BROOKHAVEN NATIONAL LABORATORY

Hidden Network Monitoring

Michael DePhillips

Brookhaven National Laboratory

BNL Virtual Symposium for ETI/MTV Consortia

April 7, 2021

BROOKHAVEN NATIONAL LABORATORY

Overview – General Statements & Hypothesis

- Analysis of Publicly Available Information (PAI) can offer valuable information for understanding nonproliferation and weapons development. AI can help this analysis.
- Specifically - Natural Language Processing (NLP) has advanced from sentiment analysis - to understanding the intent of the communicator (psychology and health).
- Expert Knowledge aka Human-in-the-loop can compensate sparse and/or sporadic events in certain domains (e.g., directed scientific discovery).
- *Using NLP with domain knowledge of SMEs; resulting sequences and clustering of queries can help identify signatures indicating some probabilistic measure of a user's intent (academic searches vs. searches for the development of a weapons program).*

BROOKHAVEN NATIONAL LABORATORY

Overview – Technique and Domain

General

- AI analysis of requests to Publicly Available Information (PAI) to help determine a user's "intent".

Specific

- Use Natural Language Processing on queries to Publicly Available repository of Nuclear Science Data at DOE national lab to reveal *questionable* intent.

"The use of publicly available data and analysis is effective at identifying high-risk nuclear trade."
"Machine learning tools can be used to dramatically enhance data analysis in terms of both speed and quality."

**E. Moniz: Signals in the Noise –*
PREVENTING NUCLEAR PROLIFERATION WITH MACHINE LEARNING & PUBLICLY AVAILABLE INFORMATION
C4STS / NTI 2021

BROOKHAVEN NATIONAL LABORATORY

Overview – Technical Approach

- Backend Data Collection
  - Operate on data while in the network and in transit
    - Collect, clean, filter, bin and hand-off to Analysis
    - Possible use includes preprocessing and some analysis
- NLP
  - Short text clustering (queries which not grammatically correct)
    - building good representation
  - Early classification
    - through multi-instance learning
  - Intent detection (probabilistically predictive / through evidence found through a sequence of queries)
- SME
  - Obtain cross-complex expert knowledge regarding query relevance to input into system

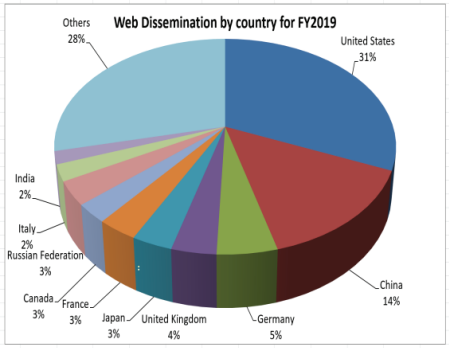
BROOKHAVEN NATIONAL LABORATORY

Goals - General

- Early Detection of a Weapons Program or Proliferation Activity
  - Gathering of Foundational Science – (as far upstream as possible)
    - e.g., Literature searches are revealing
  - Use of collective SME knowledge to aid in detection analysis
- Monitoring of PAI containing valuable information
- Creation of a silent data collection system that supports monitoring and contributes to the analysis on the fly and in near real-time

