APRIL · 2026

European airline apps

State of UX 2026

BY ED ROBERTS

Contents

Introduction

Something is off with airline app ratings.

Go to Google Play today and browse the apps of Europe's ten biggest airlines. Every app sits between 4.2 and 4.8 stars out of five. Apple's App Store shows a similar picture, averaging 4.7 across the same ten carriers. Strong numbers and the kind of ratings most product teams would be proud of.

But extract the reviews that have been written this year on those same listings and the average is 2.3.

That gap is the story of this report. Across four sources and nearly 24,000 reviews, we looked at how Europe's ten largest airlines are serving the digital side of travel in 2026. The short answer is: not very well. The app store scores say otherwise, but the headline ratings are masking the real picture. Look closer and the story is less flattering, more interesting, and in some places quite new.

At WAS we do this kind of research regularly. Most recently we looked at UK EV charging apps, where we found an industry falling short of what a newly mainstream driver population expected. Now we have turned our attention to aviation: same research approach, and as it turns out some remarkably similar patterns. The analysis covers reviews from 2023 to April 2026, with the sharpest findings concentrated in the most recent window.

Approach and scope

As with all our UX research, our goal is not to diagnose any individual airline but to describe the state of the industry as a whole. We are interested in the experience of a hypothetical "median traveller". The person who downloads the app because the airline told them to and expects it to work. The budget-flyer who has no choice but to use the app to avoid a €55 airport check-in fee.

The business traveller whose loyalty account keeps logging them out.

We analysed 23,608 reviews after cleaning, drawn from four sources:

- **Apple App Store** reviews from the UK store, pulled via the public RSS feed
- **Google Play** reviews from the UK store, via the open-source google-play-scraper library
- **Reddit** posts and comments mentioning each airline, via Reddit's public JSON endpoints
- **Trustpilot** reviews, sampled manually at around 40 per airline (used as a supporting source only)

Of the full dataset, 11,164 reviews are "digital-relevant", meaning they mention the app, the website, booking, check-in, login, payment or similar. This filtered set is the backbone of the analysis. Online reviews as a category are polarised and, as we'll show, partially inflated by bot activity; Trustpilot reviews are used here for illustration and "colour" only, not for quantitative claims.

The ten airlines we covered are Ryanair, easyJet, Turkish Airlines, Lufthansa, Wizz Air, British Airways, Air France, Vueling, Pegasus and KLM, in rank order by passengers carried (source: CAPA Centre for Aviation, 2024 data). With the exception of our discussion of the headline ratings table, this report will rarely name airlines individually, though we will often refer to budget and legacy carrier cohorts.

Headline ratings: what the public sees

This is the shop window. On 24 April 2026, the ten airlines scored as follows across the three public review channels we measured.

British Airways on Google Play jumps out, sitting at 1.3, three full stars below every other airline in the sample, but largely these are ratings most businesses would be happy with, which is exactly the problem.

Public headline ratings, 24 April 2026

Ratings and review counts as shown on each platform's public page.

Airline	Apple ★	Apple reviews	Google Play ★	Google Play reviews	Trustpilot ★	Trustpilot reviews
Ryanair	4.5	722k	4.2	382k	1.3	30k
easyJet	4.8	875k	4.2	313k	3.6	61k
Turkish Airlines	4.6	30k	4.6	145k	1.4	6k
Lufthansa	4.7	22k	4.6	96k	1.4	14k
Wizz Air	4.7	143k	4.6	702k	1.3	18k
British Airways	4.5	871k	1.3	55k	1.4	15k
Air France	4.7	17k	4.5	164k	1.3	2k
Vueling	4.6	27k	4.7	269k	1.3	10k
Pegasus	4.7	20k	4.5	162k	1.5	4k
KLM	4.7	44k	4.6	120k	1.4	8k

Sources: apps.apple.com/gb, play.google.com UK, trustpilot.com. Data captured between 09:00 and 09:20 on 24 April 2026.

