No.	Big Data topics	Details
I	**Massive X Stream Platform**	
1	IoT data stream, storage and analysis system	Ref: (1) Kafka, MQTT: streaming (2) Spark: analysis (3) MinIO, HDFS: storage (4) Architecture: DataLake + Data Warehouse (Lamda, Kappa), Lakehouse (Delta), Data Mesh.
2	Multiple stream integration with temporal data: data imputation & prediction	Ref: (1) https://arxiv.org/pdf/2311.07344.pdf (2) https://dl.acm.org/doi/pdf/10.1145/3534678.3539078
3	Water quality monitoring, imputation and prediction	Ref: (1) https://dl.acm.org/doi/pdf/10.1145/3534678.3539078 (2)
4	Rainfall, wind direction (weather) prediction	Collect data from rainfall sensors at the area in the past Collect data from some weather forecast sources in the past Building a locality rainfall prediction model Using rainfall next week of weather forecast sources, the locality rainfall prediction model gives a rainfall next week prediction for the area
5	Flooding monitoring	Collect data from water level sensors Improve IoT data quality using Kalman filter (IoT sensors calibration) Predict water level (flooding) due to rainfall, high tide
6	Telco data analytics	Real data from Telco Multiple data streams by windows ~ hour/day/week/month/year Meta data layer by Bloom filter (BFMD) Propose applications using BFMD Similarity among windows can be achieved by using Locality Sensitive Hashing (LSH) applications by using HyperLogLog
7	Mobile tracking: security, check-in, …	Access points help to detect mobiles communicating with them over time: need equipments and solutions A list of mobiles at Access point (checkpoints) over time: persons visit each zone => Bloom filter help to build meta data by windows hour/day/month/year

		Finding a mobile has visited (recent) zones: finding kids/people/dogs/… Similarity among windows can be achieved by using Locality Sensitive Hashing (LSH) https://www.linkedin.com/pulse/how-track-mobile-phone-comprehensive-guide-monitoring-apps-sxt9c/
8	Equipment maintenace prediction	Collect data from equipment monitoring sensors Multiple streams from many equipments Building model improve IoT data quality & data imputation Building a maintenance prediction model
9	UAV monitoring	Lưu lại thông tin về đường bay: vị trí, cao độ, tốc độ, gia tốc,... Giám sát thiết bị bay có bay đúng lộ trình bay quy định
II	**AI & High Performance (Big) Data Analytics**	
1	Monitoring	Multiple data streams used in system monitoring Studying Probability Data Structures: Store less data (on-the-fly) but enough to calculate F(.) in monitoring/analytics Monitoring sensors at computing nodes: application monitoring HPC team solutions: publications Collection & analytics monitoring architecture QoS service prediction: online Machine learning
2	Scheduling	Resource management tools: PBSpro, Slurm, LSF,… HPC team solutions: publications GPU-node resources High Performance Scheduling algorithms
3	Portal	Full functions for users & admin GenAI in Portal
4	Distributed ML/DL, Federated Learning, Swarm Learning => Applications	Federated Learning: member contributions => fast convergence
5	Distributed ML/DL applications: Telco, water meters	Distributed ML/DL applications: Telco, water meters Collection & Analytics for electrical meters Monitoring electric vehicle battery Swarm Learning
6	MLOps	
III	**Smart Transportation**	

1	Bus Call Button	Ref: (1) Xem kết quả về giải pháp phần cứng để phát triển giải pháp phần mềm
2	Smart Bus Box	Hardware of Smart Bus Box Communication + Streaming
3	Smart Bus Stop	Hardware of Smart Bus Stop: environment
6	Fastest-Path algorithms	Shortest traveling time: point-to-point, all-to-all Shortest traveling time prediction: in a time in future Ref: (1) Luận văn đại học - Nguyễn Duy Khang (đang triển khai)
5	Bus GPS data collection	Mobile App for bus-GPS collection: Kalman filter, auto collection (there is a tool to generate bus-GPS location based on a route in map)
5	Smart Bus Collection & Analytics	Kafka & Spark used in Smart Bus system
6	RL cho ứng dụng Traffic Light	Ref: (1) Luận văn cao học - Châu Quốc Đạt (đang triển khai)

IV FarmLab

1	FarmLab collection & analytics	A testbed area at a region with all sensors and equipments in order to collect data as well as run real applications controlled by smart solutions of Smart Village. Then results and experiences are shared to all people in the region. It is a low- Data collection solution with multiple streams Improve IoT data quality, imputation Monitoring and secure IoT sensors, devices
2	Digital Twin for FarmLab	Digital Twin Share data to customer/users
3	Virtual Farm Lab	We need a virtual farm lab because of (1) a real farm lab cannot support many scenarios used to evaluate algorithms, (2) No a real farm lab set up in a new region, (3) simulate a a Virtual Farm Lab is built to simulate a real farm lab. There are tasks as following: (1) Simulate real sensors like temperature, air humidity, brightness, water environment, wind, rainfall,... (2) Soil moisture (3) Fertilizer and nutrient content in soil ... => (1) Define input data & requirements, (2) Developing algorithms to generate sensor data, (3) Algorithms for soil moisture based on irrigation before => a Virtual Farm Lab has inputs and then generate (data stream) outputs; it is an interactive module

3	RL algorithm for Agriculture (Irrigation)	
4	Smart Village & Irrigation mobile app	
5	AI-based Smart Village Portal	
V	**AI applications**	
1	X-Count Things	Ref: (1) Construction scaffolding project
2	Detecting damage on asphalt pavement	Detecting damage on asphalt pavement based on deep learning neural network Road damage monitoring vehicle: Improve models + GPS & map location + mobile apps… Ref: (1) Luận văn Cao học - Phan Văn Hưng
4	Super resolution camera	Ref: (1) Publication of Đức Phương
7	Improving personalized recommendation models through metadata enrichment	Video/Image -> Meta data -> Reccomendation Systems: Providers & Customers (like TikTok) Applications: (1) Smart Village to introduce products Ref: (1) Luận văn Cao học - Nguyễn Công Thành
8	Early purchase prediction	Ref: (1) Luận văn Cao học - Diệp Thế Toàn
11	Session-based recommendation system in Fashion	Ref: (1) Luận văn Cao học - Nguyễn Quỳnh Anh Phương