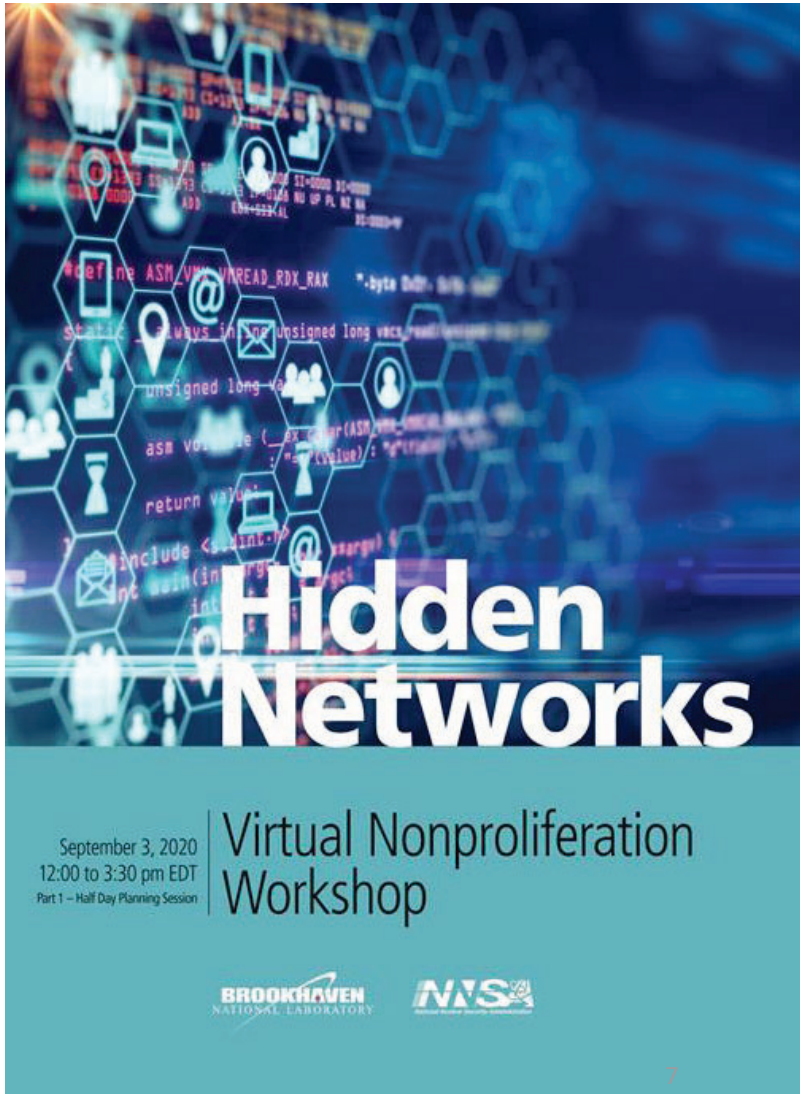
Goals - Additional Community Benefits

- A network-based filtering / cleaning / analysis tool
  - Efficient
  - Rapid Response
  - Federation
  - Clandestine
- Contributing to a Self Supervised approach toward early classification problems
- Domain Aware NLP – “human-in-loop”
  - Greater understanding on the PAI and its users
- Monitoring of relevant PAI data repository
  - Simple anomalous behavior
  - Domain relevant statistics
- Code that is adjustable, reusable and scalable
  - Modular and not data specific



