otherwise} \end{cases}$$

$$T_i = E[k_i] = \sum_{k=1}^{\phi} k.\, f(k)_i, \quad \sigma(k_i) = \sum_{k=1}^{\phi} (k - E[k_i])^2.\, f(k)_i, \qquad \text{where } f(k)_i = \frac{1}{N_{c_i}} \sum_{\theta_i \in i} t_{\theta_i}[k]$$

Cantor, B.A. et al., 2006. Mars Orbiter Camera observations of Martian dust devils and their tracks (September 1997 to January 2006) and evaluation of theoretical vortex models, Journal of Geophysical Research, vol. 111(E12)

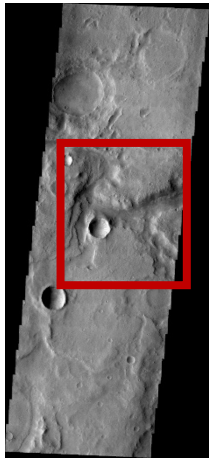
Learning Spatial Association

Based on our past observations, what other landforms co-occur with a given landform?

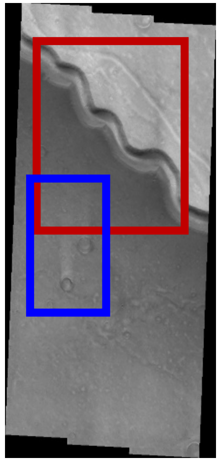


Channel

Channel



Channel



Channel,
Wind streak

$$R_i: < \textbf{expected interclass association}, \text{expected location}, \text{expected season}, \text{surface properties} >$$

For every instance θ_i in D belonging to class i:

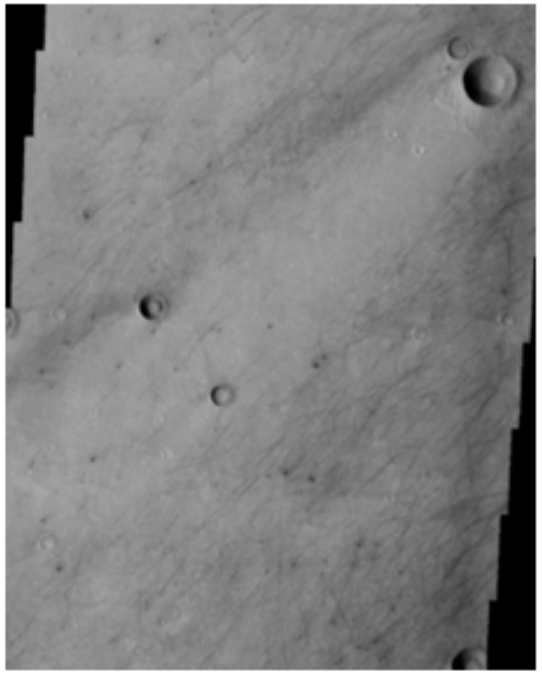
Interclass (i, j, …, n) association in each spatial cluster: $s_{i,j,\ldots,n} = \frac{|\theta_s|}{N_c}$,
where $|\theta_s|$ - frequency of co-occurrence of (i, j, …, n) in D

Refining Spatial Association through Expert Feedback

For each ILF observed association $(\mathrm{i}, \mathrm{j}, \ldots, \mathrm{n}) \in D$:

Expert feedback on how rare or important the association is

Scale: 1= (Low downlink priority/common, not interesting); 5= (High downlink priority / rare, very interesting)



class: dust devil, wind streak; lat: -15; long: 302 E; date: 2002/3/31; Mars year=25

	0	1	2	3	4	5
How important or rare is this association	○	○	○	○	○	◉

$w_{\theta_{i,e}}$

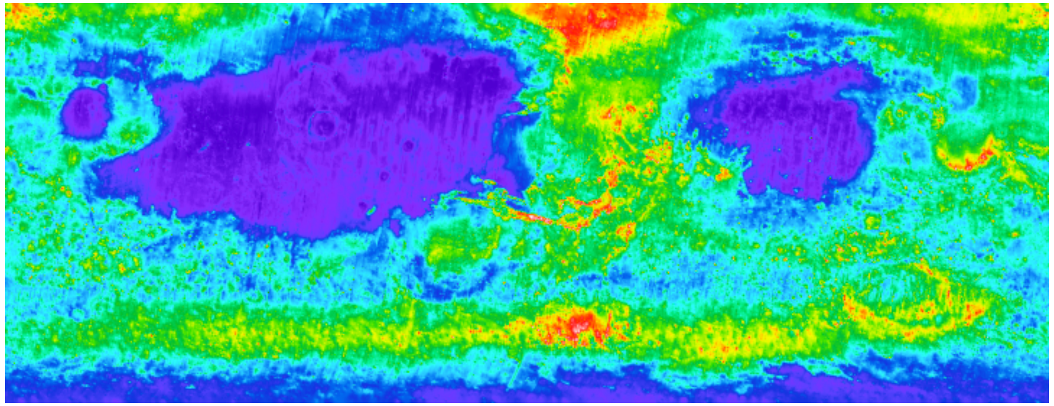
$$w_{i,j,..,n} = \frac{1}{E}\sum_{e=1}^{E} w_{\theta_{i,e}}$$

Weight or importance of the association from E experts

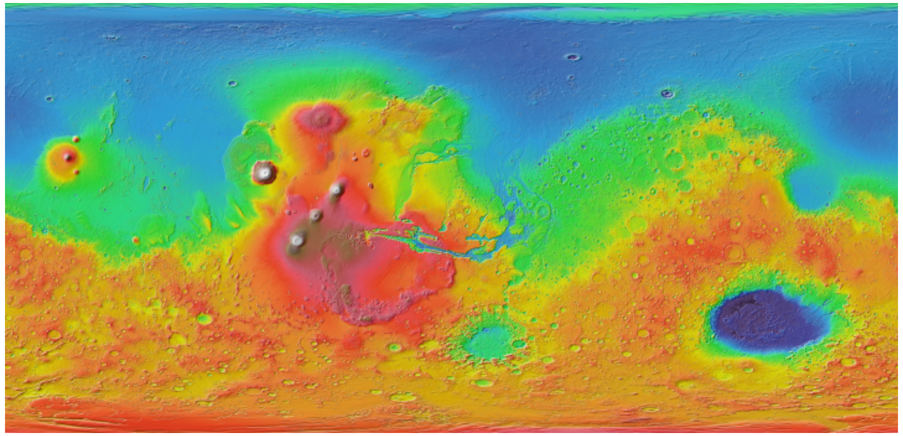
R_i: < **expected interclass association**, expected location, expected season, surface properties >

Learning Surface Properties

Based on our past observations, what can we infer about surface properties where a given landform appears?



