CONTEXT-AWARE GEOGRAPHIC INFORMATION SYSTEMS FOR REAL-TIME SECURITY EVENT FORECASTING

Richard Wen

rwen@ryerson.ca

Department of Civil Engineering
Ryerson University

August 28, 2018

Research proposal to fulfil a requirement for the degree of Doctor of Philosophy in Civil Engineering

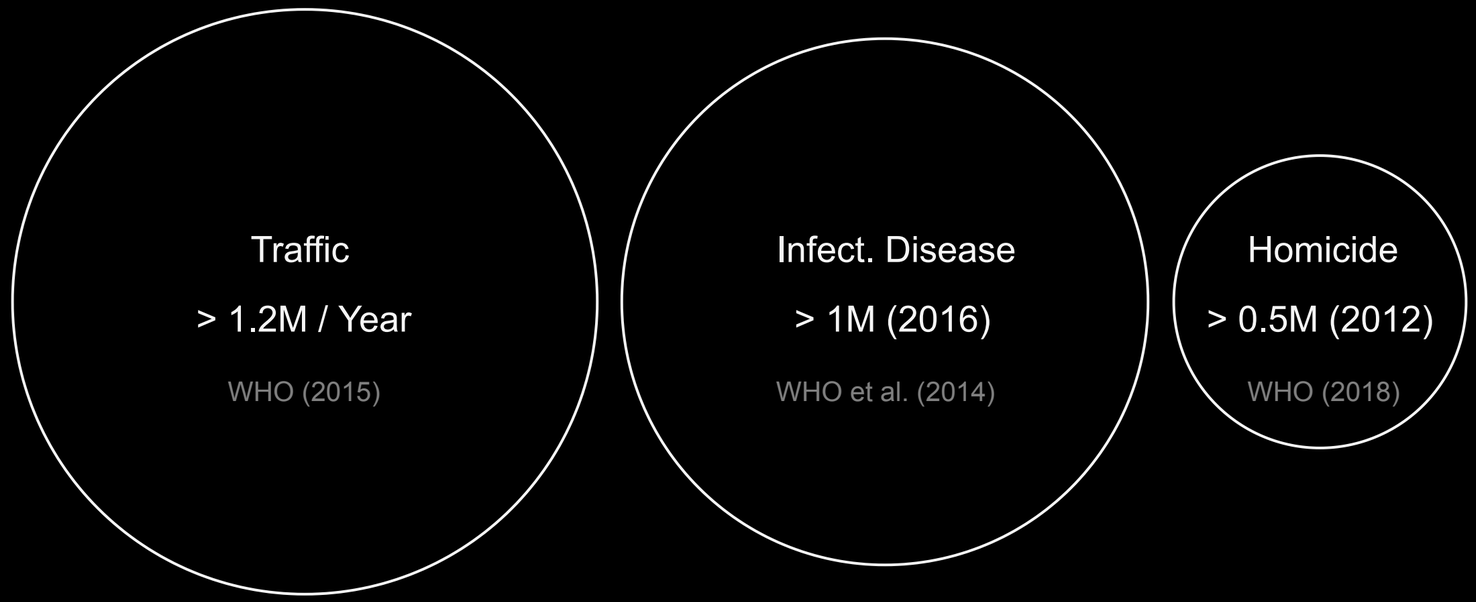
Ryerson University

OUTLINE

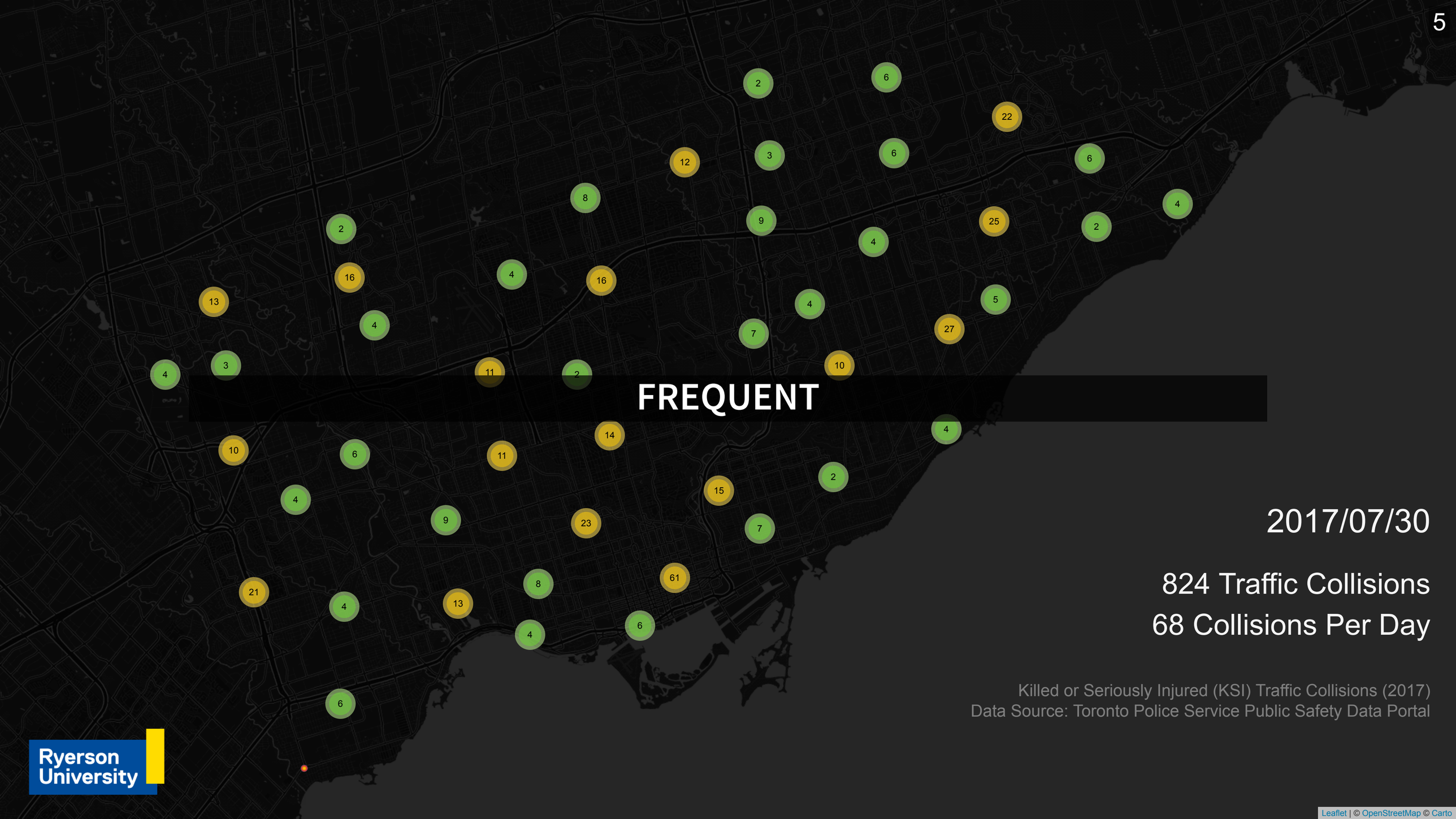
1. Introduction
2. Background
3. Methods
4. Preliminary Results
5. Conclusion
6. References

