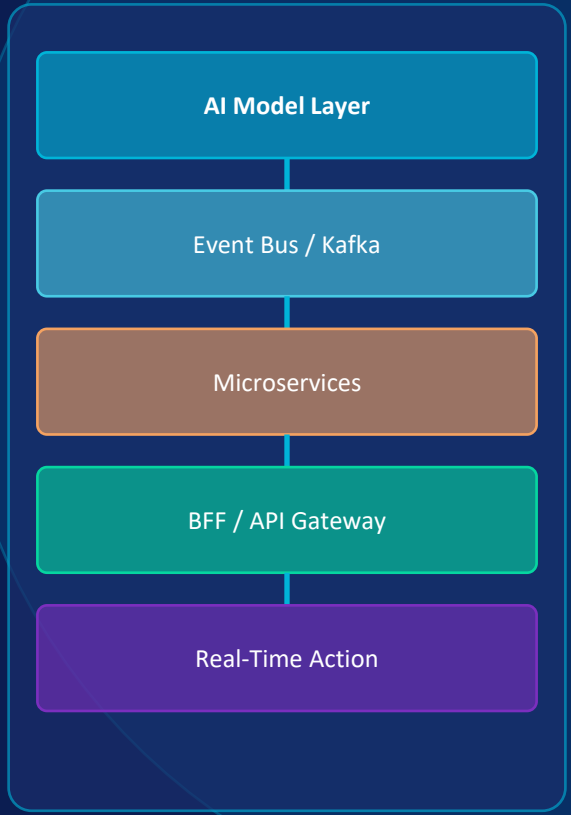
TELECOMUNICACIONES × INTELIGENCIA ARTIFICIAL

La IA no falla por el modelo. Falla por la Plataforma

AI doesn't fail because of the model.
It fails because of the platform.



@LuisRai linkedin.com/in/lraigosov

@LuisRai · linkedin.com/in/lraigosov

¿Por qué la IA no actúa a tiempo en telecomunicaciones?

Why doesn't AI act in time in telecommunications?

Predicción sin acción
Prediction without action

ES: El modelo predice churn correctamente, pero la respuesta llega días después por silos de datos y procesos lentos.
EN: The model correctly predicts churn, but the response arrives days later due to data silos and slow processes.

Arquitectura monolítica
Monolithic architecture

ES: Los sistemas heredados no soportan actualizaciones de modelos independientes ni respuestas orientadas a eventos.
EN: Legacy systems don't support independent model updates or event-driven responses.

Un canal para todo
One channel for all

ES: La misma lógica de IA responde a una app móvil, un call center y un portal web, sin adaptación por canal.
EN: The same AI logic responds to a mobile app, call center, and web portal without channel adaptation.

@LuisRai

5 dominios que exigen IA en tiempo real

5 domains that demand real-time AI

Churn Predictivo / Predictive Churn

< 200ms

ES: Detectar riesgo de fuga antes de que el cliente cancele / EN: Detect churn risk before cancellation

Atención de Reclamos / Claims & Support

Tiempo real / Real-time

ES: IA que triaga, enruta y responde reclamos por canal / EN: AI that triages, routes & resolves claims by channel

Autoservicio IA / AI Self-service

< 100ms

ES: Chatbots y portales inteligentes sin fricción / EN: Frictionless intelligent chatbots & portals

Anomalías de Red / Network Anomalies

< 50ms

ES: Detección y remediación automática de fallas / EN: Auto-detect and remediate network failures

Facturación IA / AI Billing

Stream

ES: Detección de errores y fraude en facturación / EN: Error detection & fraud in billing flows