TES Thermal inertia



MOLA Elevation

R_i: < expected interclass association, expected location, expected season, **surface properties** >

Joint representation $p_i \sim \mathcal{N}(\mu_i, \Sigma_i)$ of surface properties
Multivariate normal distribution for each class of interest at each spatial location

Provides local (terrain) context

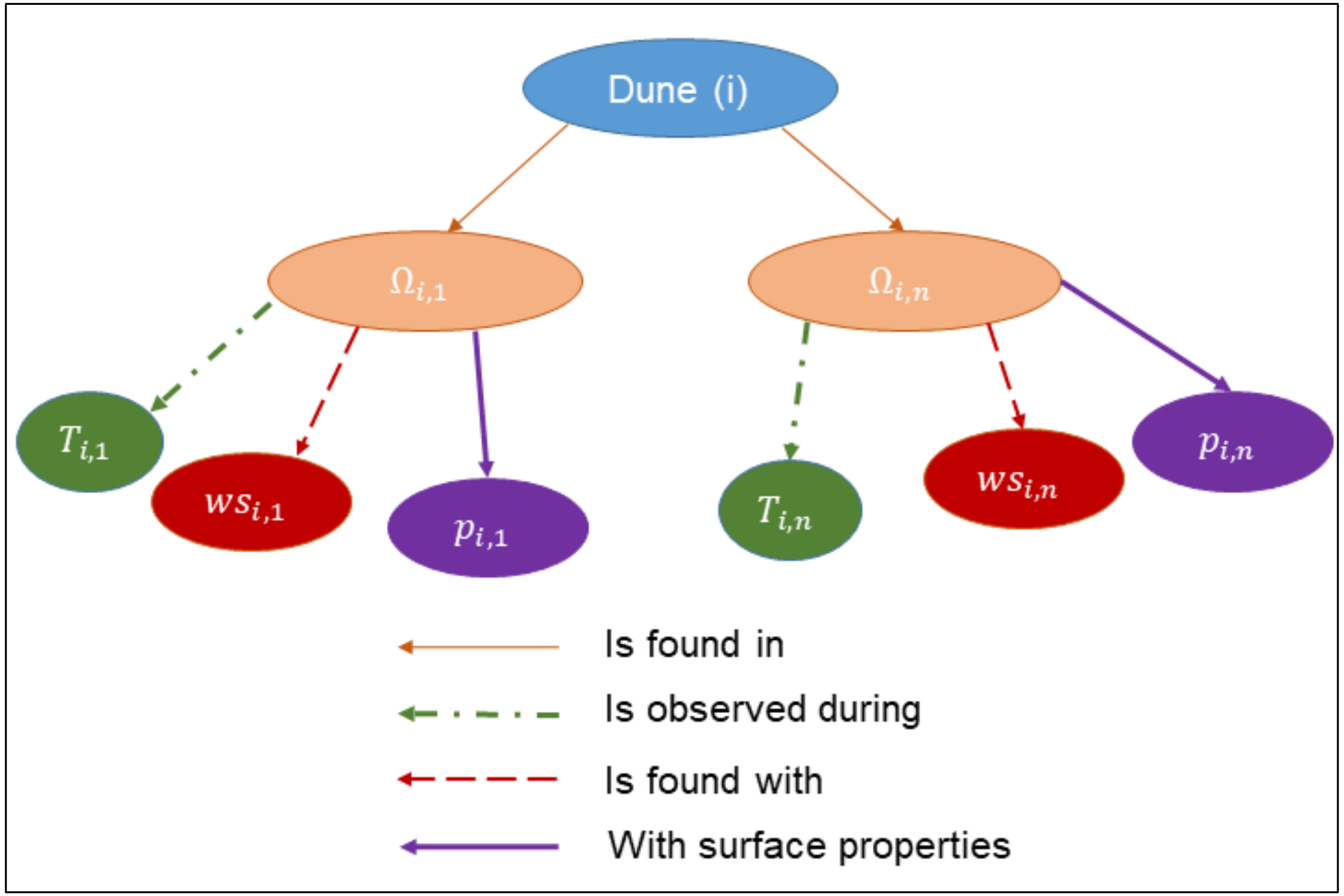
Integrating observations from multiple instruments

Christensen, P.R.; Engle, E.; Anwar, S.; Dickenshied, S.; Noss, D.; Gorelick, N.; Weiss-Malik, M.; *JMARS – A Planetary GIS*, http://adsabs.harvard.edu/abs/2009AGUFMIN22A..06C

Christensen, P. R., and H. J. Moore (1992), The martian surface layer, in Mars, edited by H. H. Kieffer, et al., pp. 686-729, University of Arizona Press, Tucson, AZ.