INTRODUCTION

OVER 40 MILLION DEATHS PER YEAR



Ryerson University



FREQUENT

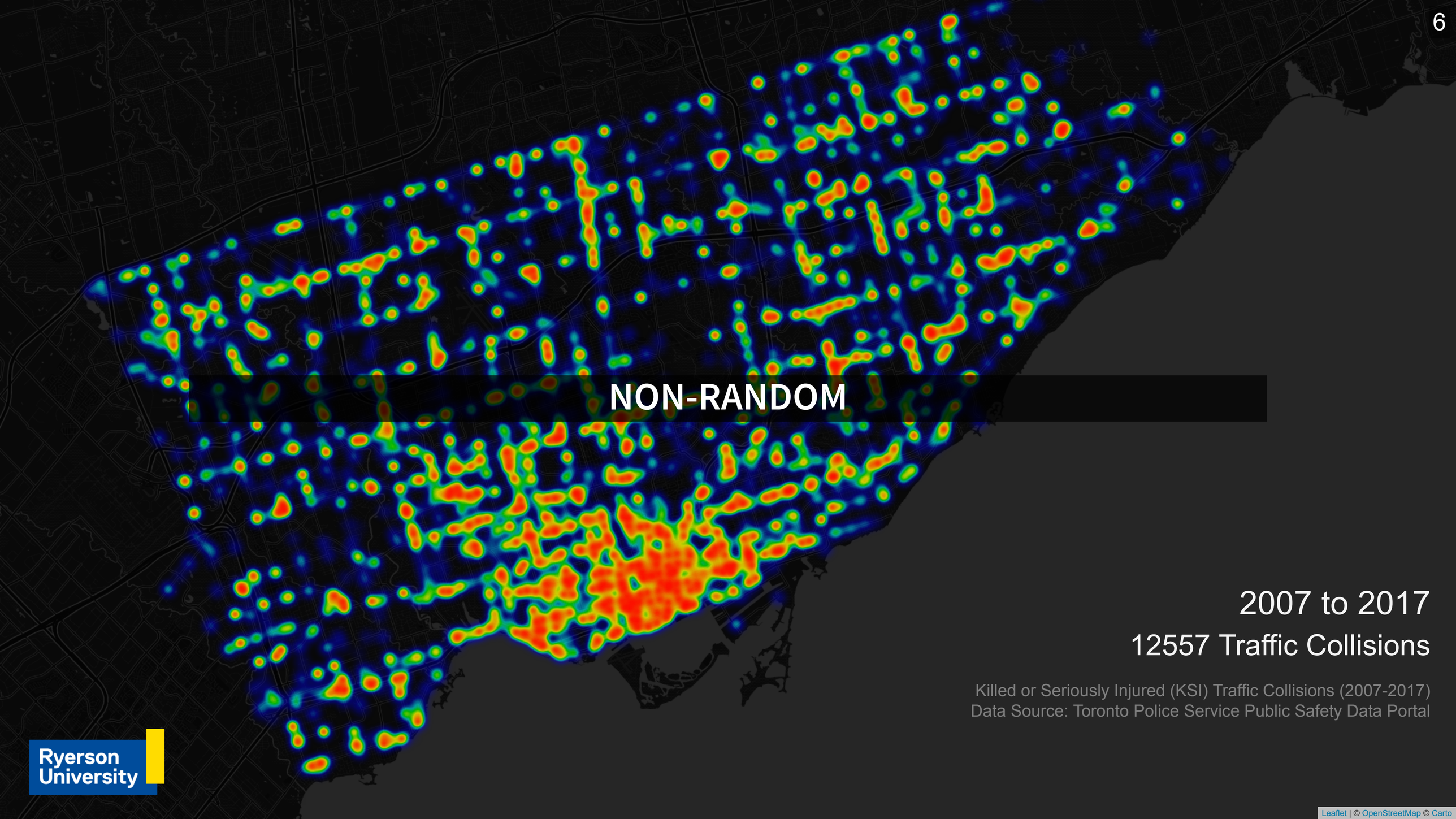
2017/07/30

824 Traffic Collisions

68 Collisions Per Day

Killed or Seriously Injured (KSI) Traffic Collisions (2017)
Data Source: Toronto Police Service Public Safety Data Portal

