SINGAPORE · IMDA GUIDELINES 2026

Singapore's AI Chatbot Transparency Guidelines 2026

On 20 July 2026 IMDA asked every consumer-facing generative AI chatbot to publish a plain-language 'info card'. This report explains the three principles and what to disclose.

Relevance | Accessibility | Timeliness

Tech Directory SG
TechDirectory.sg · Singapore B2B technology directory

Report · published 2026-08-26
Free to share · no registration

If your organisation runs a customer-facing AI chatbot — a website support widget, an in-app assistant, a companion or a general-purpose bot — what does Singapore now expect you to tell your users? The direct answer: on **20 July 2026** the **Infocomm Media Development Authority (IMDA)** issued the Transparency Guidelines for Generative AI Chatbots, which IMDA describes as among the first of their kind in the world. The guidelines are **voluntary**, but they set a clear expectation: every consumer-facing generative AI chatbot should publish a single, plain-language **'chatbot info card'** — a consolidated reference, likened to a medical label or a medicine leaflet, that tells users what the chatbot can and cannot do, how safe and reliable it is, how their data is handled and how to report a problem. This report summarises what the guidelines require, why they matter, and what any Singapore business deploying a chatbot should do about them.

> **REPORT AT A GLANCE**
>
> Issued by **IMDA on 20 July 2026**; **voluntary**, and they do not replace sector-specific laws · The core artefact is a **chatbot info card** — one consolidated, plain-language reference point · Organised around **three principles: Relevance, Accessibility and Timeliness** · The card should answer **four questions** users care about, with **at least one substantive disclosure in each** · Scope is **external-facing consumer chatbots**; purely internal tools are out of scope · A **high-level safety statement plus a link to the card at first use** is the recommended minimum · Organisations that coverage of the launch reported intend to reference the guidelines include **Google, Meta, DBS, OCBC, Singapore Airlines and Synapxe**. Full 10-page PDF available for download.

Executive summary

Adoption of generative AI has outrun trust. People use chatbots for work, study and personal life, but often despite reservations rather than with confidence — because the information they need is missing, buried or written for the wrong audience. IMDA frames this as a transparency gap: today's disclosures are either too technical (aimed at developers), too shallow (aimed at marketing), or too fragmented (scattered across terms of service, privacy policies and blog posts). The consequences are real. The guidelines cite reported cases in which users came to harm following a chatbot's dietary advice, and in which lawyers filed court documents citing case law that a chatbot had entirely fabricated.

The guidelines take a deliberately narrow first step: they address **applications, not the underlying models**, and they start with **chatbots** — the generative AI application with the widest consumer reach and the most public concern. The proposed remedy is a single, consolidated **chatbot info card**. Rather than set universal safety thresholds — impractical when risk varies so widely by use case — IMDA asks deployers to **state their own safety commitments** clearly, so that the public can measure them against their own words. That is the accountability logic of the whole document: transparency you publish is a standard you can be held to.

For a Singapore business, the practical reading is straightforward. The guidelines are not law, and there is no penalty for ignoring them. But they establish a norm that regulators, enterprise buyers and the public will increasingly expect deployers to meet — and IMDA has signalled that sectoral regulators may build on them, and that it will align with international standards bodies over time. Publishing a credible info card is

fast becoming part of what it means to deploy a consumer chatbot responsibly. This report walks through the three principles, the four questions, the disclosure pointers and the practical steps to comply.

1. What the guidelines are — and what they are not

The guidelines are **voluntary and impose no obligations**. IMDA is explicit that they do not replace, modify or affect any duties under other laws or sector-specific rules — for example, a chatbot regulated as a medical device still meets its device requirements, and financial-sector deployers still meet MAS expectations. They are a **baseline of meaningful transparency**, not a compliance regime. IMDA encourages deployers to implement at least the described minimum, then decide for themselves what and how much more to disclose.