The gap between the shop window and the street

The street tells a different story. Focusing on the most recent reviews per airline, we filtered to digital-relevant reviews and computed the average rating of what users are actually writing today.

Recent trends are much less positive

For each airline, the recent digital-relevant mean sits roughly 1 to 3 stars below the public headline. Across the industry, the public iOS average is 4.65. The recent digital-relevant iOS average is 2.00. That is a gap of 2.65 stars. On Google Play the equivalent figures are 4.18 and 2.39, a gap of 1.79 stars.

British Airways is the only airline whose Google Play headline does not sit far above recent user sentiment. Its 1.3 headline and 1.2 recent mean effectively agree. Whilst we can only speculate as to why, this shows recent reviews on BA are much more in keeping with the long term average than the rest of the cohort.

For every other airline on Play, the historical average is acting as a cushion: years of older, better scores outnumbering the recent critical ones. Ryanair's 382,000 Play ratings, for example, would barely move even against a heavily critical recent sample. And so, the headline rating does not accurately reflect recent user experience.

> *"I genuinely do not understand how this app gets good reviews. It never loads my bookings, always says there's an error and can't load... it's just useless."*
>
> — Apple App Store review, 2★, UK, March 2026

HEADLINE VS RECENT RATING

The gap, airline by airline

Each airline's public headline rating against the mean of its recent digital-relevant reviews, by platform. Δ is the change from headline to recent.

Apple App Store · Google Play · Δ = recent mean − public headline

Airline	Apple headline	Apple recent	Apple Δ	Play headline	Play recent	Play Δ
Ryanair	4.5	1.7	−2.8	4.2	1.6	−2.6
easyJet	4.8	1.3	−3.5	4.2	1.4	−2.8
Turkish Airlines	4.6	2.4	−2.2	4.6	3.4	−1.2
Lufthansa	4.7	2.5	−2.2	4.6	2.8	−1.8
Wizz Air	4.7	2.0	−2.7	4.6	2.1	−2.5
British Airways	4.5	1.4	−3.1	1.3	1.2	−0.1
Air France	4.7	2.1	−2.6	4.5	2.7	−1.8
Vueling	4.6	2.1	−2.5	4.7	2.6	−2.1
Pegasus	4.7	2.9	−1.8	4.5	2.4	−2.1
KLM	4.7	2.4	−2.3	4.6	2.9	−1.7

Sources: apps.apple.com/gb, play.google.com UK, trustpilot.com. Data captured between 09:00 and 09:20 on 24 April 2026.

A second cushion: non-organic reviews

There's a second force pushing up the review average, and it's a surprising one. Earlier WAS research in October 2025 found that a large share of Google Play reviews for European airlines showed characteristics consistent with non-organic content: short, generic, overwhelmingly four- and five-star, Android-only, and clustered in time in ways suggesting coordinated activity.

When we re-applied the same detection rules to this year's dataset, the pattern had persisted. **23% of Google Play reviews in our sample** are flagged on these criteria. The flagged reviews average 4.5 stars. The unflagged ones average 2.7.

Here's what flagged reviews look like:

"Good" · "Excellent" · "Easy to use" · "great"