Ryerson University



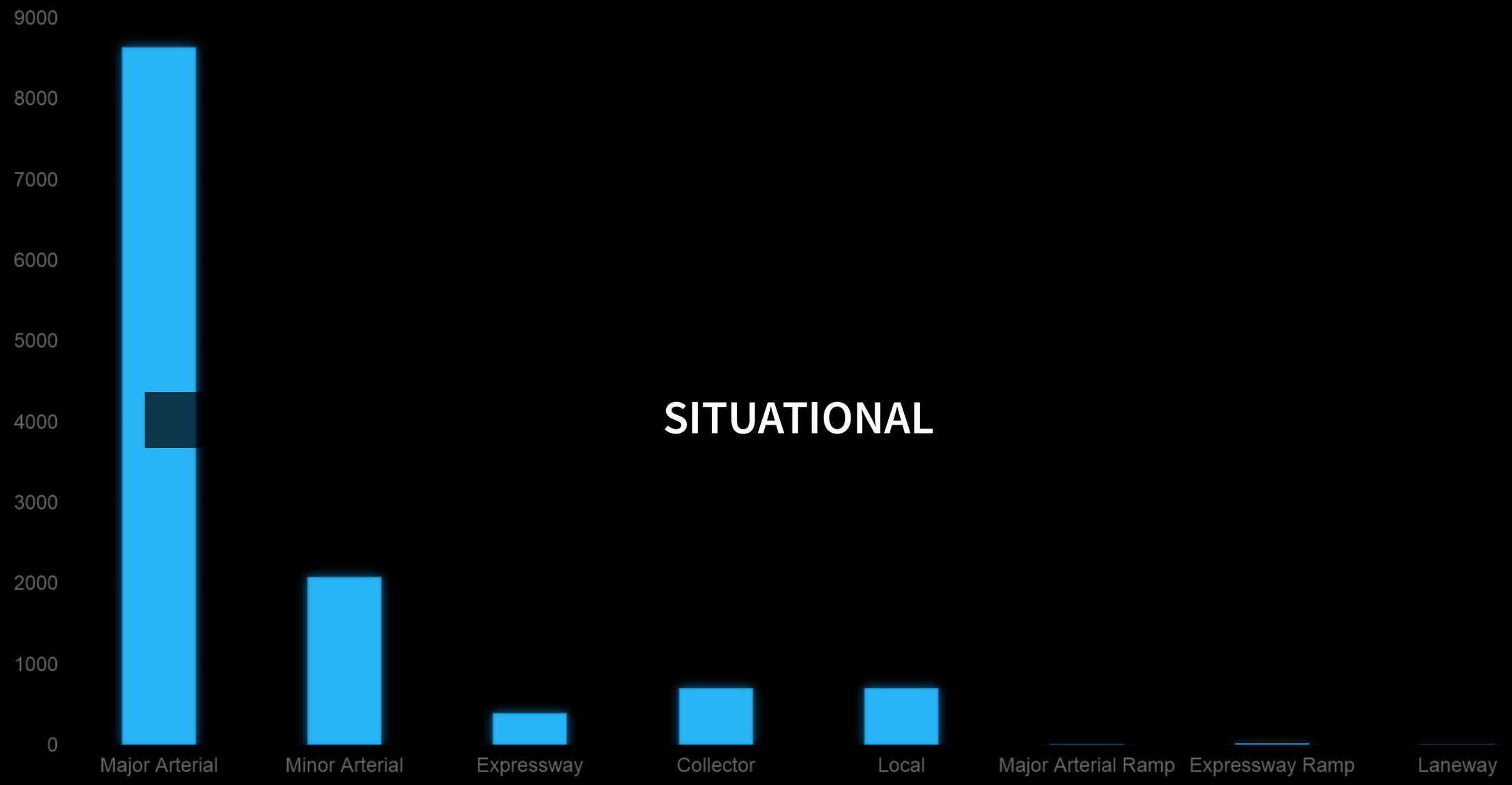
2007 to 2017

12557 Traffic Collisions

Killed or Seriously Injured (KSI) Traffic Collisions (2007-2017)
Data Source: Toronto Police Service Public Safety Data Portal

Ryerson University

SITUATIONAL



Killed or Seriously Injured (KSI) Traffic Collisions (2007-2017) from Toronto Police Service Public Safety Data Portal

Ryerson University

SECURITY EVENTS

- Disrupt societal operation and daily life
- Reduced quality of life
- Economic burden and loss of life