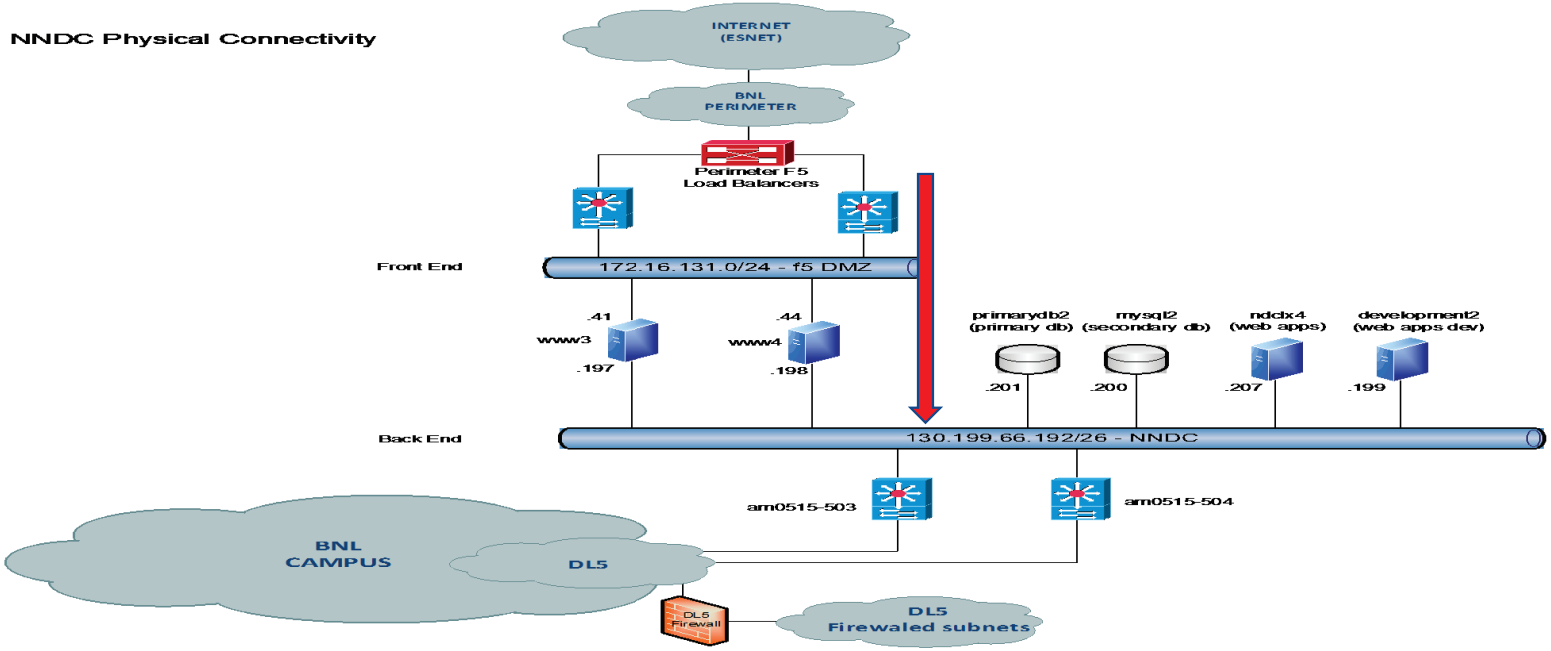
Nonproliferation (human-in-loop)

- Literature review conducted to identify materials of interest and results distributed to points of contact for feedback
- Log review – Eyeball/simple script analysis to identify clustering/significant IPs
- Following up on other recommendations from the September workshop
- Provide expert knowledge regard significance of sequences for AI tuning



Collection progress - Network

Configured and installed monitoring node in production networking architecture
Packets captured and NNDC logs reconstructed



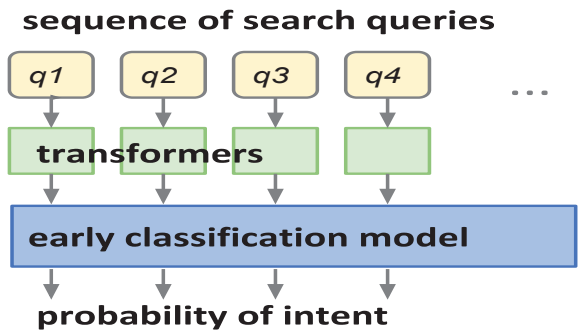
Analysis

Intent Detection from Search Sequences

Approach:

Transformers -- state of the art deep learning for natural language (Devlin et al., 2019)

Early multi-Instance classification -- statistical learning problem to produce classifications at various input sequence sample sizes. (Dennis et al., 2018)



