



[www.pipelinepub.com](http://www.pipelinepub.com)

Volume 22, Issue 11

# From Dumb Pipe to Intelligence Platform - The Telco Transformation Playbook

By: [Aleksandr Moskotin](#)

Every operator has run the same arithmetic and arrived at the same uncomfortable answer. Traffic is compounding. Revenue is not. PwC's latest outlook puts global telecom service revenue on a path from US\$1.15 trillion in 2024 to roughly [\\$1.32 trillion in 2029, a CAGR of about 2.8 percent](#), while global monthly mobile ARPU is expected to tick *down* to \$6.20 in 2029 from \$6.32 in 2024. Usage soars; willingness to pay flattens. Two decades of capital expenditure have produced a business whose unit economics drift gently downward.



The industry's standard response has been to treat this as a pricing problem – bundle harder, tier more finely, chase adjacent services. It is more usefully understood as an inference problem. The network already knows an enormous amount about demand, intent, and timing. Very little of that knowledge is currently converted into anything an operator can invoice. What has changed in the last few years is not the data; it is that the analytics and machine learning stack needed to act on it has become operationally realistic.

## The asset is the inference, not the data.

It is worth being precise about what the asset actually is, because the loose phrase “monetizing subscriber data” has caused a decade of confusion and no small amount of regulatory grief.

Raw network data has two awkward properties: its market value is close to zero, and its liability profile is close to unbounded. Nobody buys a table of location pings. What buyers want is an answer – who is likely to churn this month, which household is in-market for a device upgrade, which neighborhood justifies a new store. That answer is a modeled output, not an extract.

Two properties make network data unusually good raw material for producing those answers. It is consented, through terms of service the subscriber has actually agreed to, and it is network-verified, which means it cannot be spoofed or fabricated in the way that cookie- and SDK-derived signals routinely are. In a digital advertising market with a chronic fraud problem, verified-by-infrastructure is

a genuine differentiator rather than a marketing adjective.

The operators furthest along have understood this as a shift from selling data to selling analysis. Vodafone's Analytics business packages anonymized location intelligence for retailers, event organizers and city planners. Orange Business went broader still, building out [big data consulting and AI-based solutions](#) for enterprise clients in finance, healthcare and logistics. In both cases, the deliverable is an insight product with a service wrapper – something that can be versioned, supported and renewed.

## What operators are actually modeling

Strip away the vocabulary and the models that earn their keep fall into a small number of families.

The first is retention. Churn propensity is the oldest use of operator analytics and still the most reliably profitable, because the operator is acting on its own P&L: a subscriber saved is revenue that does not have to be re-acquired. What has improved is not the concept but the resolution – daily scoring against behavioral drift rather than quarterly scoring against demographics, and next-best-action models that choose the intervention as well as the target.

The second is commercial timing. Device upgrade cycles, plan migration, roaming, and add-on propensity are all questions about *when* rather than *who*, and the network's high-frequency signal is unusually good at *when*. A model that identifies the two-week window in which a subscriber is receptive is worth considerably more than one that identifies a broad segment with no timing attached to it.

The third family sits outside the operator's own P&L, and it is where third-party revenue lives: audience qualification for advertisers, footfall and catchment analysis for retailers, mobility patterns for transport and municipal planning, and – in markets with thin credit bureaus – alternative scoring inputs that widen financial inclusion. These are analytics products sold to someone else, which is precisely why they demand the governance discipline discussed below.

## The bottleneck is plumbing, not modeling.

The strategic case is easy. The engineering is not, and this is where most data monetization programs quietly stall.

A modern operator's signal is scattered across Call Detail Records, OSS/BSS platforms, deep packet inspection systems, location infrastructure, and subscriber management – [streams that are structurally complex and, in most estates, thoroughly siloed](#). Each was built for a different purpose, on a different cadence, under a different owner. Joining them into something a model can learn from is a multi-year, capital-intensive undertaking with very little to demo in year one.

The trap is to solve this as a storage problem. Consolidating everything into a lake produces a large, expensive, and largely inert archive. Machine learning needs features, not files: consistent entity resolution across systems, event timestamps you can trust to the minute, a governed definition of a subscriber attribute that means the same thing in the churn model and in the campaign engine, and the ability to compute all of it fresh enough to matter. A propensity score that arrives 48 hours late is not an insight; it is a report.

This is why sequencing matters more than ambition. Operators who pick two or three commercially specific questions and build only the pipeline those questions require tend to be in production while the enterprise-wide data platform is still in design review. Analytics maturity, not model novelty, is the

constraint that binds.

## Keeping the model inside the perimeter

The architectural principle that makes any of this durable is simple to state and hard to compromise on: raw personal data must not leave the operator's secure perimeter.

The pattern that satisfies this is an in-network identity layer. Machine learning runs inside the operator's own environment, and what crosses the boundary is a pseudonymous token representing audience attributes – a segment, a propensity, an eligibility flag – with no re-identifiable payload attached. The operator can participate in the programmatic ecosystem while PII stays home.

This is not merely a compliance posture. It also happens to be the version advertisers should prefer, because a token minted against network-verified behavior is harder to fake than one assembled from third-party inference. Clean-room architectures generalize the same logic to brand collaborations: two parties measure a shared audience without either handing over its underlying records.

Consent, in this architecture, has to be a running system rather than a signature collected once. Preferences change, regulations change, and a subscriber who agreed to service messaging has not necessarily agreed to third-party advertising. The operators that hold up under audit treat consent state as a live attribute sitting in the same feature store as everything else – queryable, versioned, and enforced at the moment the audience is assembled rather than reviewed after the fact.

## Paying for outcomes, not for rows

The final step is the one that separates a data science capability from a revenue line, and the commercial model does most of the work.

Selling data creates regulatory and reputational exposure that scales with volume. Selling *outcomes* does not. The models that hold up are performance-based: the operator delivers targeted campaigns through its own channels – SMS, RCS, IVR, in-app – priced on cost-per-acquisition or revenue share. Incentives align with the advertiser's result, the operator retains custody of the underlying data, and there is no dataset in circulation to be misused later.

Operators running this model have reported ARPU uplift in the region of 5 percent. In a market where ARPU is otherwise drifting downward, a five percent move on the base is material. The compounding element is the feedback loop rather than the initial model: every campaign labels the training data for the next one, an advantage that a third-party audience vendor structurally cannot replicate.

What makes such a number defensible, though, is measurement discipline rather than model sophistication. Performance pricing only works if both sides agree on what was caused rather than merely correlated, which means holdout groups, incrementality testing, and an agreed attribution window written into the contract before the first campaign runs. Programs that skip that step tend to find their reported uplift disputed at renewal.

## What the playbook does not promise

Honesty about the failure modes is worth more than another success anecdote.

Subscriber attention is a finite, shared resource, and an operator that over-monetizes its own messaging channels will damage a customer relationship worth considerably more than the incremental campaign. Attribution disputes are common and get harder as channels multiply. Consent frameworks continue to diverge by jurisdiction, which means an approach that is compliant

in one market may need re-architecting in the next. And none of this is a product an operator can buy in a quarter – even with partner infrastructure, it is a capability built with real organizational cost, spanning legal, network, marketing and data teams that do not habitually share a roadmap.

But the direction of travel is not seriously in doubt. The data moving through these networks is worth more than the connectivity fees charged to carry it. The operators who work that out first will not be selling data. They will be selling what they have inferred from it.

Not for distribution or reproduction.