FORECASTING SECURITY EVENTS

1. **Data**: news, social media, government
2. **Context**: situation surrounding event
3. **Certainty**: probability, details

GEOGRAPHIC INFORMATION SYSTEMS (GIS)

- Set of hardware and software
- Store, manage, and present spatial data
- Enable information extracted from data

CONTEXT-AWARE GIS

- Enable GIS to use contextual data
- Adapt to context
- Act based on context

RESEARCH OBJECTIVES

1. Develop methods/models for context-aware GIS
2. Conduct experiments with context-aware GIS

OBJECTIVE 1 COMPONENTS

1. **Data Source**: contextual data extraction
2. **Database**: store and query 1.
3. **Data**: process and manipulate 2.
4. **Model**: forecast security events using 3.
5. **Visualization**: present outputs of 4.

OBJECTIVE 2 CRITERIA

1. **Relevant**: involves security event forecasting
2. **Practical**: uses non-artificial data
3. **Measurable**: uses qualitative or quantitative forecasting model assessments

CONTRIBUTIONS

1. Conceptual framework for context-aware GIS
2. Software framework for context-aware GIS
3. Software architecture for context-aware GIS
4. Methods for forecasting security events

BACKGROUND

WEB GIS AND ARCHITECTURES

- Client server
- Service-oriented
- Cloud-based

CONTEXT-AWARE SYSTEMS

"A system is context-aware if it uses context to provide relevant information and or services to the user, where relevancy depends on the user's task."

Ref: Dey (2001)

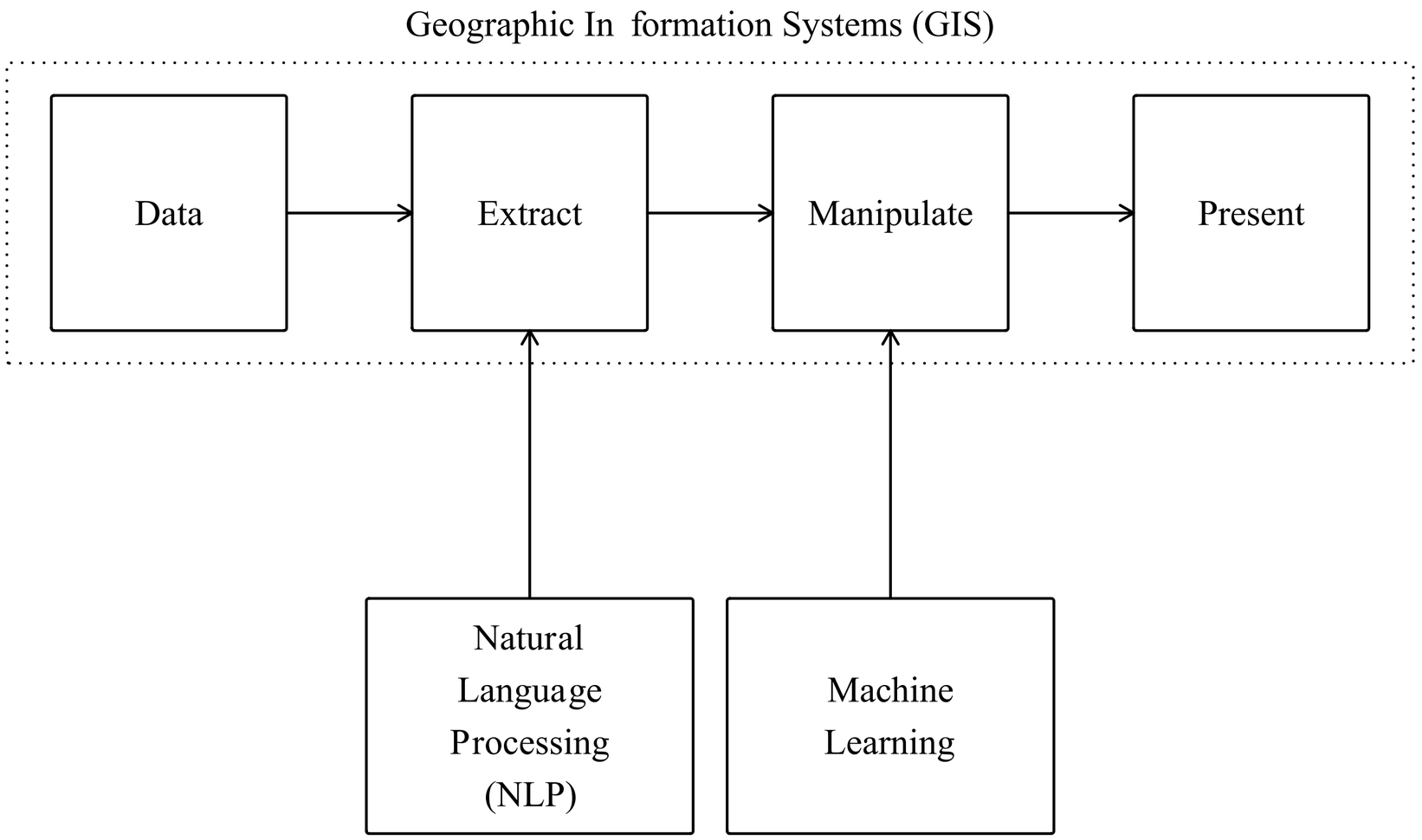
NATURAL LANGUAGE PROCESSING (NLP)