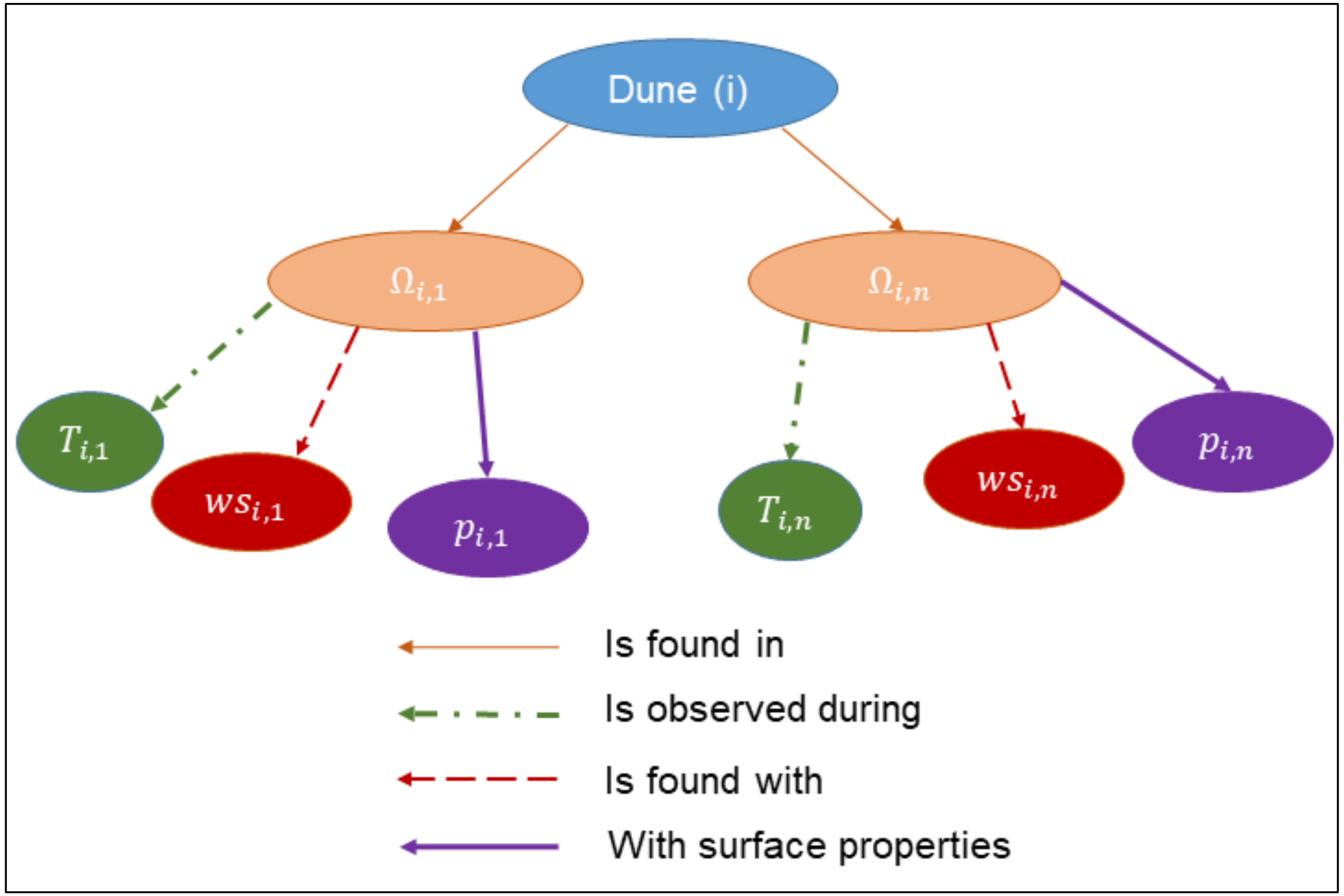
https://pds-geosciences.wustl.edu/missions/mgs/mola.html