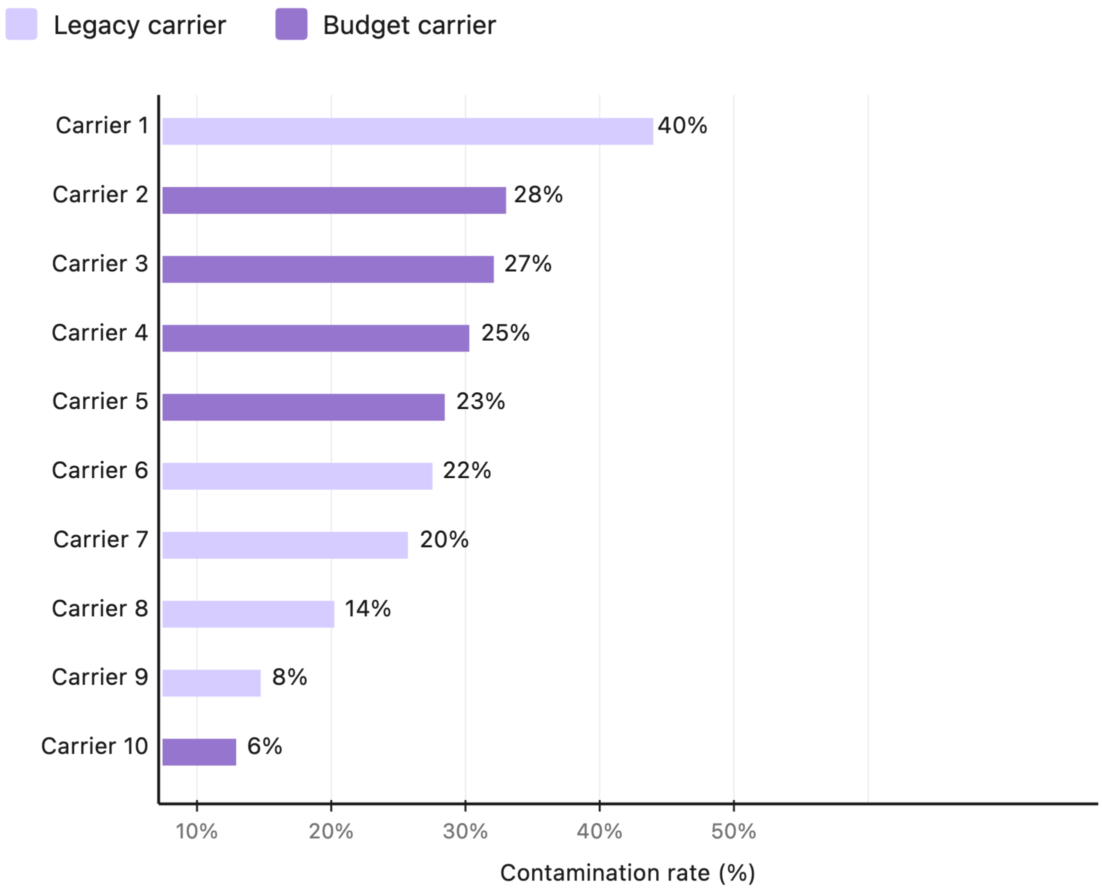
Google Play · 5★ · Android · 2026

Individually, each of these is an inoffensive one-word review. In aggregate, they form a distinct pattern, because there are thousands of them: almost entirely on Android, almost entirely top-rated, concentrated in specific time periods.

NON-ORGANIC REVIEW DETECTION

Google Play review contamination, by carrier

Percentage of scraped Play reviews matching the bot-detection pattern.



Sources: apps.apple.com/gb, play.google.com UK, trustpilot.com. Data captured between 09:00 and 09:20 on 24 April 2026.

The contamination rate varies by airline. One legacy carrier's Play sample is as high as 40%, while several budget carriers sit between 24% and 28%. Two carriers, one from each cohort, sit below 10%. If flagged reviews are removed, Google Play means drop by between 0.1 and 0.6 stars depending on airline.

The origin of these reviews is not something we can prove from public data alone, and we are not asserting that the airlines themselves commission them. If you'd like to see what we did, we can share the detection methodology on request. We have applied it consistently across all ten carriers in the sample.

What the data does show is a persistent layer of Google Play reviews that do not behave like organic user content. They are short, top-rated, Android-only and time-clustered, and this materially inflates the public ratings of nine of the ten airlines we measured.

Responsibility for the integrity of the review system sits with Google, but the effect on what the public sees is straightforward: most of the industry's Google Play headlines are cushioned both by historical averaging and by a persistent layer of reviews that do not appear to reflect organic user experience. The one exception in our sample is low in part because neither cushion is doing much work for it.

