

ปรีณา ทองสกุล

Java Spark Developer

✉ java-spark-developer@example.co.th

☎ +66 80 555 2026

📍 อุบลราชธานี

👤 สรุป

Data engineer ประสบการณ์ 6 ปีสร้าง Java Spark pipeline ขนาดใหญ่ ทำ data quality gate, replay และ cost control ให้ทีมวิเคราะห์ใช้งานข้อมูลได้ภายในวันเดียว

🏢 ประสบการณ์การทำงาน

- **Java Spark Developer**
Mekong Data Platform 11/2019 - ปัจจุบัน
 - ปรับ Spark job รองรับ event 620 ล้านรายการต่อสัปดาห์ ลด runtime จาก 9 เหลือ 2.5 ชั่วโมง
 - เพิ่ม data quality rule 42 ข้อ ลด schema drift จาก 5.4% เหลือ 0.7%
 - แบ่ง Hive table ตามมิติธุรกิจ 3 ด้าน ลด EMR cost 33%
 - ทำ Airflow rerun พร้อม owner ให้ job 9 ตัว เพิ่ม complete delivery เป็น 99.4%
- **Java Spark Developer**
SignalCraft Analytics 8/2014 - 10/2019
 - เขียน Java transformation ให้ source system 18 แห่ง ตัดงานกลางคืนด้วยมือ 24 ชั่วโมงต่อสัปดาห์
 - ทำ schema versioning ให้ table 31 ชุด ลด analyst table break 58%
 - ทำ dependency map 67 รายการให้ analyst 4 คน ลดเวลาหาสาเหตุจาก 1 วันเหลือ 90 นาที
 - ทดสอบ backfill บน staging 2.1 พันล้านแถว 6 รอบและยืนยันผลไม่ต่างกัน

<> โครงการ

Customer Feature Lake pipeline รายวันที่แปลง event จากเว็บร้านค้า และ support เป็น feature ที่ตรวจสอบย้อนกลับได้

Java Spark
Hive
Airflow

🎓 การศึกษา

วิทยาศาสตรมหาบัณฑิต สาขา
วิทยาการข้อมูล
มหาวิทยาลัยอุบลราชธานี
10/2013 - 6/2017

☆ ทักษะ

ทักษะทางเทคนิค

Java, Apache Spark, Hadoop, Hive,
Airflow, AWS EMR

🗣 ภาษา

ภาษาไทย **เจ้าของภาษา**
ภาษาอังกฤษ **คล่องแคล่ว**



การรับรอง Cloudera Spark and
Hadoop Developer
องค์กรรับรองวิชาชีพ
2025