Context-aware Novelty Detection

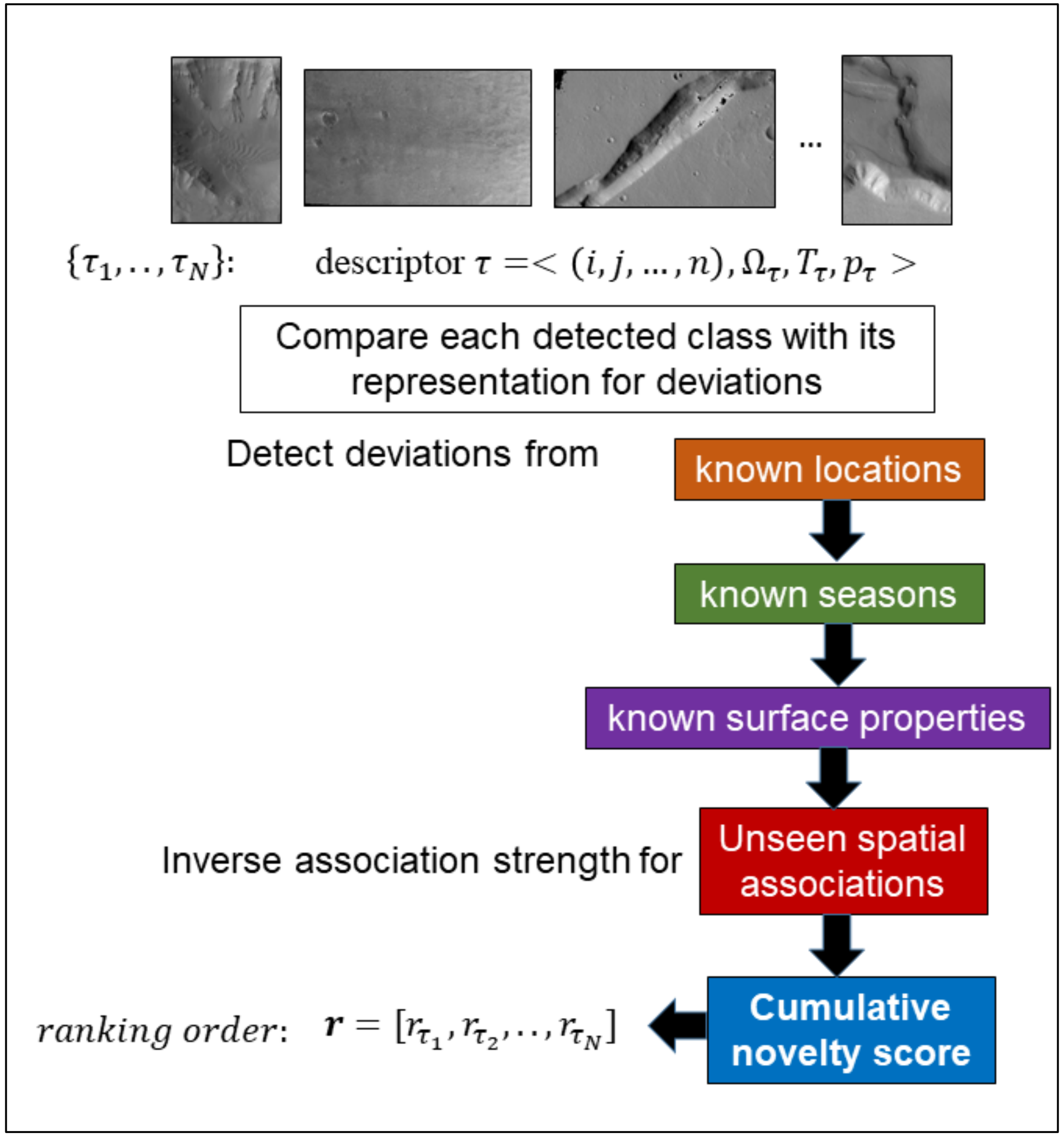


Learned representations (for each class/ event of interest)

Context-aware Novelty Detection

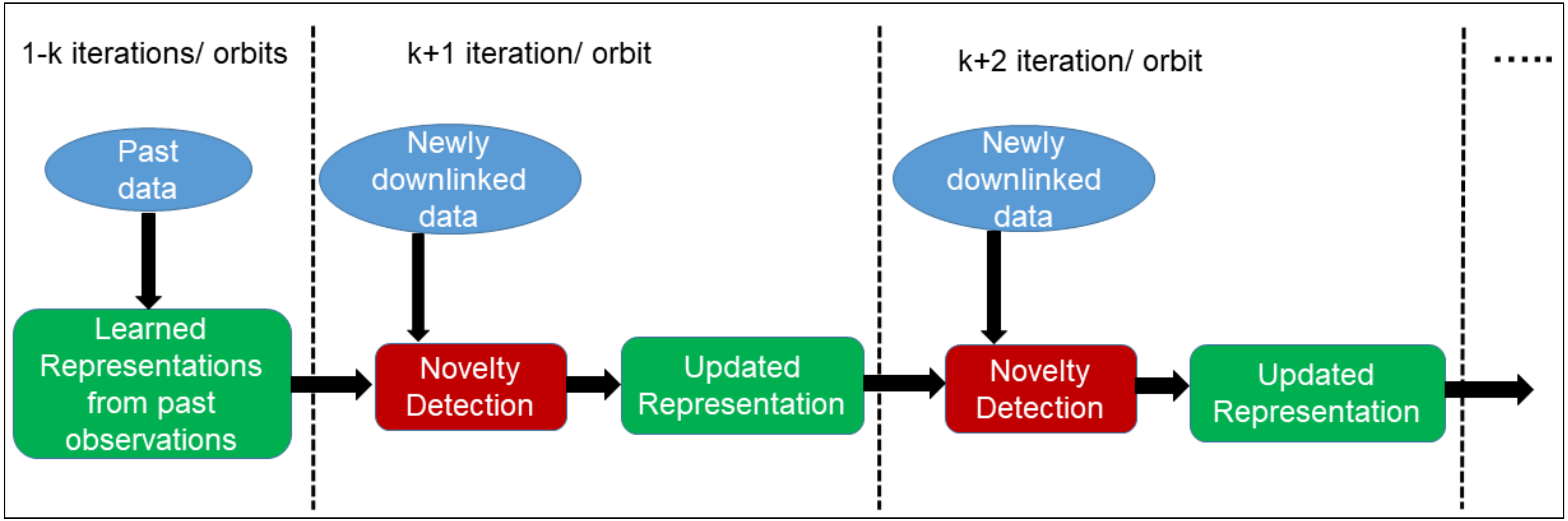


Learned representations (for each class/ event of interest)



Novelty detection from learned representations

Iterative Representation Update



Updated with expert guidance

- Merging clusters
- Updating weights of spatial association importance
- Updating list of interesting classes/ events

Acceptability of Novelty Detection to Experts

Mean average precision (MAP): Average precision at every position b where $r_\tau = e_\tau$, over all test set batches Q

$$MAP = \frac{1}{Q.B}\sum_{q=1}^{Q}\sum_{b=1}^{B}\gamma_b$$

r_τ- rank generated by the module

e_τ- rank from experts

where $\gamma_b = \begin{cases} p_b, & r_\tau = e_\tau \\ 0, otherwise \end{cases}$

Spearman Rank Correlation (SRC) between r_τ and e_τ

Acceptability of novelty ranking to experts

Representation	MAP	SRC
Standalone	0.1864	0.5482
Expert Guided	**0.765**	**0.984**

Acceptability of Novelty Detection to Experts

Mean average precision (MAP): Average precision at every position b where $r_\tau = e_\tau$, over all test set batches Q

$$MAP = \frac{1}{Q.B}\sum_{q=1}^{Q}\sum_{b=1}^{B}\gamma_b$$

r_τ- rank generated by the module

e_τ- rank from experts

where $\gamma_b = \begin{cases} p_b, & r_\tau = e_\tau \\ 0, otherwise \end{cases}$