- Extracting useful data from text
- Structure contextual data
- N-grams, word distributions

MACHINE LEARNING

- Discover and apply patterns from data
- **Supervised**: predict from known values
- **Clustering**: grouping similar data
- **Incremental Learning**: continuous model updates

CONTEXT-AWARE GIS INTEGRATION



PROBLEM STATEMENT

- **Real-time GIS**: large continuous spatial data
- **Event-driven Architecture**: react to events
- **Context-aware GIS**: react and adapt to context

METHODS

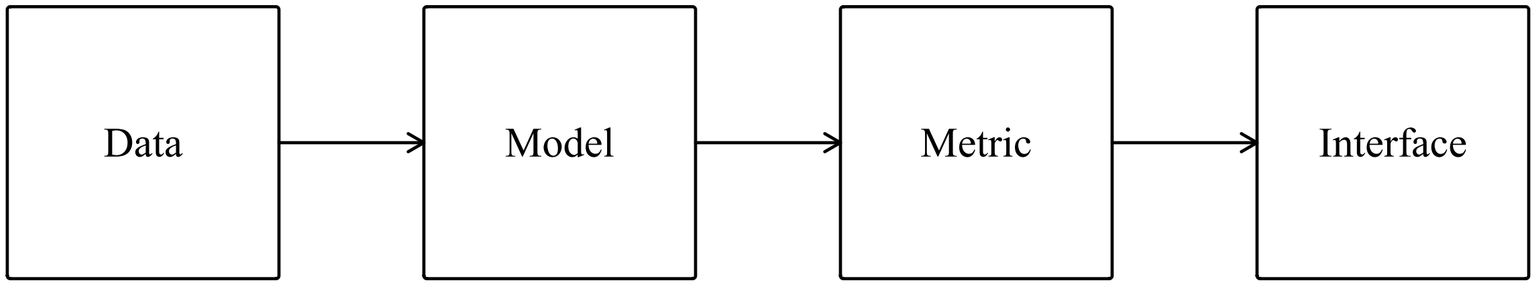
SOFTWARE FRAMEWORKS

- Produce applications in a standardized way
- Component interaction
- Reusable, consistent, comparable

