A Data Integration Architecture For Smart Cities

Murilo Borges Ribeiro
Kelly Rosa Braghetto

Department of Computer Science
Institute of Mathematics and Statistics
University of São Paulo

Introduction

- Smart cities
- Heterogeneity in data format and structure
- Low integration between data sources
- Data difficult to access for applications and non-technical users

Research Questions

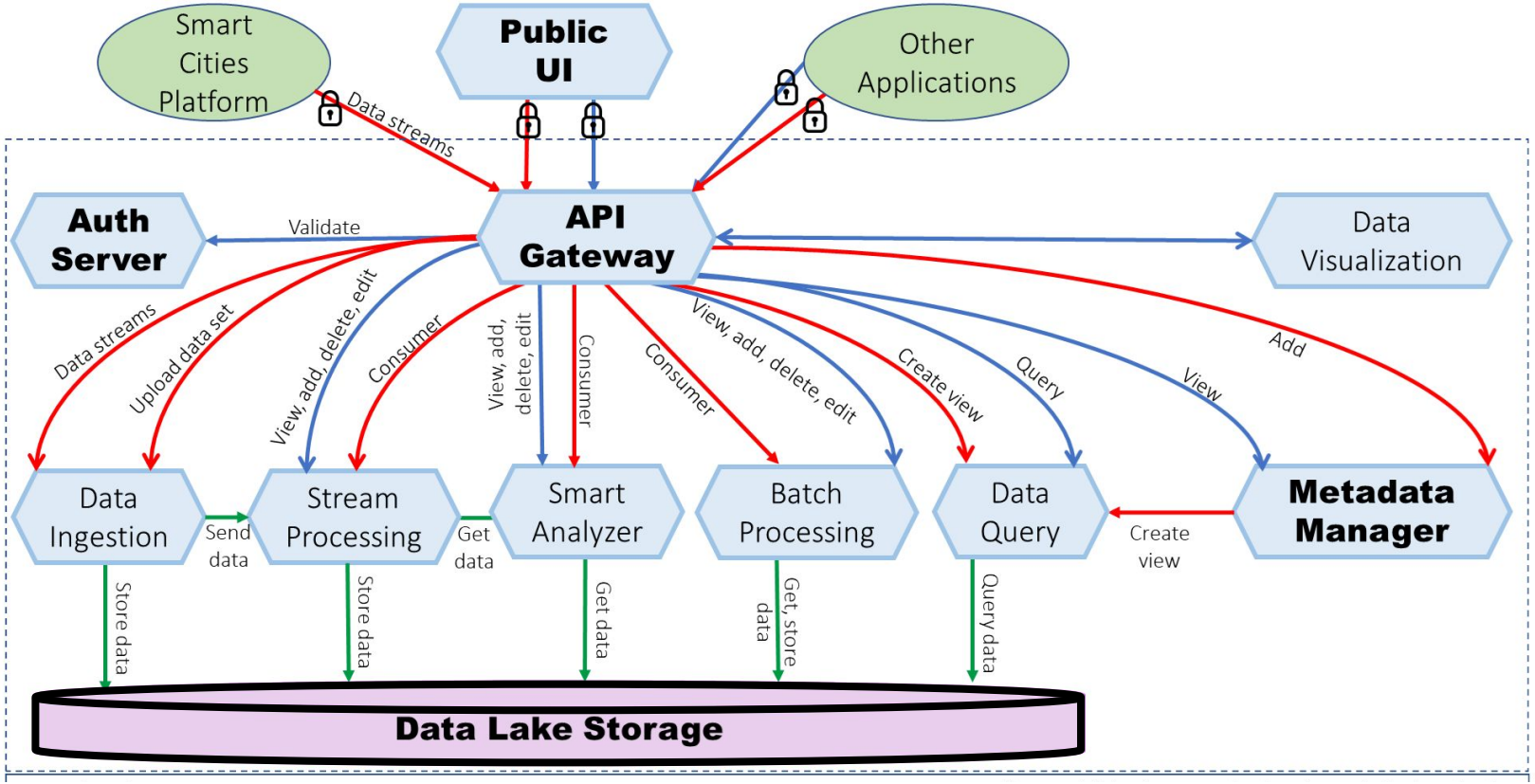
- What are the main challenges and issues identified by the researchers for data integration in smart cities?

- What are the functional and non-functional requirements of a software platform for data integration in smart cities?