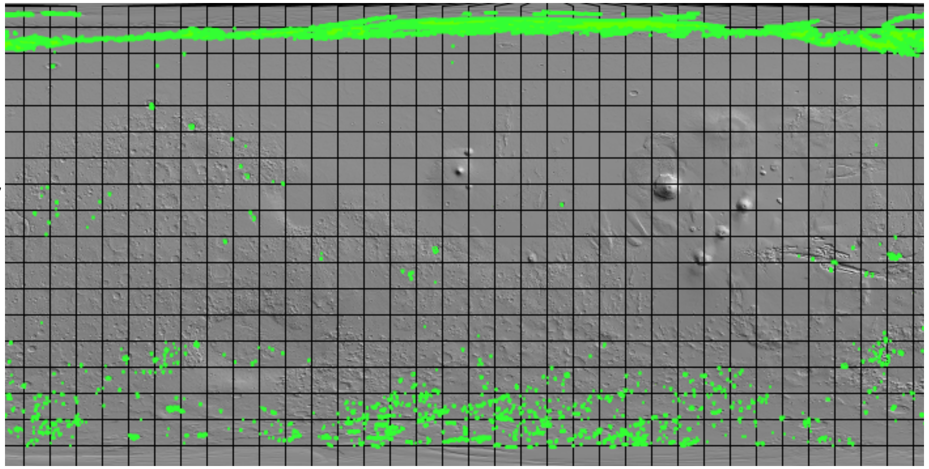
Spearman Rank Correlation (SRC) between $r_\tau\ and\ e_\tau$

Acceptability of novelty ranking to experts

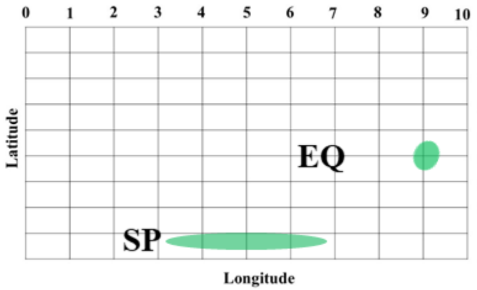
Representation	MAP	SRC
Standalone	0.1864	0.5482
Expert Guided	**0.765**	**0.984**

Examples of Detected Novelties:

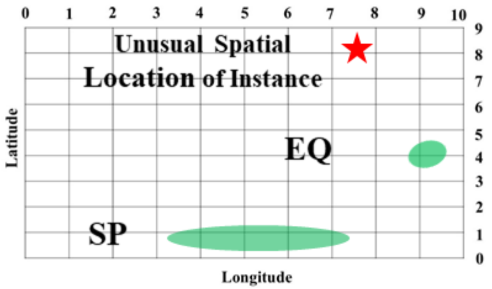
EQ- Equatorial; SP- South Polar; NP – North Polar



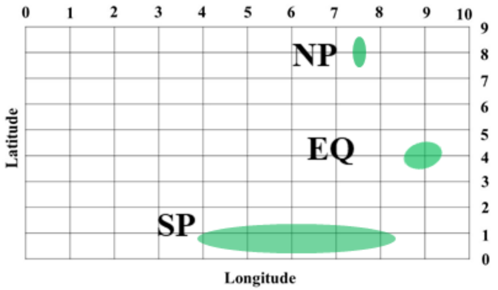
Dune locations:18 years of THEMIS observation



