**Current Draft:**

**NBD(NIST Big Data) Requirements WG Use Case Template**

|  |  |  |  |
| --- | --- | --- | --- |
| **Use Case Title** | | Object identification and tracking from Wide Area Large Format Imagery (WALF) Imagery or Full Motion Video (FMV) – Persistent Surveilance | |
| **Vertical (area)** | | Defense (Intelligence) | |
| **Author/Company/Email** | | David Boyd/Data Tactics/dboyd@data-tactics.com | |
| **Actors/Stakeholders and their roles and responsibilities** | | 1. Civilian Military decision makers 2. Intelligence Analysts 3. Warfighters | |
| **Goals** | | To be able to process and extract/track entities (vehicles, people, packages) over time from the raw image data. Specifically, the idea is to reduce the petabytes of data generated by persistent surveillance down to a manageable size (e.g. vector tracks) | |
| **Use Case Description** | | Persistent surveillance sensors can easily collect petabytes of imagery data in the space of a few hours. It is unfeasible for this data to be processed by humans for either alerting or tracking purposes. The data needs to be processed close to the sensor which is likely forward deployed since it is too large to be easily transmitted. The data should be reduced to a set of geospatial object (points, tracks, etc.) which can easily be integrated with other data to form a common operational picture. | |
| **Current**  **Solutions** | **Compute(System)** | | Various – they range from simple storage capabilities mounted on the sensor, to simple display and storage, to limited object extraction. Typical object extraction systems are currently small (1-20 node) GPU enhanced clusters. |
| **Storage** | | Currently flat files persisted on disk in most cases. Sometimes RDBMS indexes pointing to files or portions of files based on metadata/telemetry data. |
| **Networking** | | Sensor comms tend to be Line of Sight or Satellite based. |
| **Software** | | A wide range custom software and tools including traditional RDBM’s and display tools. |
| **Big Data  Characteristics** | **Data Source (distributed/centralized)** | | Sensors include airframe mounted and fixed position optical, IR, and SAR images. |
| **Volume (size)** | | FMV – 30-60 frames per/sec at full color 1080P resolution.  WALF – 1-10 frames per/sec at 10Kx10K full color resolution. |
| **Velocity**  **(e.g. real time)** | | **Real Time** |
| **Variety**  **(multiple datasets, mashup)** | | Data Typically exists in one or more standard imagery or video formats. |
| **Variability (rate of change)** | | **Little** |
| **Big Data Science (collection, curation,**  **analysis,**  **action)** | **Veracity (Robustness Issues)** | | The veracity of extracted objects is critical. If the system fails or generates false positives people are put at risk. |
| **Visualization** | | Visualization of extracted outputs will typically be as overlays on a geospatial display. Overlay objects should be links back to the originating image/video segment. |
| **Data Quality** | | Data quality is generally driven by a combination of sensor characteristics and weather (both obscuring factors - dust/moisture and stability factors – wind). |
| **Data Types** | | Standard imagery and video formats are input. Output should be in the form of OGC compliant web features or standard geospatial files (shape files, KML). |
| **Data Analytics** | | 1. Object identification (type, size, color) and tracking. 2. Pattern analysis of object (did the truck observed every weds afternoon take a different route today or is there a standard route this person takes every day). 3. Crowd behavior/dynamics (is there a small group attempting to incite a riot. Is this person out of place in the crowd or behaving differently. 4. Economic activity    1. is the line at the bread store, the butcher, or the ice cream store,    2. are more trucks traveling north with goods than trucks going south    3. Has activity at or the size of stores in this market place increased or decreased over the past year. 5. Fusion of data with other data to improve quality and confidence. |
| **Big Data Specific Challenges (Gaps)** | | Processing the volume of data in NRT to support alerting and situational awareness. | |
| **Big Data Specific Challenges in Mobility** | | Getting data from mobile sensor to processing | |
| **Security & Privacy**  **Requirements** | | Significant – sources and methods cannot be compromised the enemy should not be able to know what we see. | |
| **Highlight issues for generalizing this use case (e.g. for ref. architecture)** | | Typically this type of processing fits well into massively parallel computing such as provided by GPUs. Typical problem is integration of this processing into a larger cluster capable of processing data from several sensors in parallel and in NRT.  Transmission of data from sensor to system is also a large challenge. | |
| **More Information (URLs)** | | Motion Imagery Standards - <http://www.gwg.nga.mil/misb/>  Some of many papers on object ident/tracking: <http://www.dabi.temple.edu/~hbling/publication/SPIE12_Dismount_Formatted_v2_BW.pdf>  <http://csce.uark.edu/~jgauch/library/Tracking/Orten.2005.pdf>  <http://www.sciencedirect.com/science/article/pii/S0031320305004863>  General Articles on the need: <http://www.militaryaerospace.com/topics/m/video/79088650/persistent-surveillance-relies-on-extracting-relevant-data-points-and-connecting-the-dots.htm>  <http://www.defencetalk.com/wide-area-persistent-surveillance-revolutionizes-tactical-isr-45745/>  <http://www.defencetalk.com/wide-area-persistent-surveillance-revolutionizes-tactical-isr-45745/> | |
| **Note:** <additional comments> | | | |