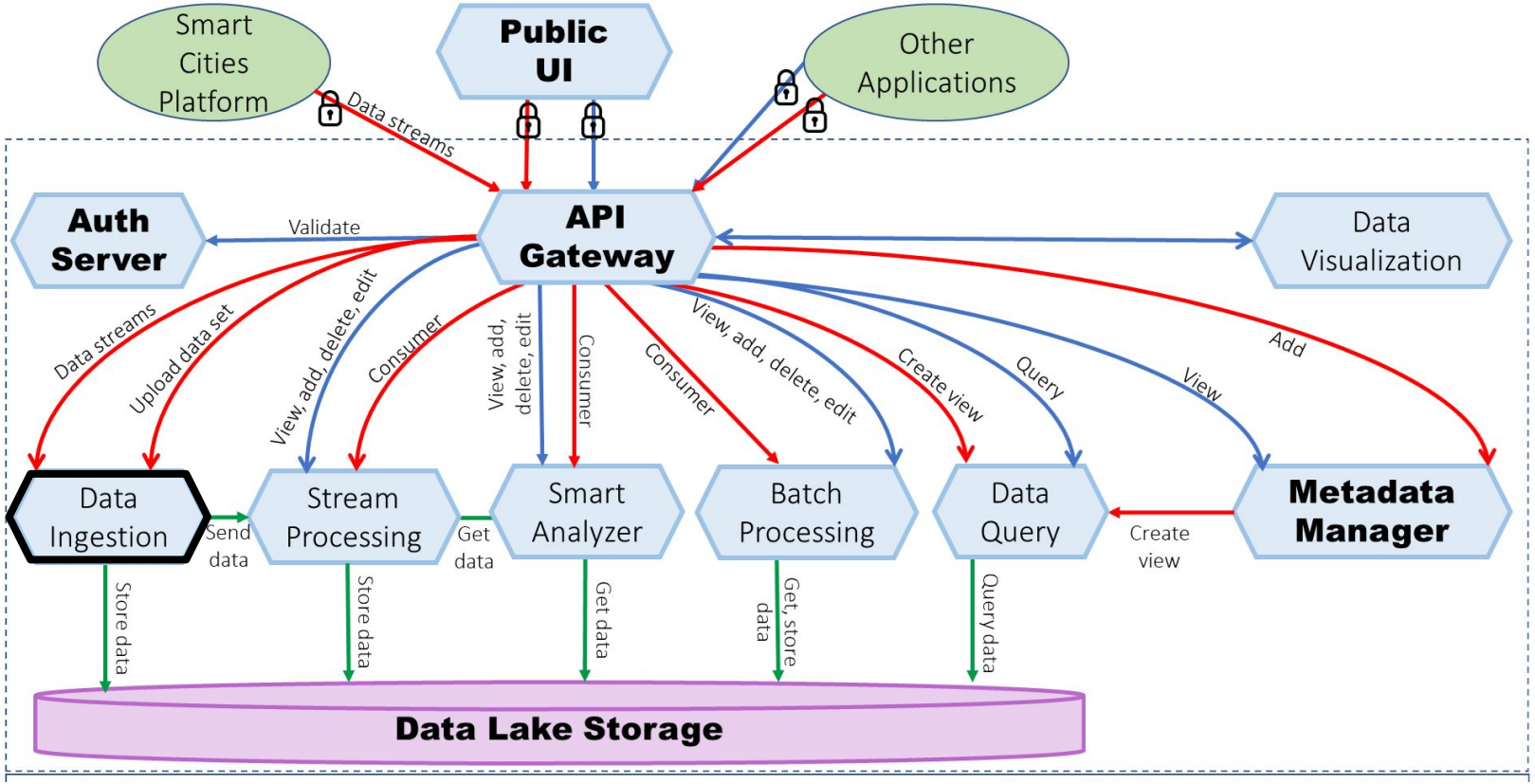
Requirements for a Data Integration Platform

	Psyllidis et al. 2015	Consoli et al. 2015	Cheng et al. 2015	Rathore et al. 2016	Hashem et al. 2016	Costa and Santos 2017	Mehmood et al. 2019
Ingestion	X		X	X	X	X	X
Metadata	X	X					X
Processing			X	X	X	X	X
Machine Learning				X	X		
Analysis and Visualization	X					X	X
External Access		X	X		X		
Scalability	X		X	X			X
Availability					X		X
Security and Privacy				X		X	

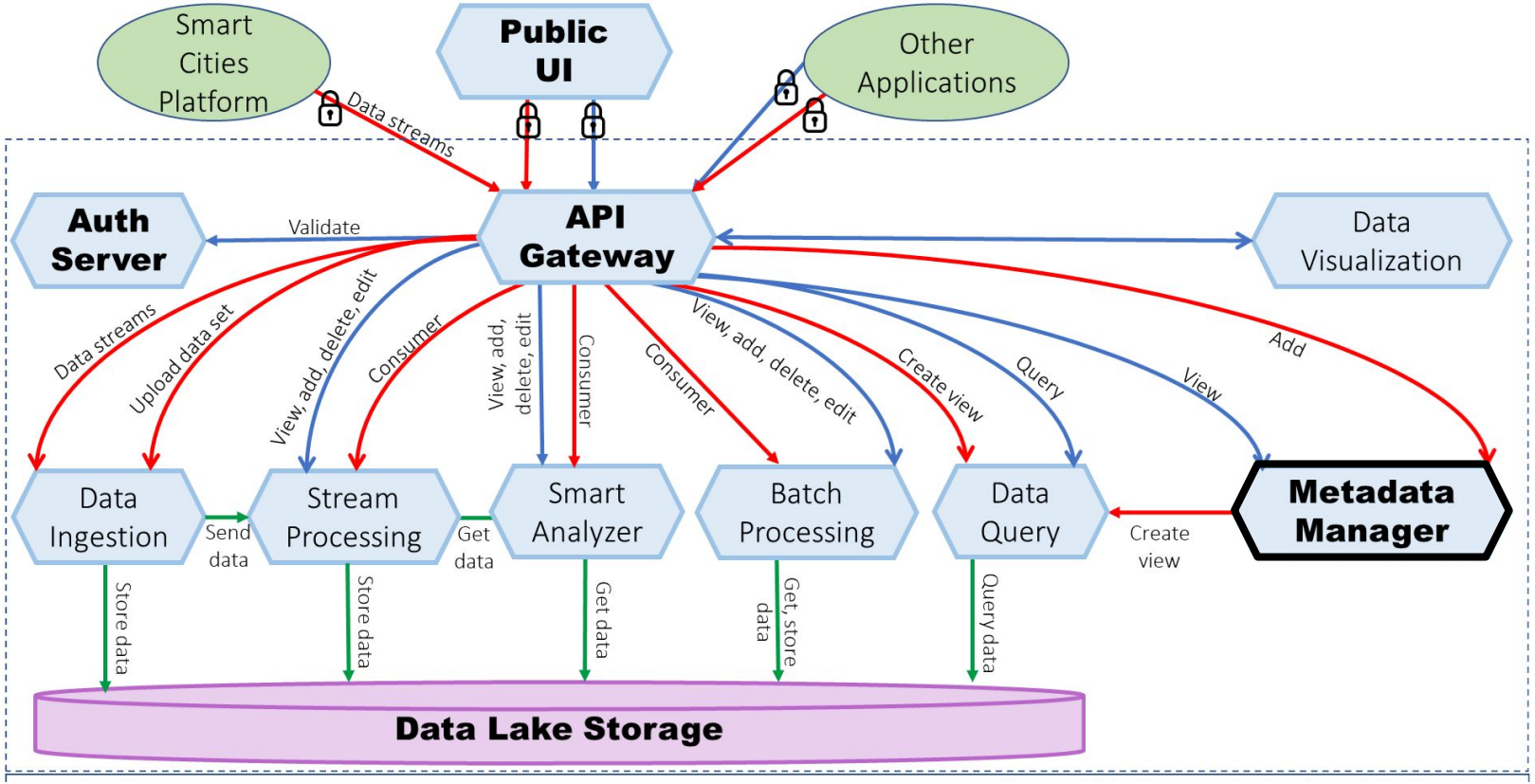
Proposed Architecture For Data Integration