So the big takeaway: **the 4-to-5 star ratings shoppers see when choosing or trusting an airline app become 2-to-3 when you focus on recent, genuine human experience.**

Two trends worth highlighting

Before we dig into what users are trying to do and where it breaks down, two broader patterns emerged from the data that deserve spotlighting quickly.

1. The tone of reviews is increasingly angry

> *"We tried doing online check in and it didn't work. We then proceeded to do it at the airport and we had to pay an extra €45 per person. Complete and total SCAM."*
>
> — Trustpilot review, UK, April 2026

Our dataset shows strong anger language ("disgraceful", "shambles", "worst", "unacceptable", "scam", "joke") rising 75% between 2023–24 and 2025–26. Nearly 500 explicit uses of that vocabulary in the recent window, up from below 200 in the earlier one.

This is not industry-specific. Mainstream consumer reviewing has become more polarised across categories.

In our recent EV charging review (UK public EV charging apps: state of UX 2025, March 2026), we observed a shift from patient frustration to principled anger, captured in what we called 'grumpy cliché' phrases: *"it's 2025"*, *"it's not rocket science"*, *"how hard can it be"*. Echoes of that vocabulary appear in this dataset too. *"In this day and age"* appears 59 times. Variants of *"how hard can it be"* and *"should just work"* appear in the hundreds.

> *"Why in this day and age do I have to type in my email address manually. Why scroll through buggy drop downs. Why are you trying to guilt trip and trick me into buying more. This is app old fashioned [sic]."*
>
> — Apple App Store review, 1★, UK, December 2024

> *"Can't change the email address associated with my account. How hard can it be, technically?"*
>
> — Google Play review, 3★, UK, April 2026

These phrases are not unique to airlines. They are the vocabulary of a mainstream consumer who is comparing their airline app to their food delivery app, their banking app, their streaming app.

Users don't just compare within a market, they compare to the best and most frequently-used software on their phone. The bar of quality is rising across categories, and customers will call out the apps that haven't kept up.

2. Users are increasingly well equipped and well informed

AirTags and Tile trackers come up often in the 2025–26 data: passengers are using them to track their luggage and in some cases to argue with customer service when their account of a bag's location contradicts the airline's.

> *"They lost our luggage on the way. When it was located and flown to our final destination, they didn't deliver it to our hotel (they are obliged to do so by law), and the communication was useless and almost non-existing. I could just see it on my AirTag."*
>
> — Trustpilot review, UK, April 2026

> *"If it wasn't for the airtags I would be going frantic. Appalling service."*
>
> — Reddit, airline-specific subreddit, March 2026

The same pattern shows up in the language of consumer rights. Citations of EU261, the European Union's passenger-rights regulation, rose from a single mention in our 2023–24 window to 56 in 2025–26, outpacing the 1.66x growth in review volume between the two windows. Consumers are getting better at naming the regulation that applies to their case, sometimes via claims-management services, sometimes via word-of-mouth, sometimes via the forum culture we describe later. The absolute volume is still modest but the direction of travel is unmistakable. Five years ago the airline's knowledge of a delayed bag was the authoritative source. Today the passenger's own phone might be, and the passenger increasingly knows which regulation applies to whatever the airline does next.

What airline app users are actually trying to do

The discussion so far has been about ratings and tone. Next we turn to what users are trying to do: the jobs that take them to the app in the first place, and the moments where the journey breaks.

For this we use the Jobs To Be Done (JTBD) framework, which asks what outcome the user is trying to achieve rather than what solution is in their hand. For each job we score importance and satisfaction, and we read the gap between them as an indicator of how under-served the job currently is. We extracted the jobs that were emergent from the reviews themselves rather than imposing a pre-conceived set on the data. We identified one overarching job, eight small jobs within it, and one emotional job that sits across everything.