Iteration 1:
- Known spatial locations of dunes

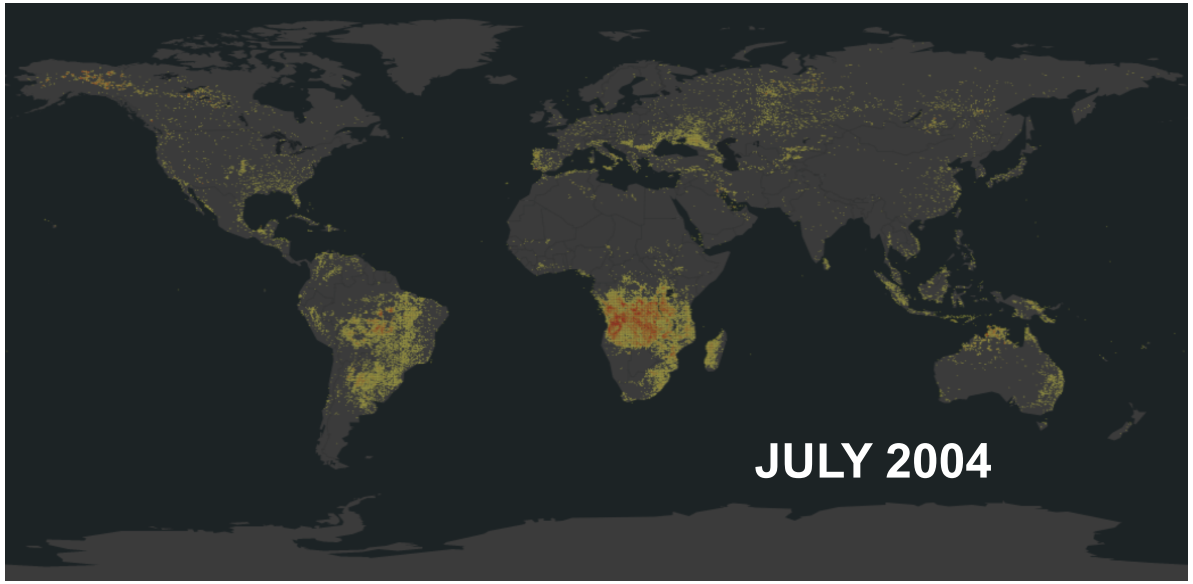


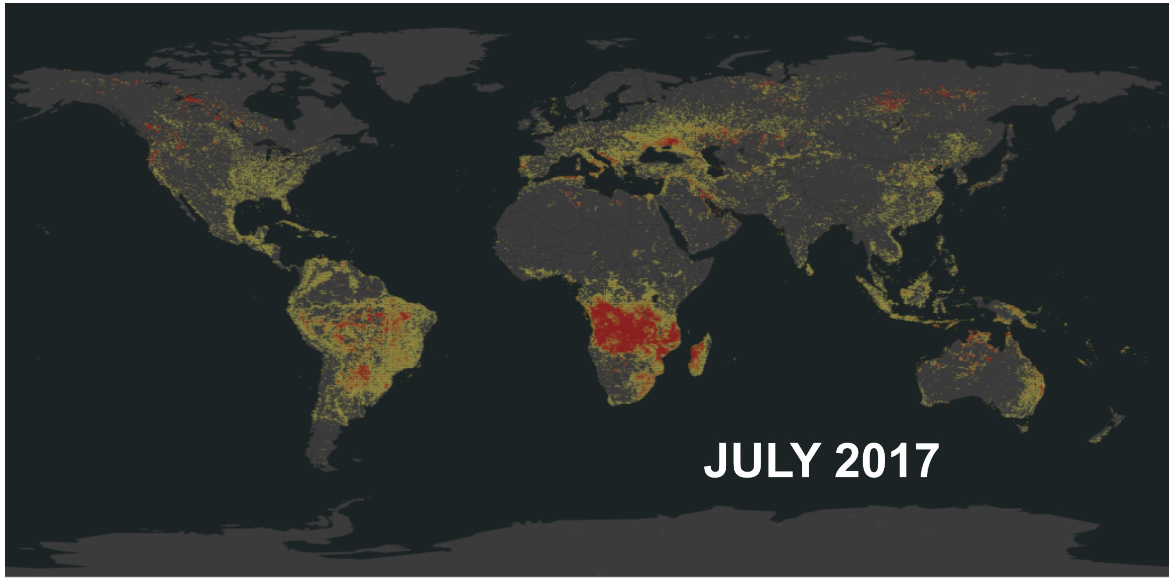
Iteration 2:
- Known spatial locations of dunes modified
- Dune observed at an unusual location



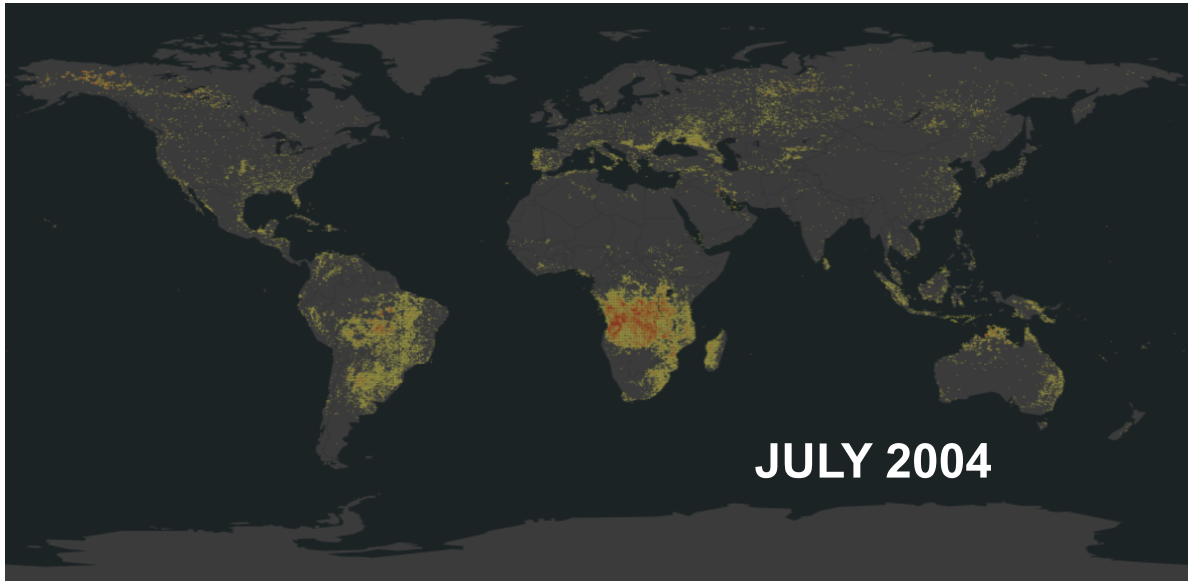
Iteration 3:
- Known spatial locations of dunes modified to include north pole

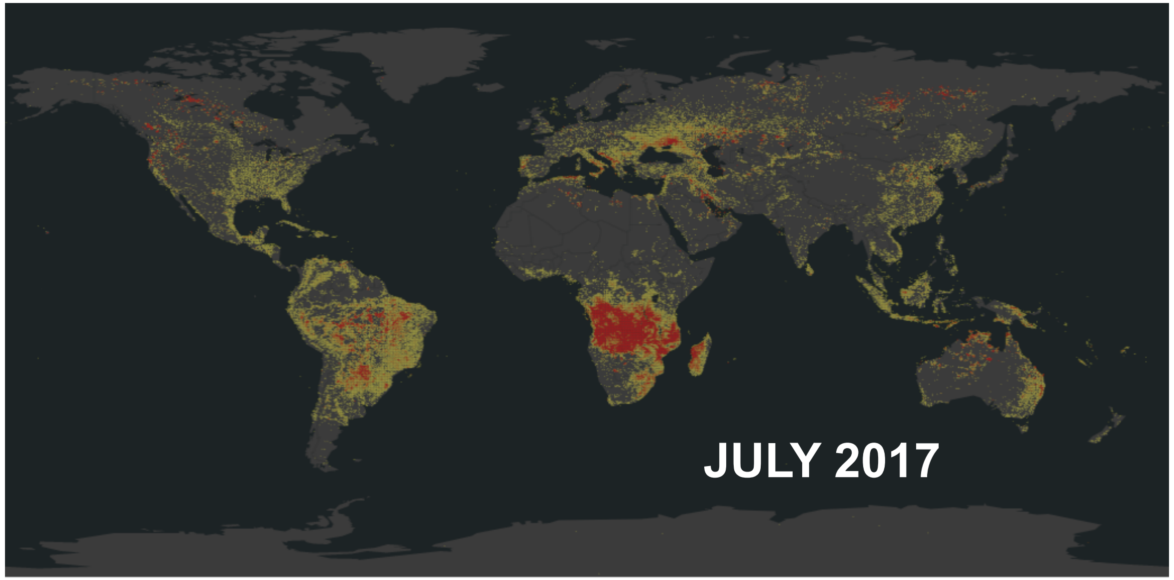
Context-Based Representation of Natural Hazards