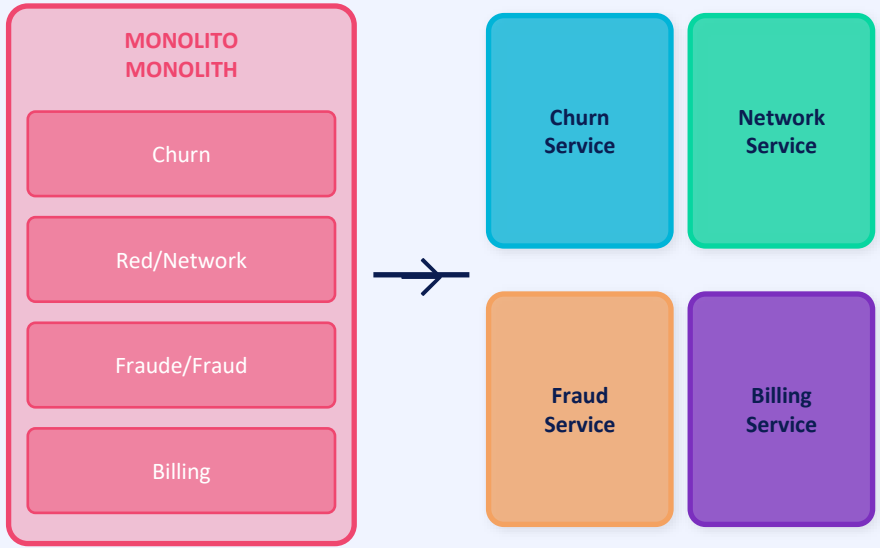
@LuisRai

Separar dominios de IA

Decouple AI domains

1. ES: Cada dominio (churn, red, fraude) despliega su modelo de forma independiente
   EN: Each domain (churn, network, fraud) deploys its model independently
2. ES: Actualizaciones sin detener la operación completa
   EN: Updates without stopping the full operation
3. ES: Escalado horizontal por carga de cada servicio
   EN: Horizontal scaling per service load
4. ES: Fault isolation: un modelo fallido no cae toda la plataforma
   EN: Fault isolation: one failed model doesn't bring down the platform

Monolito → Microservicios

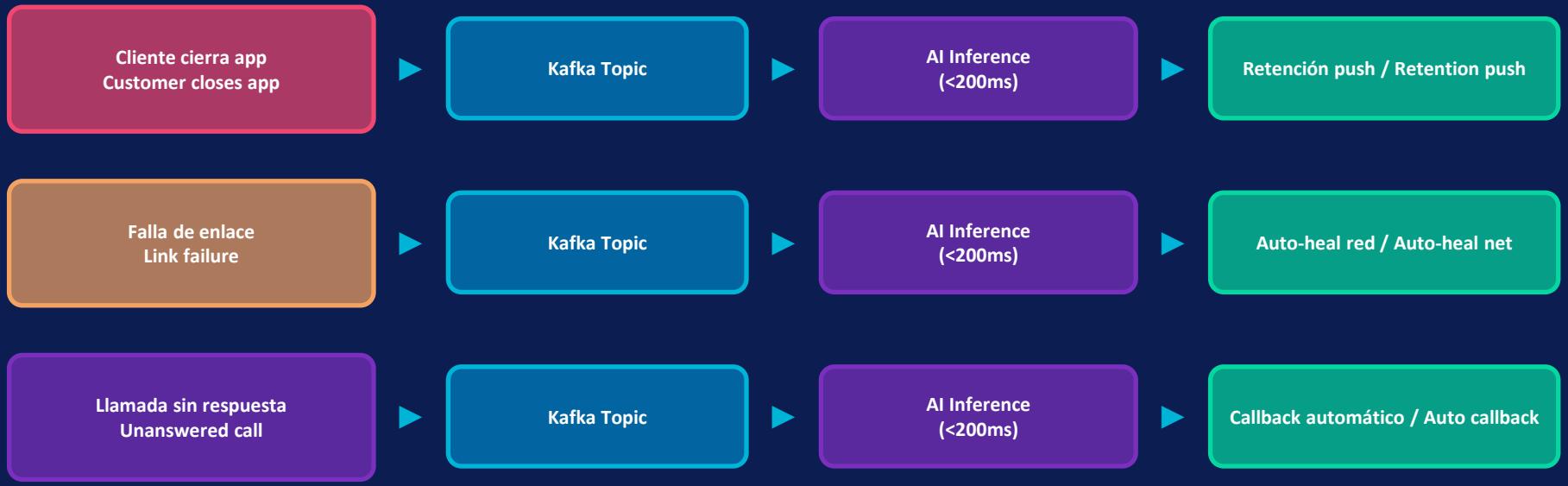


Ref: Martin Fowler – Microservices Architecture (martinfowler.com) | IBM – Microservices in Practice

@LuisRai

Cuando un evento ocurre, la IA actúa — sin polling, sin espera.

When an event occurs, AI acts — no polling, no waiting.