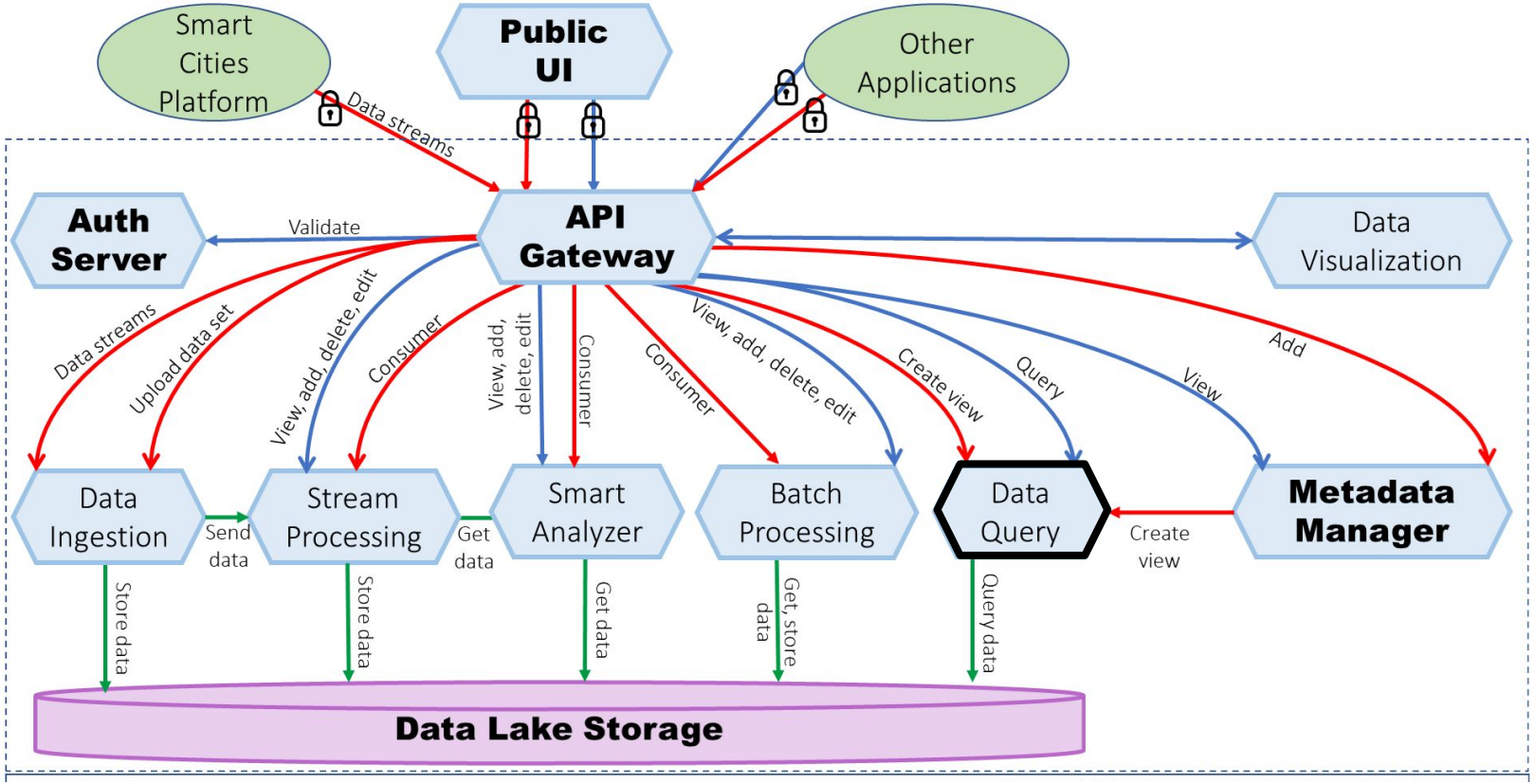
Proposed Architecture For Data Integration