Context-Based Representation of Natural Hazards





- Location
- Temporal information
- Temperature
- Precipitation
- Slope
- Land Cover Class
- Vegetation/ Fuel type
- Regional Deviation from Driving Factors

Factors that are known to cause or trigger the natural hazard

- Other Disturbances in Spatial/ Temporal Neighborhood

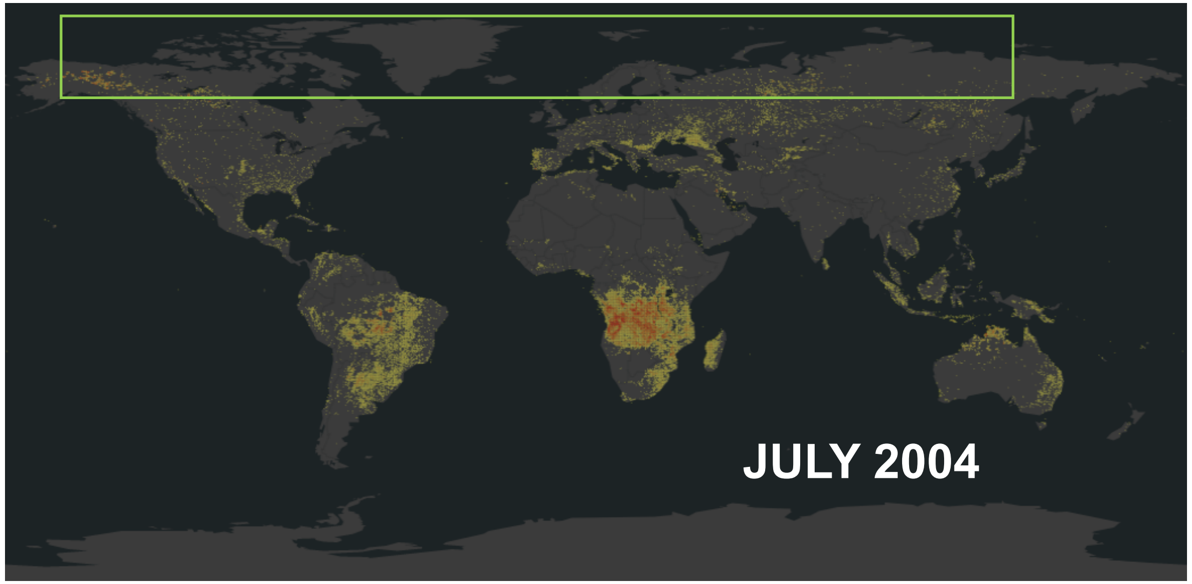
Contributing factors & effects

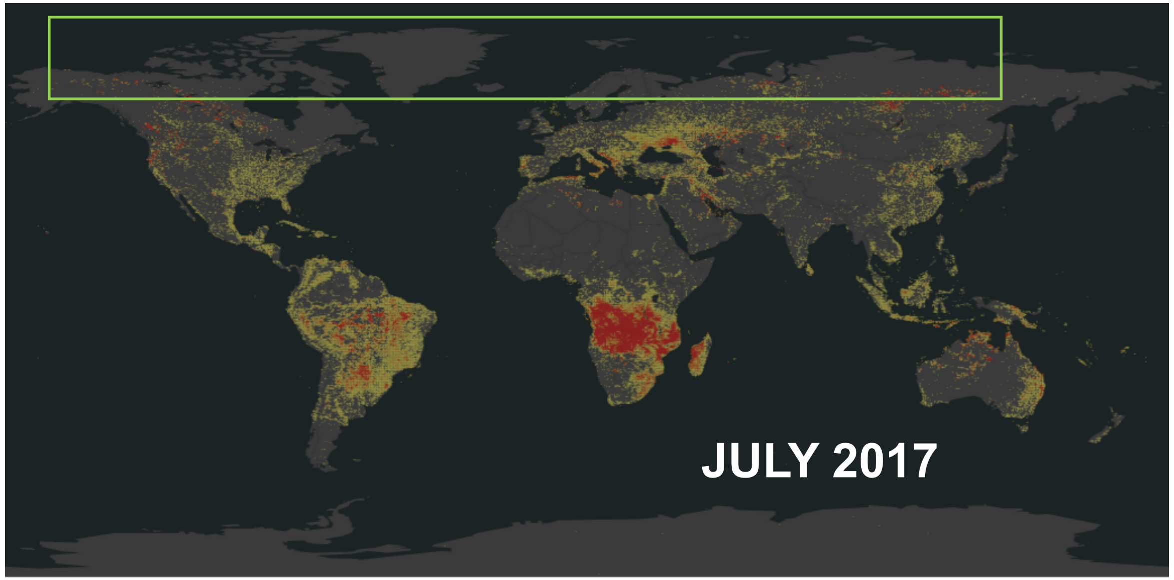
- Emission and air quality
- Burned Area

