



## **DE JUISTE STREAMING ENGINE FOR THE JOB**

### ***Context:***

Voor vele bedrijven is de batch-geïntendeerde architectuur van Big Data verwerken, waarbij data gecapteerd worden in grote, schaalbare databases, waarna de data later verwerkt wordt, gewoon te traag. Moderne 'fast data' architecturen zijn geëvolueerd naar stream-gebaseerde architecturen, waarbij data meteen verwerkt wordt terwijl die verzameld wordt. Dit leidt tot een competitief voordeel voor deze bedrijven.

SEr zijn echter vele stream processing tools, en de vraag is dan uiteraard welke te gebruiken? Daarvoor kan het helpen om volgende factoren af te wegen:

- Latency: welke snelheid van verwerking is nodig?
- Volume: hoeveel volume moet kunnen verwerkt worden?
- Integratie met andere tools: welke en hoe?
- Processing: welke soorten data moet er verwerkt worden? En dient dit eerder in bulk of per individueel event te gebeuren?

### ***Nood:***

In deze opdracht willen we een aantal criteria onderzoeken die belangrijk zijn voor het kiezen van specifieke technologieën, en willen we ook voorbeelden uitwerken van 3 streaming tools: Reactive Streams (beschikbaar in Java sedert JDK 9), Kafka Streams en Apache Spark, die allen op een ander niveau toelaten om te werken met een continue stream van data.

### ***Taak:***

- Naast een situering van het streaming landschap, en een overzicht van de 3 bovenvermelde technologieën, is het ook de bedoeling om concreet werkende implementaties te bouwen. Een voorstel hier is om een tool te bouwen die een stream van inzichten geeft in cryptomunten:  
markt inzichten & visualisatie: trade volume, order books, price changes, market cap, ...
- anomaly detectie: detecteren wanneer een cryptomunt een significante prijs en/of volume aanpassing meemaakt, om zo een potentieel interessant inkoop moment te signaliseren
- alerts: intelligente notificaties die getriggered worden gebaseerd op bepaalde cryptocurrency markt condities

### ***Resultaat:***

Aantonen welke tools en welke niveaus van abstractie bruikbaar zijn voor een set van Streaming vraagstukken.

### ***Contactpersoon:***

Sam Waegeman – [sam.waegeman@xt-i.com](mailto:sam.waegeman@xt-i.com)  
XTi NV (<https://www.xt-i.com>)  
Guldensporenpark i  
9820 Merelbeke