Jobs by importance, satisfaction and opportunity

#	Statement	Importance	Satisfaction	Frequency	Opportunity
000	Organise and complete my air travel with minimal friction (overall)	10	2	5	18
004	Check in and have my boarding pass ready when needed	10	3	4	17
003	Pay for travel and optional extras	9	2	3	16
001	Find and book a flight that fits my needs	9	3	3	15
009	Feel in control of my travel experience (emotional)	9	3	8	15
007	Get my problem resolved when something goes wrong	8	2	2	14
006	Stay informed about my flight's status and any changes	8	3	7	13
002	Change or cancel my booking when plans shift	7	2	2	12
005	Select and confirm my seating preferences	6	2	3	10
008	Access my account and manage my loyalty benefits	5	1	7	9

Importance, Satisfaction and Frequency scored on a 1–10 scale. Opportunity = Importance + max(Importance – Satisfaction, 0). Scores of 15 or above are critical under-served; 10 to 14.9 are under-served; below 10 are served right or over-served. Sources: apps.apple.com/gb, play.google.com UK, trustpilot.com. Data captured between 09:00 and 09:20 on 24 April 2026.

Scoring the opportunity

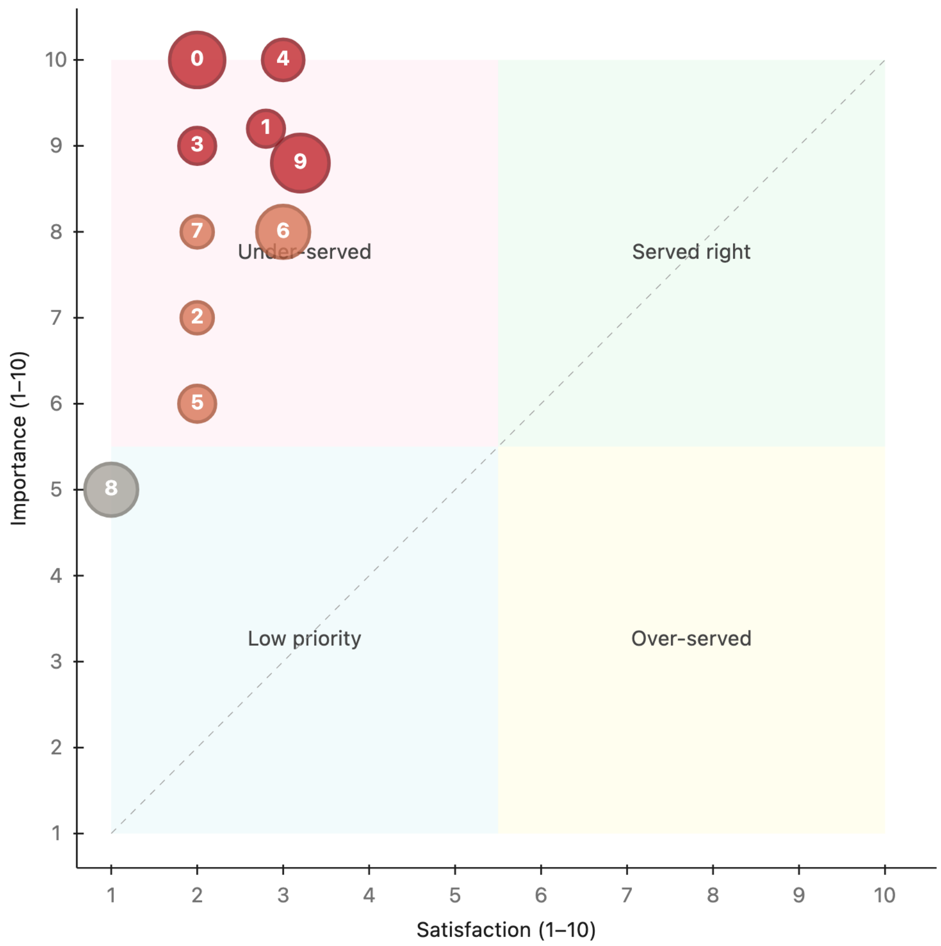
Using Ulwick's Outcome-Driven Innovation scoring we can rank the jobs by "opportunity" (a very strategist-facing term), but essentially a measure of how well served each job is by the airline apps against the user priority. Scores of 15 or above are critical under-served; 10 to 14.9 are under-served; below 10 are served right or over-served.