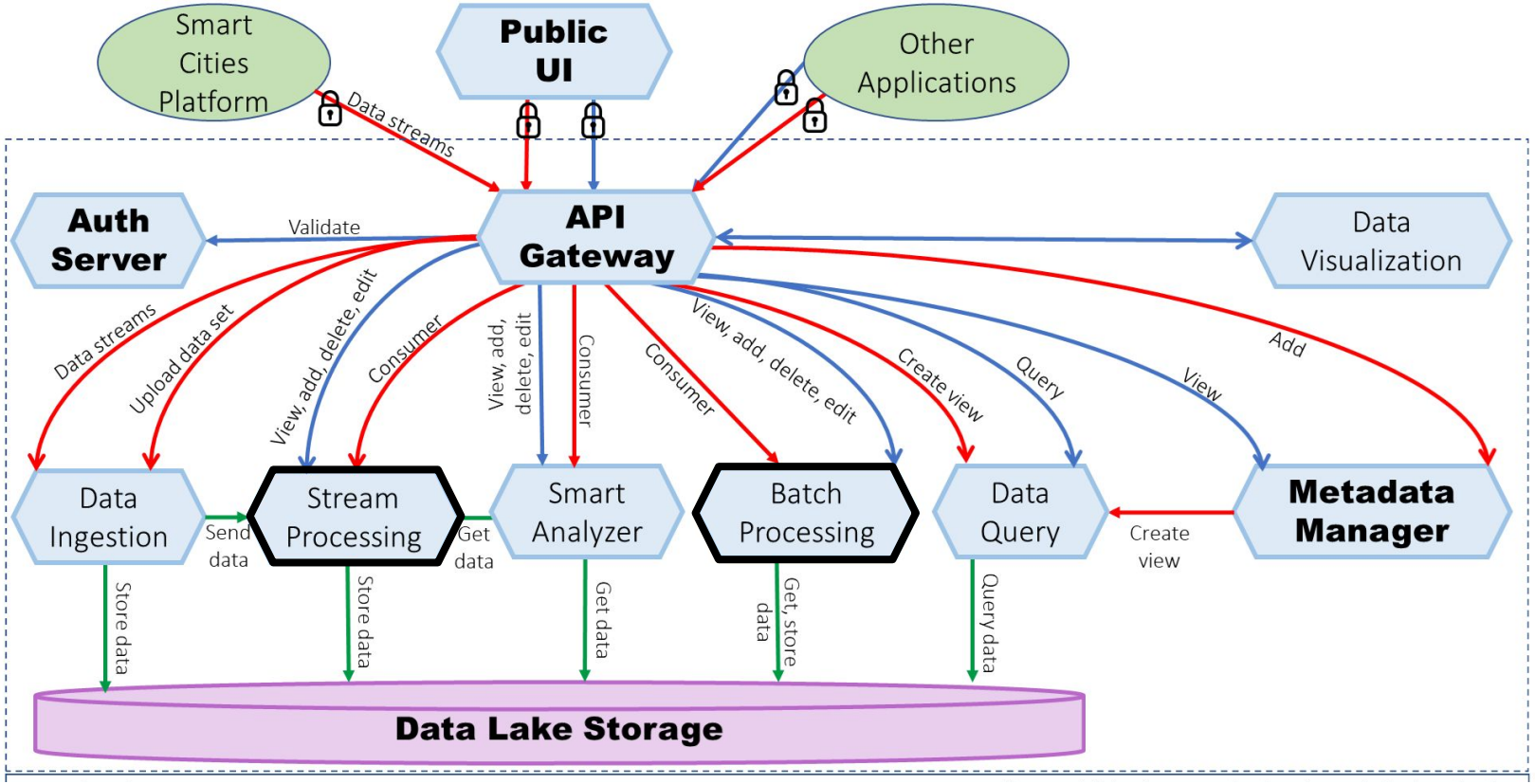
Proposed Architecture For Data Integration