Two boundaries define the scope. First, the guidelines target the **application layer** — the chatbot most consumers actually interact with — rather than the foundation model beneath it, complementing the global conversation that has so far centred on model disclosure. Second, they target **external-facing chatbots** used by customers or the public. Chatbots deployed *solely* for internal use, such as employee productivity tools or internal knowledge assistants, fall outside the scope. For the purposes of the guidelines, a generative AI chatbot is any application — standalone or embedded — that uses foundation models to produce adaptive, human-like responses through a conversational text or voice interface, including bots that blend generative AI with rules-based or retrieval components, and those that handle images or other modalities.

Responsibility sits primarily with the **deployer** — the organisation that makes the chatbot available. When something goes wrong, users look to the deployer first, not to the upstream model provider, so the deployer is both accountable for and best placed to provide transparency. One point is easy to miss: transparency sits **on top of** the deployer's safety work, not in place of it. IMDA expects deployers to conduct **risk assessment, mitigation and application testing before deployment**; the info card discloses that work, it does not substitute for it. Upstream providers are not off the hook either: they should furnish deployers with the model information needed to assess risk. But deployers need only disclose what is **reasonably available** to them — they are not expected to attest to a model provider's internal safety testing they cannot verify. IMDA groups external-facing chatbots into three broad, non-exclusive categories:

Category	Designed for	IMDA's cited examples
General-purpose	Open-domain interactions across many tasks — writing, creative and learning work	ChatGPT, Claude, Gemini
Companion	Social interaction and emotional engagement; persona-driven, empathetic simulation	Character.AI, Replika
Domain-specific enterprise	A particular field or task, most commonly customer service	Foodpanda's in-App Support, Singapore Airlines' Kris

The label matters less than the risk and context of use. IMDA asks deployers to weigh how consequential the chatbot is (answering a query versus acting on the user's behalf), the domain it operates in (general

versus a specialised field such as health or finance), and who its users are — in particular, whether minors can reach it. Newer forms such as agentic assistants may not fit any category neatly; what governs is the risk, not the name.

2. The chatbot info card: a 'medical label' for AI

The centrepiece of the guidelines is the **chatbot info card**: a single, consolidated reference point where a consumer can find the essential facts about the chatbot they are using. IMDA is deliberately flexible about its form — it can be a disclosure document, a dedicated information page, or a pop-up that appears when a user clicks an information icon. What matters is not the format but that the card is **relevant, accessible and timely**:

- **Relevant** — it contains the essential information users need to use the chatbot responsibly and with confidence.
- **Accessible** — it presents that information so it is easy to understand, easy to navigate and easy to find.
- **Timely** — it reflects the chatbot's latest capabilities, risks and safety policies.

> THE MINIMUM IMDA ENCOURAGES
>
> Cover the **four key areas** users care about most, with **at least one substantive disclosure in each** · Surface a **high-level safety statement together with a clearly identifiable link to the info card at first use** of the chatbot · Keep the card **easily reachable from within the chatbot** thereafter. Beyond this minimum, deployers decide what and how much to disclose, and how to design and present it.

3. Principle 1 — Relevance: the four questions users care about

IMDA's user research identified four questions that let people engage with a chatbot responsibly. A credible info card answers all four, with at least one **substantive disclosure** in each — a concrete, specific statement a user can act on. A generic line such as ‘we take safety seriously’ does not meet the bar; ‘the chatbot can produce factual errors, so verify its responses before relying on them for important decisions’ does. A shared minimum matters because it lets users find the same core facts across different chatbots instead of patchy, inconsistent disclosure.