The jobs most under-served, highest opportunity for improvement, are the umbrella job of **completing the trip (18)**, **check-in and boarding pass (17)**, and **paying (16)**. Unfortunately for the airline apps, eight of ten jobs sit at or above the under-served threshold. However the real story is revealed when we visualise the jobs on a simple 2×2 matrix.

Suddenly we see nearly every single job sits in the danger zone, and importance exceeds satisfaction in every case. "Loyalty" appears outside the pink quadrant only due to its more moderate importance. This demonstrates starkly how poorly our cohort's digital channels are meeting customer needs.

JOBS-TO-BE-DONE · VISUALISATION

Almost every job sits in the danger zone



Each job plotted by user importance against current satisfaction; the shaded quadrant marks where importance outstrips satisfaction. Source: JTBD analysis in this report.

What's breaking, and where

The three highest-opportunity jobs are the structural spine of the digital journey: **check-in** (28% of digital-relevant reviews mention it, 18% satisfaction), **payment** (only 10% scoring 4–5★), and **booking** (18% satisfaction). Fail any one of these for a passenger and the trip itself is at risk.

> *"Check in doesn't work. Baggage tracking doesn't work. Setting notifications for flight status updates never sends any updates. This is simply useless."*
>
> — Google Play review, 1★, UK, February 2025

Sitting just behind these three is **authentication**, the lowest satisfaction score in the study at 5%. Login doesn't make Tier 1 because importance is moderate (losing your login is not the same as losing your flight), but it is the highest-frequency job in the set: every session starts there. Reset-password loops, 2FA failures, post-device-switch lockouts, verification emails that never arrive. This is the job that wears users down most, even when no single instance is catastrophic.

> *"Switched phones and now I cannot log in. After receiving the sms/email code it just reverts back to the login screen of the app... without logging in."*
>
> — Google Play review, 1★, UK, February 2026

Underneath every functional job sits an emotional one: users want to **feel in control** of their travel experience. Every functional failure erodes this a little more. We score this emotional job 15 on the opportunity scale, in the critical under-served band alongside the Tier 1 functional jobs.

Where the design falls

Eleven per cent of digital-relevant reviews use explicit design vocabulary: "interface", "design", "layout", "navigate". Users are not just reporting that something doesn't work, they are framing the failure as a design problem. That vocabulary used to belong to the design-literate but now more and more belongs to the median traveller.

Performance complaints lead at 5% (apps that hang or fail to load), followed by navigation at 3% (users who can't find what they came for). Visual-design complaints are smaller in volume but sharp in tone.

> *"You will not enjoy using this app. It's ugly and difficult to navigate. Definitely one of the worst flight apps I've had the pleasure of using."*
>
> — Apple App Store review, 1★, UK, 2019

Forty reviews in the digital subset mention accessibility directly: special assistance, screen-reader compatibility, font sizes, support for older devices. Forty is a small number but it almost certainly under-represents the underlying need.

> *"There is no clear way to access special assistance via the app."*
>
> — Apple App Store review, 1★, UK, 2024

Visual design is the lowest-satisfaction category in our scan at 4%. When digital channels fall short, ordinary users will often now reach for design vocabulary to describe why.

Cohort comparison: budget vs legacy carriers

Across every theme we measured, a consistent split emerged between the five budget carriers in our sample (Ryanair, easyJet, Wizz Air, Vueling, Pegasus) and the five legacy carriers (British Airways, Lufthansa, Air France, KLM, Turkish Airlines). Legacy carriers outperform budget carriers on every one of ten digital themes we measured. The cohort gap on combined mean rating, across our digital-relevant app sample, is 0.58 stars (budget 1.98, legacy 2.56).