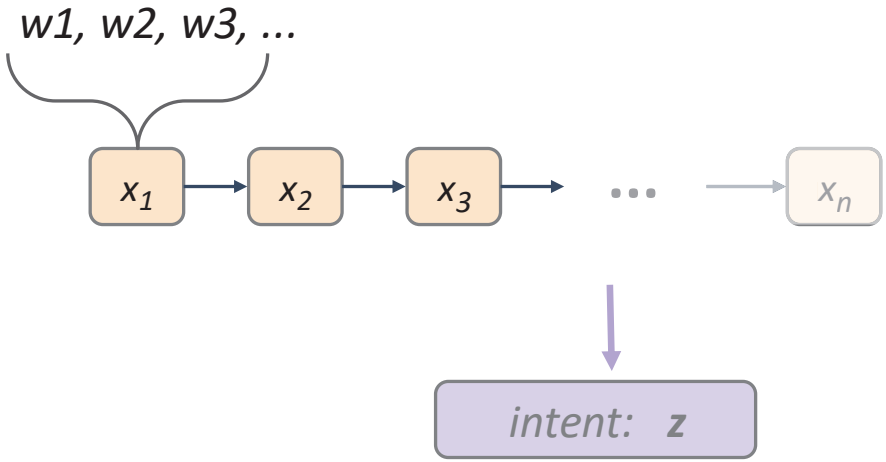
July, 2020	September	November	January, 2021
Related work survey	Transformers adapted / fine-tuned to problem	Unsupervised approach development	**Deliverable:** Documented Code via Jupyter Notebook

Progress - Analysis

- Analysis Code – Nutshell
  - Creates NL Higher Dimension Representations with primary keys then