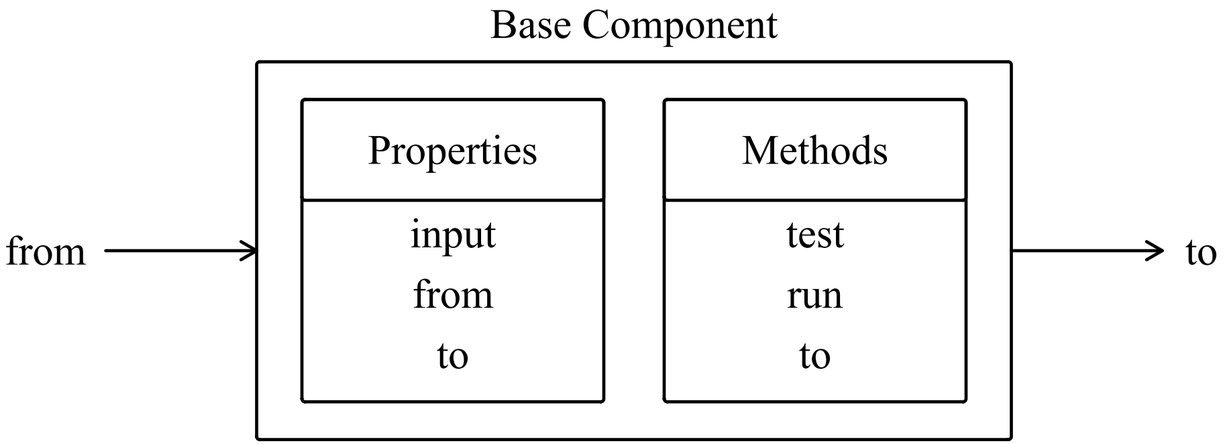
SOFTWARE DEVELOPMENT

- Object-oriented programming
- Test-driven approach
- Prototyping

SOFTWARE COMPONENTS



BASE COMPONENT



DATA COMPONENT

- Extract, store, process data
- Location, time, and numbers
- NLP methods for text

NLP WORD DISTRIBUTION

Given `a a b c c d`:

word	count
a	2
b	1
c	2
d	1

MODEL COMPONENT

- Statistics and machine learning
- **Supervised classification**: linear regression, naive bayes, decision trees
- **Clustering**: k-means, mixture models
- **Incremental Learning**: neural networks

METRIC COMPONENT

Metric	Description
Accuracy	Proportion of correct values
Precision	Correct values using model categories
Recall	Correct values using actual categories
F1 Score	Accuracy using precision and recall
RMSE	Scaled diff. of actual and model values

INTERFACE COMPONENT

- Map
- Dashboard

POTENTIAL EXPERIMENTS

1. Traffic Collision Forecasting
2. Crime Event Forecasting
3. Health Symptom Monitoring and Forecasting

TRAFFIC COLLISION AND CRIME EVENT FORECASTING

- Data: social media, open data, government
- Methods: NLP, supervised learning
- Outcomes: web app and models for forecasts