You might expect the opposite. Budget carriers tend to sit on newer digital infrastructure than the legacy ones. Technical modernity is not the obvious explanation.

Where the gap is widest

Looking theme by theme, the widest cohort gap by some distance is on **notifications and flight information**: legacy carriers satisfy 31% of users on this theme, budget carriers satisfy 7%. That's a 24 percentage point difference, more than twice the next-largest gap.

> *"Quite a standard airline app. All the basic functions are there apart from being updated about the flight status, like opening a gate. This is something I expect to be a core service, that's why 2 stars down."*
>
> — Google Play review, a budget carrier, August 2025

Why the gap exists

Two industry-level differences are well documented and may each contribute:

- **Loyalty infrastructure.** Legacy carriers operate mature, tier-based frequent flyer programmes; the budget cohort largely does not. This may drive more regular engagement and reduced friction overall.
- **Commercial model.** Budget commercial models based around upsells make for dense booking flows: bags, seats, priority, insurance, transfers, fast-track, food, carbon offset. Each is a potential friction point.

Chatbots: rising fast, struggling to deliver

Mentions of chatbots and AI support in our dataset rose from 5 in the 2023–24 window to 63 in 2025–26, a more than twelvefold increase, and reflects a real industry shift. Budget carriers automated earlier with narrower, scripted scope. Legacy carriers arrived later with LLM ambition, bringing with it a more complex quality control challenge. In 2025, a comedian's TikTok showing a major UK legacy carrier's chatbot failing to recognise its own hub airport went viral and was picked up by Frommer's, who independently reproduced the failure.

In February 2024, a Canadian tribunal ruled in *Moffatt v Air Canada* that an airline could not disclaim responsibility for its chatbot's misleading information. It remains the most widely-discussed precedent on airline chatbot liability to date.

The capability cliff

A structural problem sits at the heart of how airlines have deployed chatbots. They handle pre-sales queries and general FAQ. Complex requests route to a separate, less-automated lane: phone queues, legalistic web forms, or 30-day response windows that routinely lapse. The effect is a capability cliff. A user arrives at the chatbot with a straightforward-seeming question ("my flight was cancelled, what happens to my compensation?") and hits a system designed for pricing pre-sales. At precisely the moment statutory rights become relevant, the digital path ends.

> *"Bad experience. Dropped in another airport with nowhere to get info apart from chat bot. The refund you get is ridiculous. Missing customer services."*
>
> — Trustpilot review, UK, April 2026

Working around chatbots is now common

A whole sub-culture has formed around getting past the chatbot to a human being. UK frequent-flyer threads on sites such as Head for Points and Lemon Fool share tactics: users will repeat "human" or "supervisor" until the bot yields. They might choose a high-value intent like visa or check-in to reach a human queue more quickly. The most common advice is to escalate via social media when the chatbot won't connect to a human operator. That users are publishing these tactics, rather than simply using them, is itself a signal. The digital support journey is now experienced as something to be navigated around, not through.

Why this matters beyond aviation

The chatbot capability cliff is not an aviation problem. It is found wherever GenAI support is being deployed faster than improvements to underlying service infrastructure. A similar gap can be seen in financial services under the FCA's Consumer Duty and EU Directive 2023/2673, and in energy retail under Ofgem complaint-handling standards. Treating GenAI support as a layer over an existing service stack, rather than taking the opportunity to redesign and rethink where the automation boundary should sit, is sowing the next generation of two-star reviews.

Conclusions: what the data tells us

The first and greatest takeaway from this study is that the public app store rating is not a meaningful measure of current digital UX quality for European airline apps. Bot contamination inflates one end of the averaging, historical weighting slows the other. The gap between what a casual observer sees and what users are currently saying has become structural, and the mechanisms that produce it are not going away unless platforms and airlines address them deliberately.