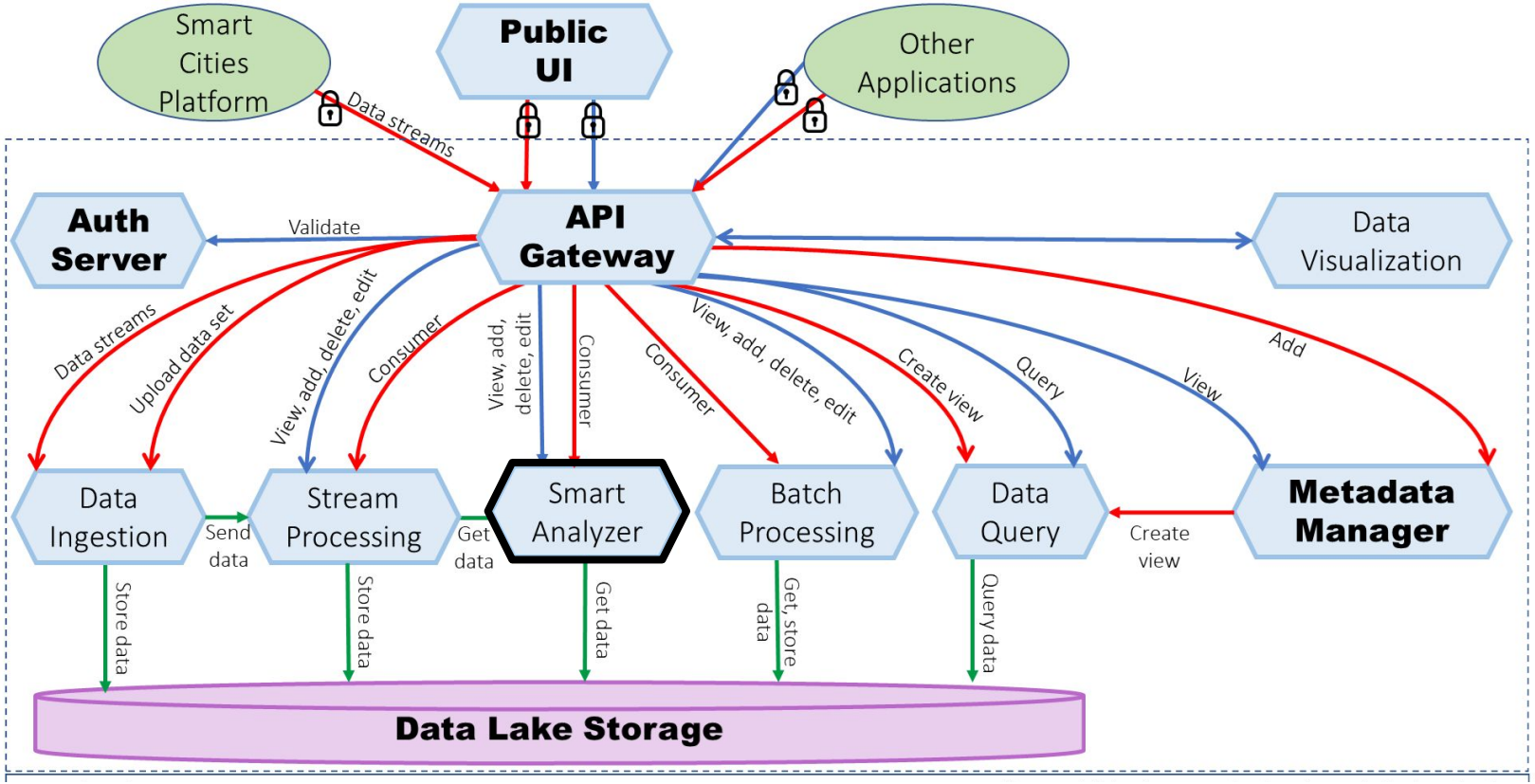
Proposed Architecture For Data Integration



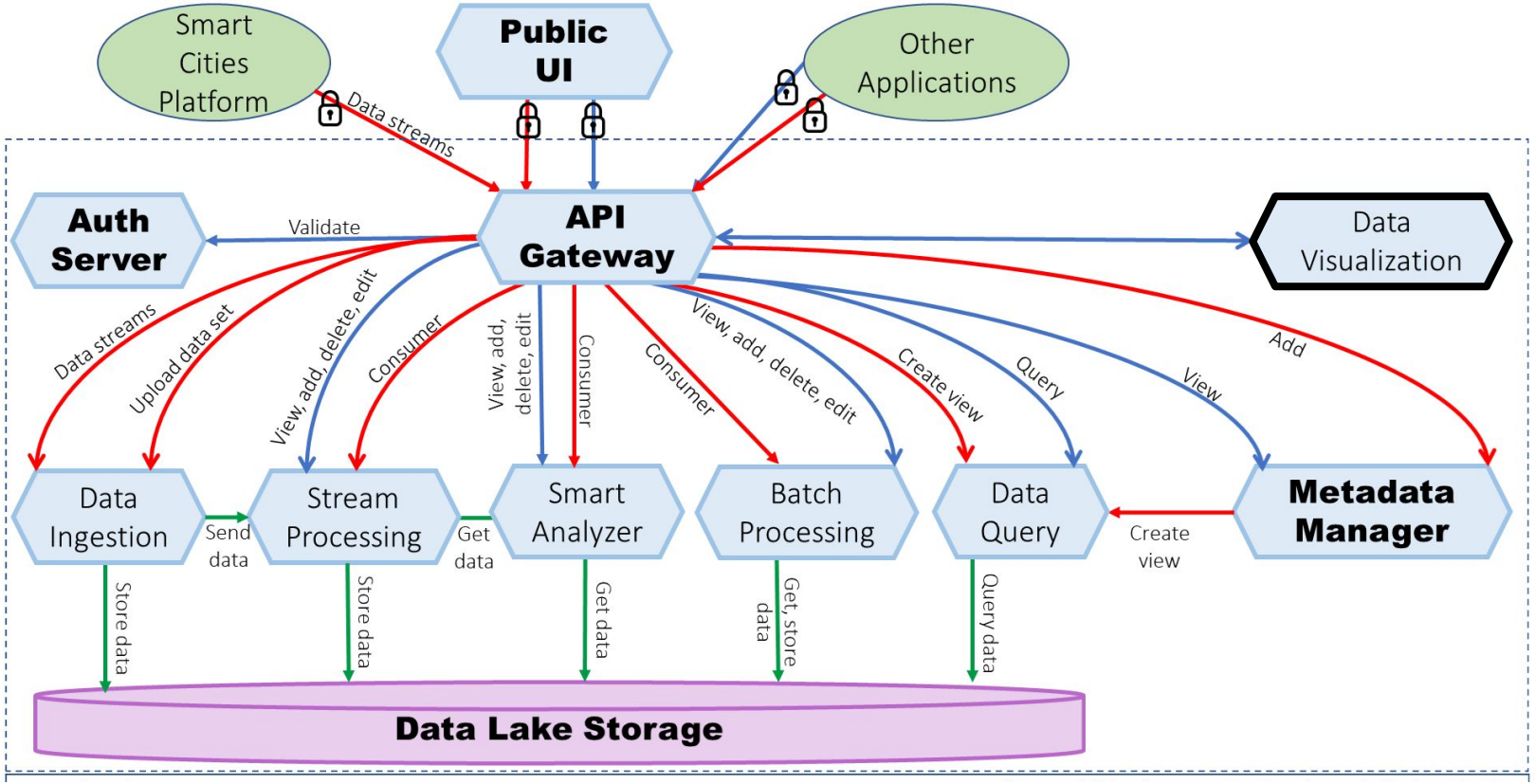
Proposed Architecture For Data Integration