Effects of the natural hazard

Wildfire representations updated over time with every novel detection

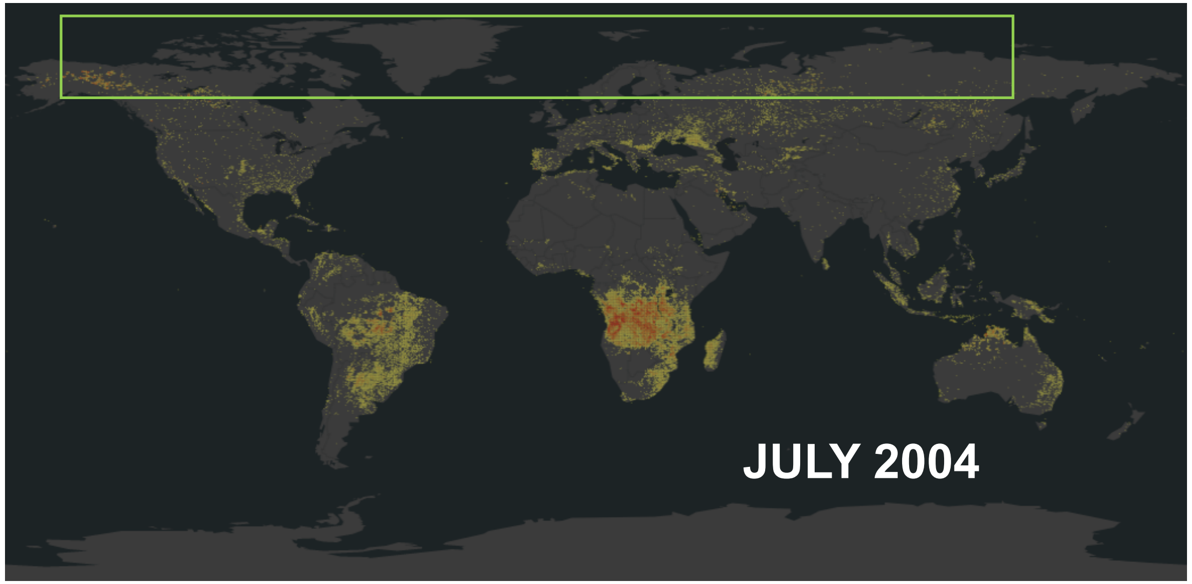
Novelty Detection from Context-Based Representation

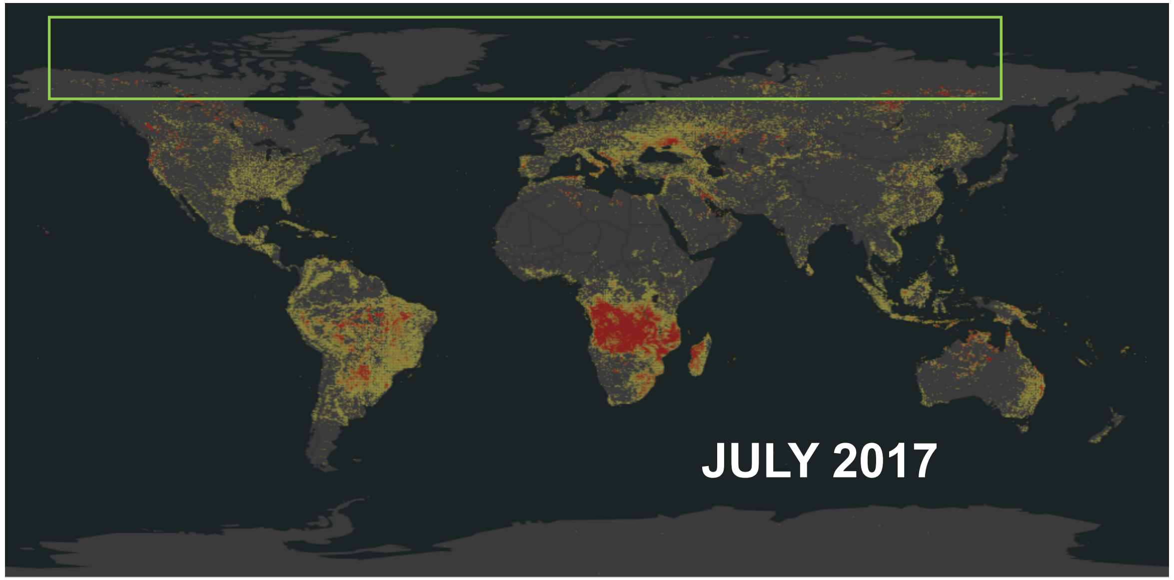


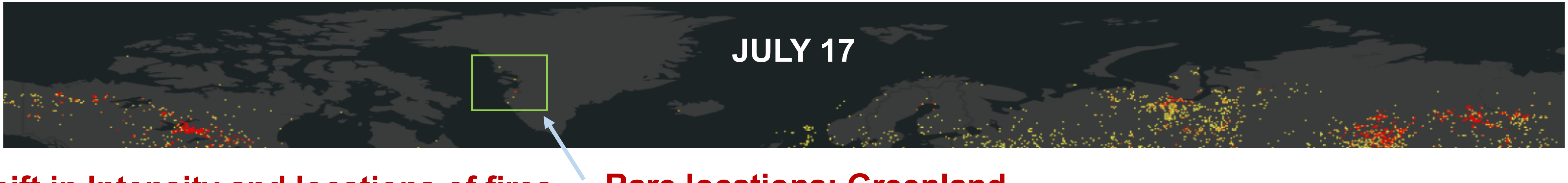


Shift in Intensity and locations of fires

Novelty Detection from Context-Based Representation







Shift in Intensity and locations of fires

Rare locations: Greenland

Future Directions

- Towards a more interpretable representation of novelties
- Applicable for retrieval, identifying context-based rare observations
- Incorporate additional relevant features and models (different modality)
  - data dimensionality
- Learn representations from varying volumes of data
  - early stages of an instrument/ less explored surfaces – low data volume
  - increasing data volume over time
- Expert feedback for adapting to an application (attributes, weights), evaluation
  - “interestingness” often relative, difficult to quantify

Acknowledgements

- JMARS: https://jmars.asu.edu/
- https://themis.mars.asu.edu/
- https://search.earthdata.nasa.gov/search
- Philip Christensen
- Mars Space Flight Facility