sets into sized vector. Creates clustering of significant sequences.
    - Queries pass through transformers (RoBERTa-base) to be embedded and then the sequence of embeddings (produced from the transformers) are fed to an RNN-LSTM to do the classification.

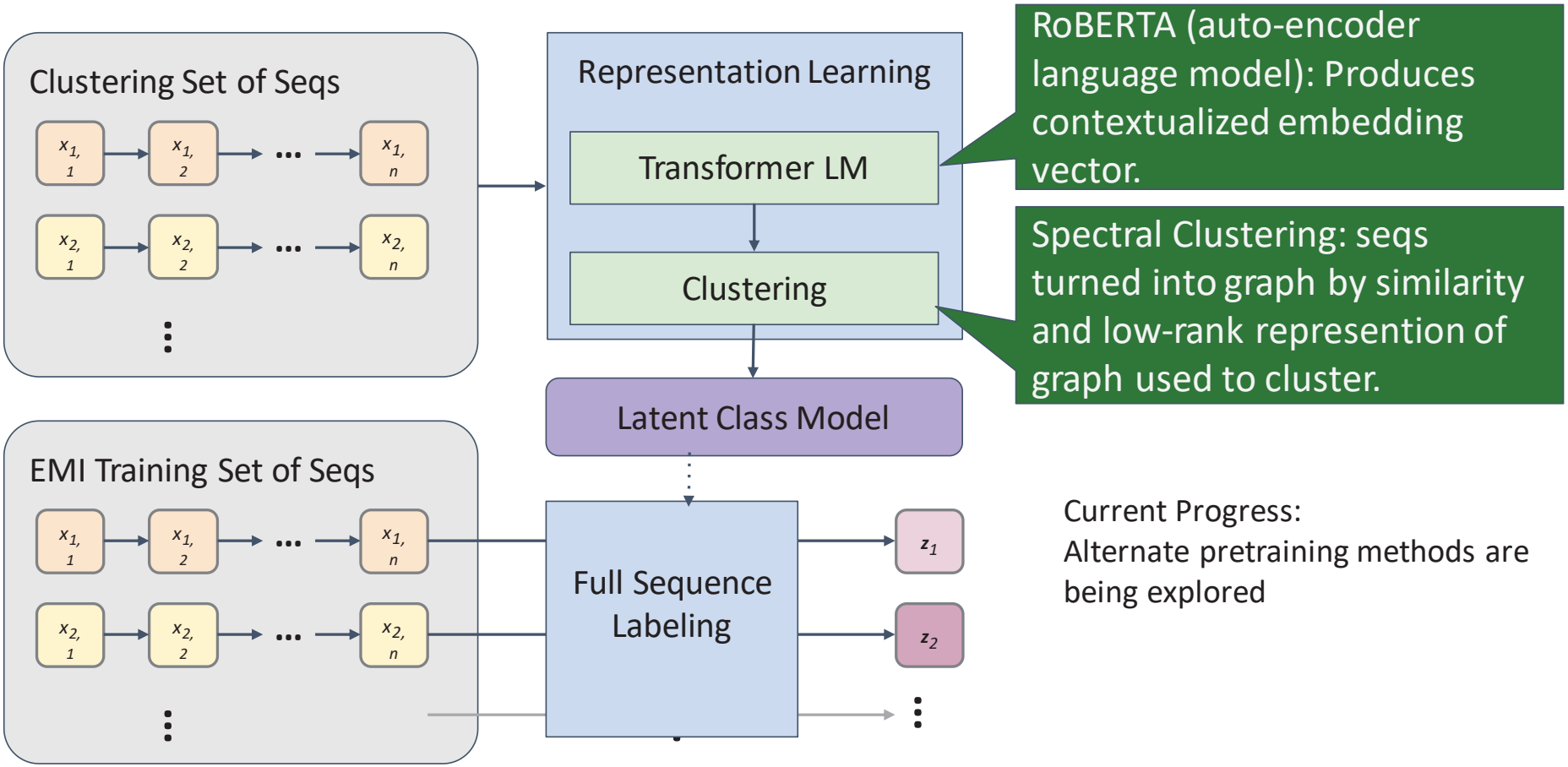
Each query is a sequence of *keywords or natural language*