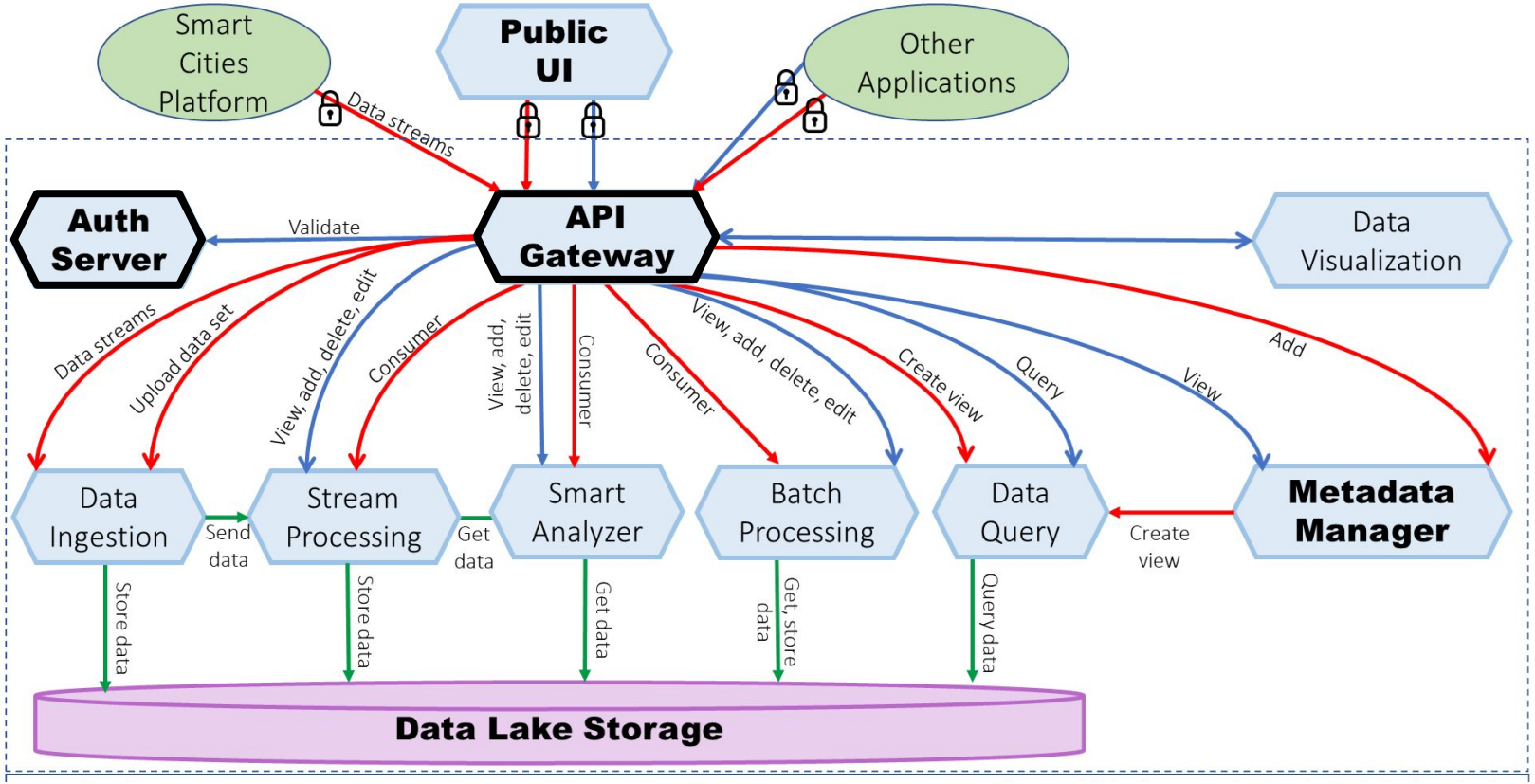
Proposed Architecture For Data Integration