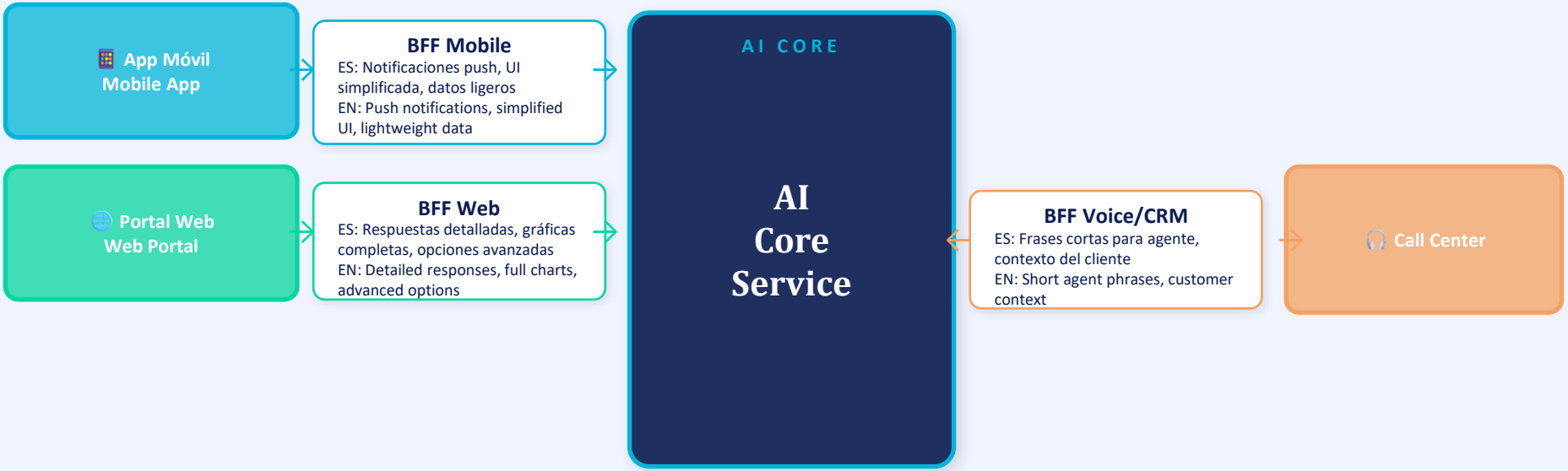


💡 Con Kafka/Pulsar: latencia < 50ms extremo-a-extremo — imposible con arquitectura batch o polling | With Kafka/Pulsar: latency < 50ms end-to-end — impossible with batch or polling architecture

@LuisRai

La misma IA, experiencias distintas por canal

Same AI, different experiences per channel



ES: El BFF desacopla la presentación de los datos de la lógica de negocio, permitiendo que la IA se adapte al canal sin duplicar modelos | EN: BFF decouples data presentation from business logic, letting AI adapt per channel without duplicating models

Ref: Sam Newman – Building Microservices (O'Reilly) | Phil Calçado – Pattern: BFF (philcalcado.com/2015)

@LuisRai

Microservicios + Event-Driven + BFF en Telecomunicaciones

Microservices + Event-Driven + BFF in Telecommunications

FUENTES / SOURCES: OSS · BSS · CRM · Network Probes · CDRs · Billing · App Events

EVENT BUS: Apache Kafka / Apache Pulsar — Real-time streaming, exactly-once delivery, topic partitioning per domain

MICROSERVICES LAYER

- Churn MS
- Network Anomalies MS
- Fraud Detection MS
- Claims AI MS
- Billing AI MS

BFF LAYER (Backend for Frontend)

BFF Mobile	BFF Web	BFF CRM/Voice	BFF Partner API
Mobile App	Web Portal	Call Center	Partners

@LuisRai

La adopción de IA en telecomunicaciones crece —> pero la plataforma es el cuello de botella

AI adoption in telecommunications is growing — but the platform is the bottleneck

78%

de operadoras de telecomunicaciones pilotean IA
telecommunications operators piloting AI

NVIDIA Telecommunications AI Report 2024

3x

más rápido con event-driven vs. procesamiento batch
faster with event-driven vs batch

Confluent State of Data Streaming 2024

67%

del tiempo de IA perdido en latencia de integración
of AI time lost in integration latency

IBM Global AI Adoption Index 2024

<50ms

latencia requerida para remediación de red en tiempo real
latency required for real-time network remediation

3GPP TS 28.554 (Network Mgmt)

AI Adoption in Telecommunications (%) — Source: NVIDIA, IBM, Ericsson

El debate no es
"el modelo no funciona"
Es:
"la plataforma no esta diseñada para operar IA en tiempo real"

The debate isn't
"the model doesn't work"
It is:
"the platform isn't designed to operate AI in real time"

01 Microservicios: desacopla dominios de IA / Decouple AI domains

02 Event-Driven: actúa en < 200ms sobre cada evento / Act in < 200ms on every event

03 BFF: adapta la IA al canal, no al revés / Adapt AI to channel, not the reverse

@LuisRai linkedin.com/in/lraigosov