Each query is part of a sequence of multiple queries with an intended target.



BROOKHAVEN NATIONAL LABORATORY

Self Supervised Approach – Early Classification



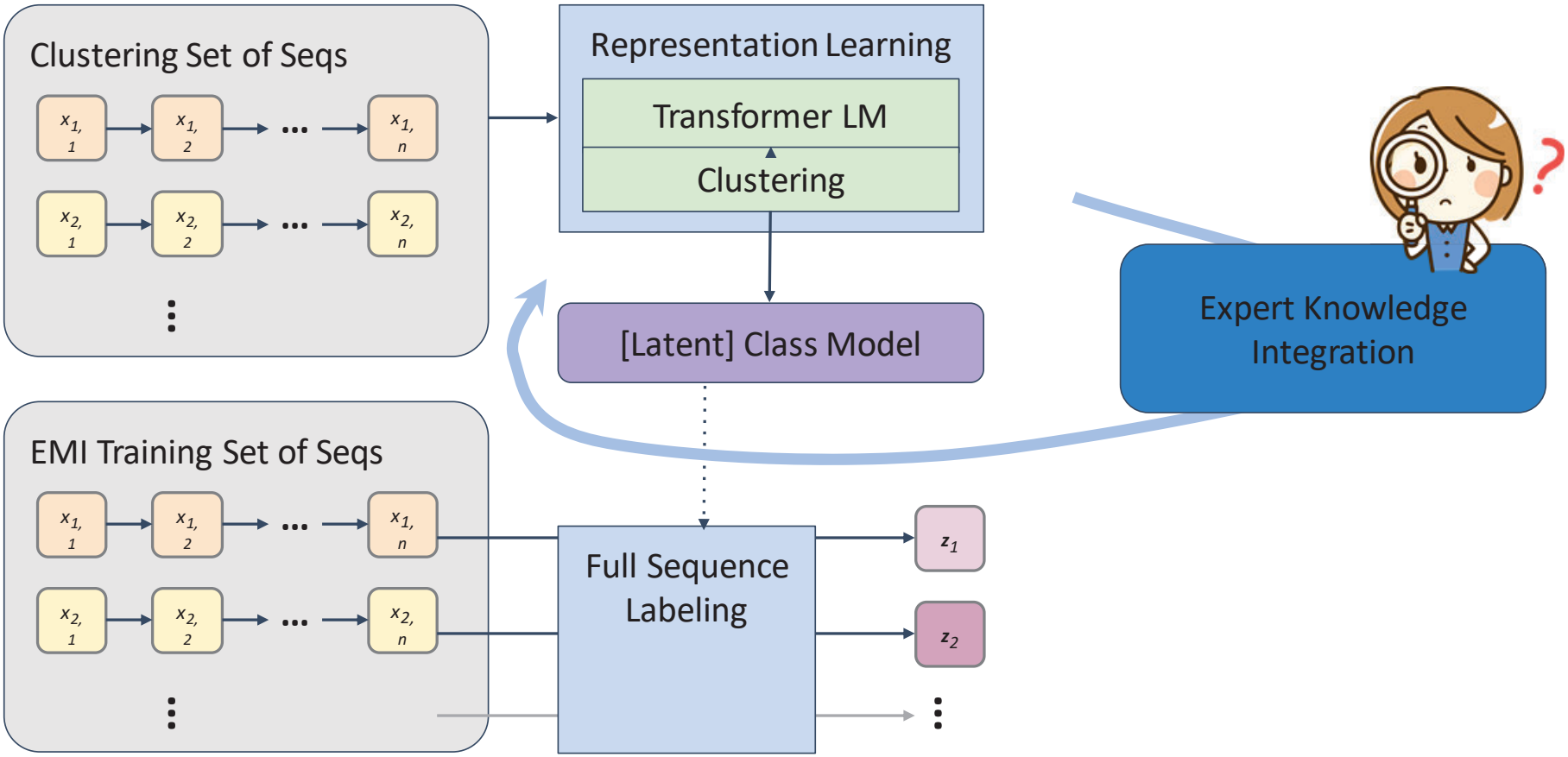
RoBERTA (auto-encoder language model): Produces contextualized embedding vector.

Spectral Clustering: seqs turned into graph by similarity and low-rank represention of graph used to cluster.