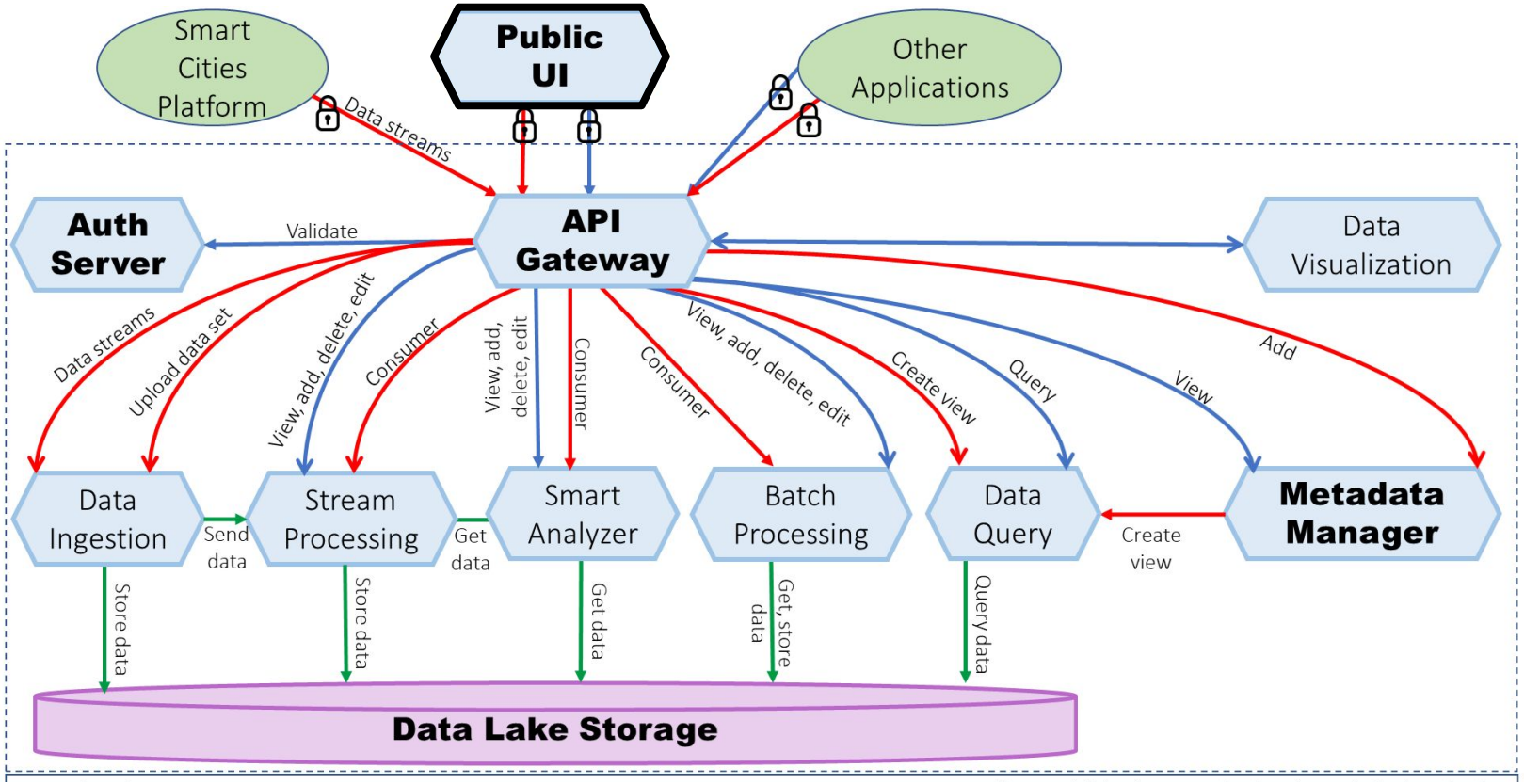
Proposed Architecture For Data Integration



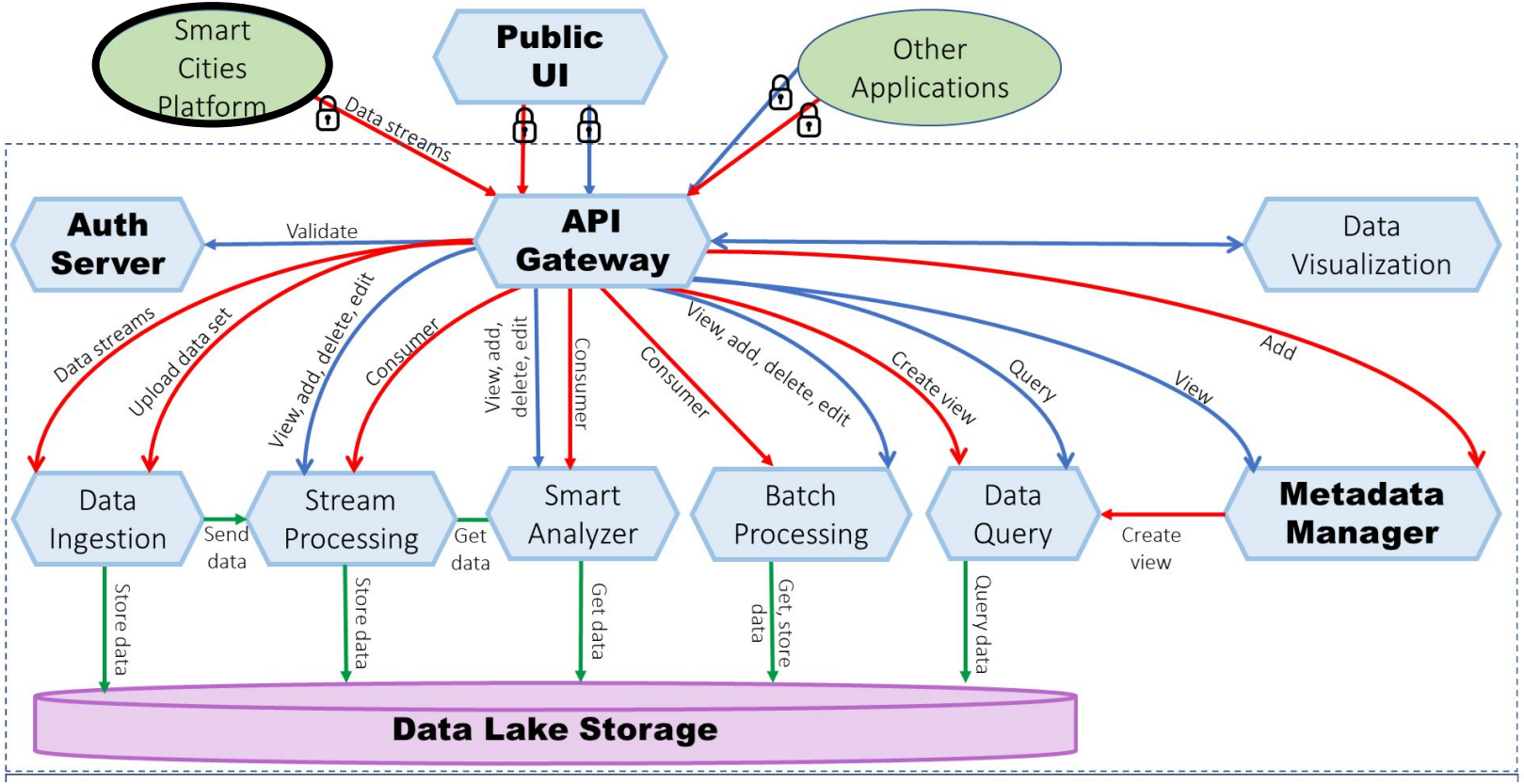
Proposed Architecture For Data Integration