The friction also creates a real cost burden. Budget carriers charging €55 for airport check-in when the app-based check-in fails are levying a direct consumer tax on digital product failures. As rights-awareness rises and regulation tightens, the cost is also accruing on the airline side of the ledger. The passenger who cites EU261 today is the passenger who writes the Trustpilot review tomorrow. A meaningful share of the industry's angriest reviews come from customers who have been forced through a digital journey that did not end in a resolution.

> *"I'm on my 5th attempt of booking a flight on [a legacy carrier].com. Every time it fails. How can a website in 2026 be this bad? Incredible. Year after year [a legacy carrier] does nothing about their terrible outdated IT and website."*
>
> — Trustpilot review, UK, April 2026

This doesn't read like fury, more like resignation. And that is probably the worst place for an airline's digital experience to leave them.

A few things we'll be watching for 2027

Some patterns are small now but worth keeping an eye on for the next edition.

1. **Digital sentiment is still falling.** Average ratings of digital-relevant reviews dropped from 2.64 in 2023 to 2.28 in 2025, with the year-to-date through 22 April sitting at 1.70. Partial-year data, but the direction of travel is consistent.
2. **Bot contamination on Google Play** has plateaued at 22–25% since 2022. We'll be interested to see if industry bodies or Google itself address it publicly.
3. **Regulatory infrastructure is catching up.** The EU's Digital Fairness Act, expected as a Commission proposal in Q4 2026, may extend consumer protections to B2C chatbot interactions. A narrower right to human intervention already exists for financial services chatbots under Directive (EU) 2023/2673.

If you're running a digital channel in a regulated, customer-facing business across aviation, financial services, energy, mobility and other sectors, the patterns in this report likely apply to you in some form. The capability cliff between automated support and statutory-rights handling is not an aviation problem. If you'd like to benchmark your own digital experience against this kind of analysis, get in touch at hello@wearesystematic.com or find us on LinkedIn.

About the author

Ed Roberts

PARTNER & CHIEF STRATEGY OFFICER

Ed Roberts leads strategic direction at We Are Systematic across its work in regulated and operationally complex sectors. His focus is how design decisions carry commercial and increasingly legal consequences for the businesses that ship them. He writes regularly on the business of strategy and design in complex industries.

ed@wearesystematic.com · linkedin.com/in/ed-roberts-49272b33/

About We Are Systematic

We Are Systematic is a UK-based strategic UX consultancy. We do strategy-first, evidence-based design work for organisations building safety-critical, regulated or operationally demanding products: automotive and mobility, IoT and AI products, financial systems and energy infrastructure. The throughline across our work is that the surface most users interact with is the smallest visible part of a much larger system, and the design decisions on that surface carry consequences far beyond aesthetics. Our job is to reduce the risk in those decisions: to help teams get the high-stakes ones right before they are expensive to change.

European airline apps: state of UX 2026 is our view on the state of customer-facing app design in commercial aviation as of 2026. It follows UK public EV charging apps: state of UX 2025, and applies the same evidence-based discipline our team brings to regulated, operationally complex products to one of the most heavily used consumer app categories in Europe.

This report draws on large-scale qualitative analysis of public app store and Trustpilot reviews, supported by the Claude Opus 4.7 language model with multi-layered human QA and a persistent evidence-preservation protocol. Any interpretive errors remain our own. A 2027 follow-up is planned.

Research by We Are Systematic. Data gathered April 2026 from public reviews on Apple App Store (UK), Google Play (UK), Reddit and Trustpilot. Sources, methodology notes and anonymisation approach available on request.

SECTORS

Automotive and mobility · IoT and AI · Financial services · Energy and electrification

CAPABILITIES

Product strategy · UX/UI design · User research · Rapid prototyping · Proposition testing · Product optimisation

CONTACT

wearesystematic.com
hello@wearesystematic.com

Benchmark your UX against the market.

wearesystematic.com/contact