HEALTH SYMPTOM MONITORING AND FORECASTING

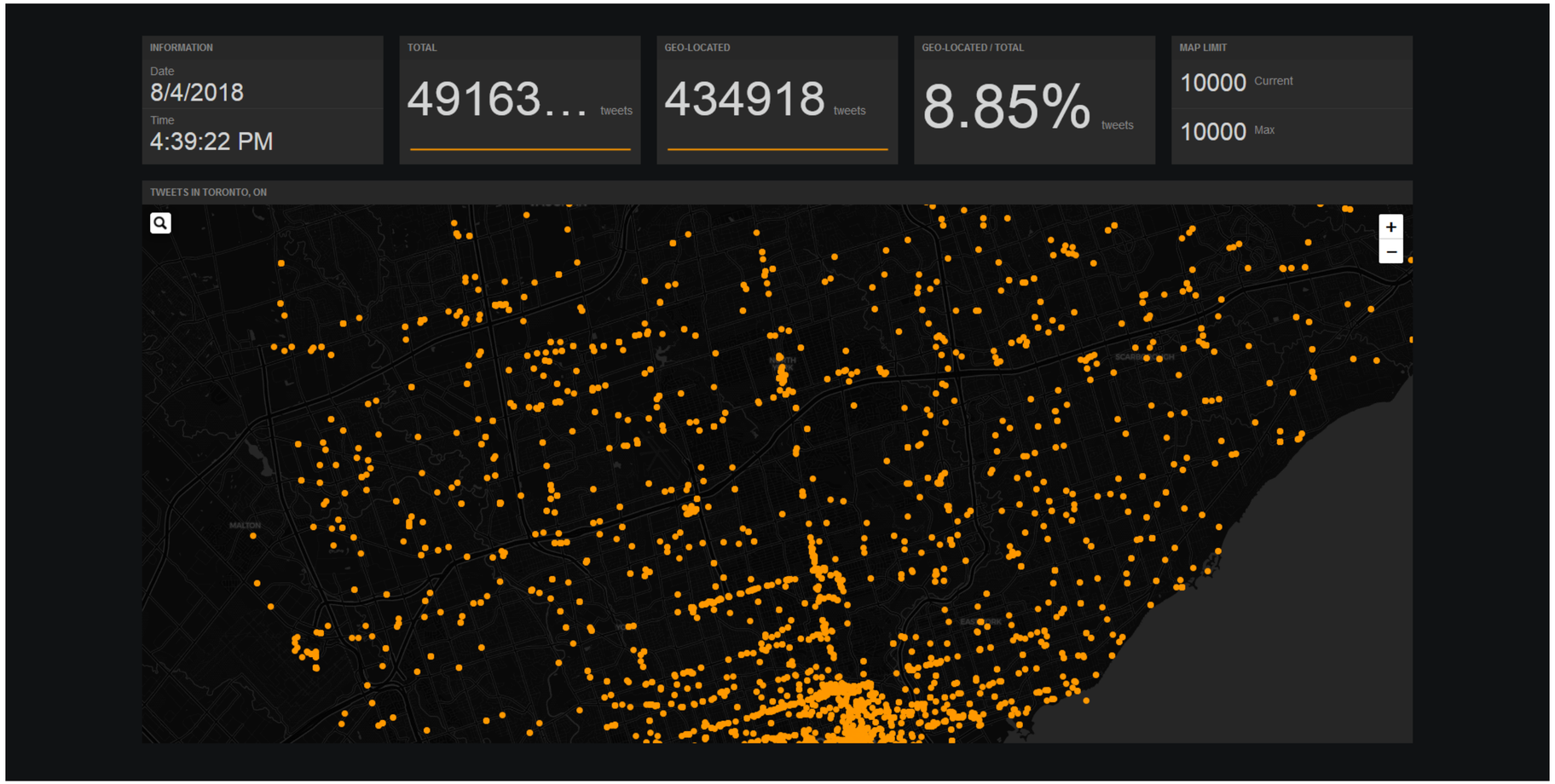
- Data: social media, open data, government
- Methods: NLP, supervised learning, clustering
- Outcomes: web app and models for monitoring and forecasts

PRELIMINARY RESULTS

SOFTWARE

- Developed Node.js packages for Twitter data, MongoDB, and PostgreSQL
- Explored potential software for framework components
- Hbase, GeoMesa, scikit-learn, Apache Kafka, D3.js

PROTOTYPE DASHBOARD



Ryerson University

CONCLUSION

- Context-aware GIS framework and architecture
- Experiments of context-aware GIS
- Methods for forecasting and monitoring security events

REFERENCES

- World Health Organization. (2015). Global status report on road safety 2015: Monitoring health for the sustainable. Retrieved June 26, 2018, from http://www.who.int/violence_injury_prevention/road_safety_status/2015/en/
- World Health Organization. (2018). World health statistics 2018: Monitoring health for the sdgs. Retrieved June 26, 2018, from http://www.who.int/gho/publications/world_health_statistics/2018/en/

- World Health Organization, United Nations Office on Drugs and Crime, & United Nations Development Programme. (2014). Global status report on violence prevention 2014. Retrieved June 26, 2018, from http://www.who.int/violence_injury_prevention/publications/violence/en/
- Dey, A. K. (2001). Understanding and using context. Personal and ubiquitous computing, 5(1), 4–7.