Current Progress:
Alternate pretraining methods are being explored

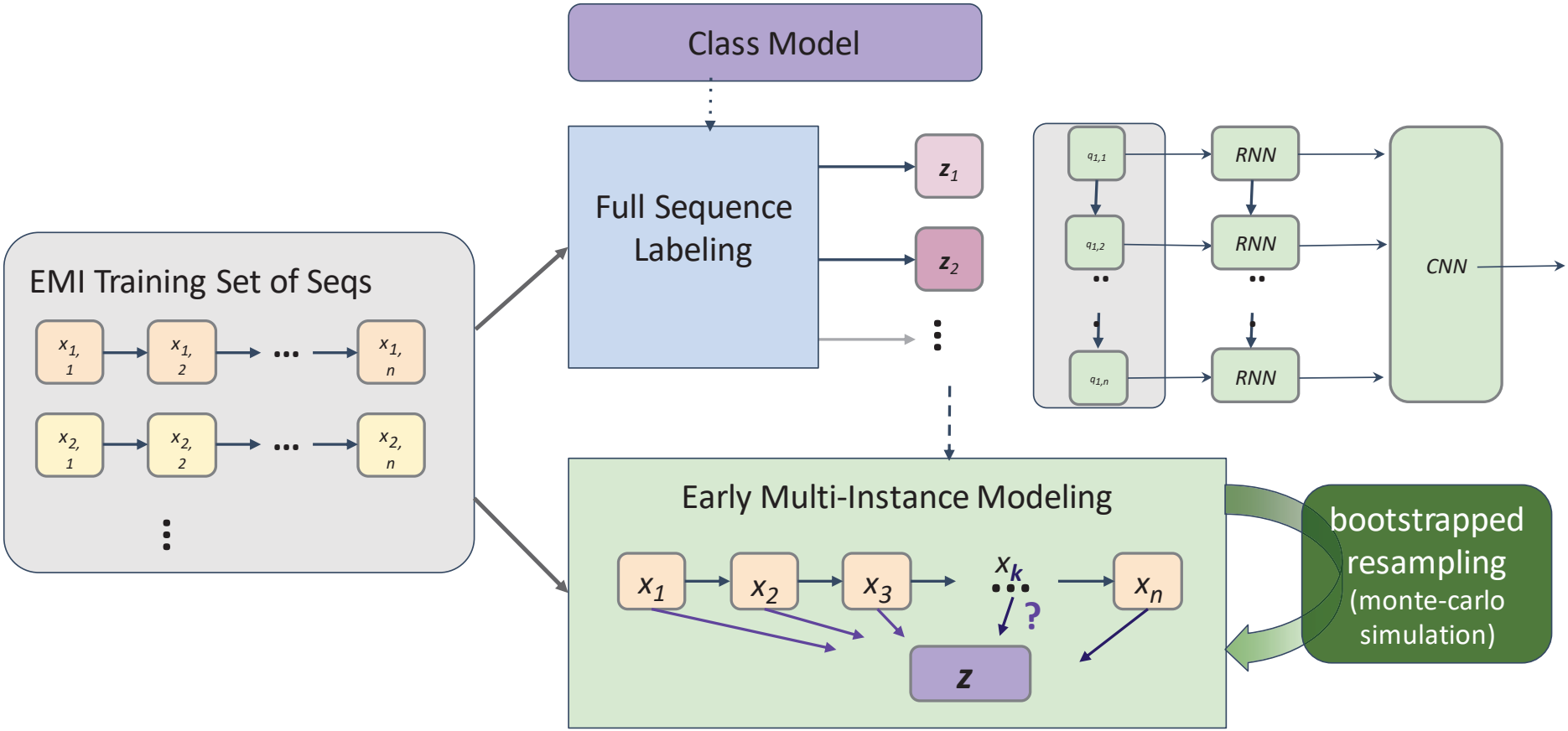
BROOKHAVEN NATIONAL LABORATORY

Fine Tuning Early Classification with Expert Knowledge



BROOKHAVEN NATIONAL LABORATORY

Prediction Confidence Evaluation (95%)



Clustering –Sanity Check

Reviewing the clustering results by looking at queries from 5 sequences from each cluster

In [258]:
```
records = queryData_1.data["values"]
for cluster_num in range(num_clusters):
    random.seed(42)
```

```
Upto first 10 queries of 5 random sequences from cluster 0 (separated by ;):

[(1, 'Barium; 116Barium; 116 Barium; Barium 116; Barium; Barium,production'),
 (2, '32 IMME; 32 IMME; IMME'),
 (3, 'helios,gas cell; helium,implanted; 3he,implanted'),
 (4, 'Slater approximation'),
 (5, 'beam energy,spin; spin')]
```


```
[(1, 'Barium; 116Barium; 116 Barium; Barium 116; Barium; Barium,production'),
 (2, '32 IMME; 32 IMME; IMME'),
 (3, 'helios,gas cell; helium,implanted; 3he,implanted'),
 (4, 'Slater approximation'),
 (5, 'beam energy,spin; spin')]
```

BROOKHAVEN NATIONAL LABORATORY

Next steps

- Establish information of interest
- Establish baseline on historic data
- Forward information directly to the algorithm – streaming logs
- It may be possible to correlate multiple queries/responses to extract such information, but this moves from filtering to more intelligent processing and it remains to be decided where and how this should take place.
- Implementing a more state-of-the-art machine learning technique to do the classification
- Integrating expert knowledge to improve the classes/labeling/clusters
- Iterative evaluation - integrating additional data
  - Accuracy / performance
  - Metadata
- Process historic data