Proposed Architecture For Data Integration



Proposed Architecture For Data Integration



Unique Features

- Single point of access to microservices by external applications
- Authentication and access authorization service
- Centralized interface for the services
- Creation of new data collections based on existing ones
- Compatibilization of data in collections that have suffered structural or semantic changes over time

Performance Evaluation

- Guideline based on the Cloud Evaluation Experiment Methodology (CEEM) [Li et al. 2013]
- Set of experiments designed to:
  - Measure the microservice's performance under normal workloads
  - Identify the maximum number of users and concurrent requests that a single instance can handle
  - Assess the capacity of the microservice to self-adjust to the current workload, increasing or decreasing the number of instances

Contributions

- A panorama of the requirements of data integration for smart cities and state-of-the-art solutions
- A Microservice architecture to integrate smart cities' data
- An (ongoing) implementation of the architecture on top of the InterSCity platform, using open-source tools
- A guideline to evaluate the performance of the architecture under both normal and above-normal workload conditions

References

- Cheng, B., Longo, S., Cirillo, F., Bauer, M., and Kovacs, E. (2015). Building a big data platform for smart cities: Experience and lessons from santander. In 2015 IEEE International Congress on Big Data, pages 592–599.
- Consoli, S., Mongiovic, M., Nuzzolese, A. G., Peroni, S., Presutti, V., Reforgiato Recupero, D., and Spampinato, D. (2015). A smart city data model based on semantics best practice and principles. In 2015WWW '15 Companion: Proceedings of the 24th International Conference on World Wide Web. Association for Computing Machinery.
- Costa, C. and Santos, M. Y. (2017). The suscity big data warehousing approach for smart cities. In Proceedings of the 21st International Database Engineering Applications Symposium, IDEAS 2017, page 264–273, New York, NY, USA. Association for Computing Machinery.
- Hashem, I. A. T., Chang, V., Anuar, N. B., Adewole, K., Yaqoob, I., Gani, A., Ahmed, E., and Chiroma, H. (2016). The role of big data in smart city. International Journal of Information Management, 36(5):748 – 758.
- Li, Z., OBrien, L., and Zhang, H. (2013). CEEM: A practical methodology for cloud services evaluation. In IEEE 9th World Congress on Services, pages 44–51.
- Mehmood, H., Gilman, E., Cortes, M., Kostakos, P., Byrne, A., Valta, K., Tekes, S., and Riekki, J. (2019). Implementing big data lake for heterogeneous data sources. In 2019 IEEE 35th International Conference on Data Engineering Workshops (ICDEW), pages 37–44.
- Psyllidis, A., Bozzon, A., Bocconi, S., and Titos Bolivar, C. (2015). A platform for urban analytics and semantic data integration in city planning. In Celani, G., Sperling, D. M., and Franco, J. M. S., editors, Computer-Aided Architectural Design Futures. The Next City - New Technologies and the Future of the Built Environment, pages 21–36, Berlin, Heidelberg. Springer Berlin Heidelberg.
- Rathore, M. M., Ahmad, A., Paul, A., and Rho, S. (2016). Urban planning and building smart cities based on the internet of things using big data analytics. Computer Networks, 101:63 – 80.

INCT InterSCity

USP

Thank you!!!

muriloribeiro@usp.br

This research is part of the INCT of the Future Internet for Smart Cities funded by CNPq proc. 465446/2014-0, Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001, FAPESP proc. 14/50937-1, and FAPESP proc. 15/24485-9.

CAPES

FAPESP

CNPq