What users want to know	The disclosure area	Example of a substantive disclosure
What can it do — and not do?	Capabilities, limitations and prohibitions	Answers FAQs on our travel policies; cannot give legal advice on incidents; not for users under 13
How reliable and safe is it?	Risks, safeguards and user precautions	Can produce factual errors — verify important outputs; report harmful content via the flag button
What happens to my data?	Data collected, who has access, training use, controls	Conversations may be used to train our models unless you opt out; you can delete your history
How do I report a problem?	Support channel and response process	Email support@example.sg; we acknowledge reports within three working days and share the outcome

The four areas break down as follows. **What the chatbot can be used for** sets expectations up front — its capabilities, its limitations and caveats on performance, and its prohibitions (age restrictions, out-of-scope uses) — so users choose appropriate tasks and know when to seek human help. **How reliable and safe the chatbot is** gives users a clear view of the risks, the safeguards in place, the residual risks that remain, and the precautions to take (fact-check, never share sensitive information, use parental controls). **How user data will be used and protected** surfaces the highlights of the privacy policy: what data is collected, who can access it, whether it trains models, and what controls (deletion, opt-out) users have. **How users can report issues** tells users the right channel, the types of issues they can raise, and how a report will be acknowledged and followed up.

Transparency is not the same as disclosing everything. IMDA sets out sensible limits on what a deployer must publish. **Proportionality** balances the usefulness of information to users against the effort to produce it. **Proprietary information** need not be revealed — where a safety measure is a competitive advantage, describe what protection is in place and how effective it has been, without the implementation detail. **Security** is protected: excessive disclosure of system internals or vulnerabilities can invite attack, so disclosure should be purposeful and calibrated. And deployers building on third-party models need only disclose what is **reasonably available** to them.

4. Principle 2 — Accessibility: easy to understand, navigate and find

Even the most relevant information fails if users cannot access or understand it. On being **easy to understand**, IMDA asks deployers to be **specific and substantive** rather than reassuring — explain what a safeguard is, how it works and what users can expect — and to **use plain language** a general adult audience can follow, broadly a secondary-school reading level (the Flesch–Kincaid index is suggested as a check). Technical terms should be explained: if the card says the chatbot can ‘hallucinate’, it should add that this means the chatbot can state false information confidently as if it were fact.

On being **easy to navigate**, IMDA recommends **layered disclosure** — a summary up front, with expandable sections or click-throughs for detail — content **structured for scanning** with headers, bolded terms, bullet points and short sections, and **visualisations** such as comparison tables or icons showing what the chatbot can and cannot do reliably. On being **easy to find**, the card should be a **one-stop shop**

linked to related documents; **surfaced at onboarding**, where a substantive safety statement (for example, that the chatbot can make mistakes and outputs should be verified) appears with a clear link at first use; **reachable from within the chatbot** through a persistent access point such as an information icon; and **standardised across platforms** so the card is equally accessible on web and mobile. Deployers offering a family of related bots may publish a single **consolidated card** that documents shared characteristics while noting the differences between variants.

5. Principle 3 — Timeliness: publish at launch, keep current, version

An info card that reflects outdated information is of limited value. Deployers should **publish the card by the time the chatbot is made available** to external users — including in beta or pilot form — so users and stakeholders (a parent researching a chatbot before a child uses it, for instance) can review it first. For chatbots already public before the guidelines, IMDA encourages publishing a card **as soon as reasonably practicable**. Thereafter, the card should be **updated within a reasonable time** whenever a development meaningfully affects users — whether from a change to the chatbot or from a newly identified external risk — and reviewed **periodically (for example, annually)** even when nothing has changed.

Update the info card when…	An update is usually not needed for…
A model change significantly alters reasoning or responses on safety-sensitive topics	A minor upgrade to a newer version of the same model with no real change to capabilities or safeguards
Guardrails or content filters change so the chatbot intervenes much more or much less	Interface and visual-design changes that do not alter the chatbot's core behaviour
New capabilities are added, such as image generation or web browsing	Backend, infrastructure, routine security patches or performance improvements
The chatbot is widely adopted for a high-risk use it was not designed for	—

Finally, IMDA encourages **versioning** so users can verify they are reading current information: a **last-updated date** and, where the chatbot uses version identifiers, the **chatbot version** the card describes.

6. What to disclose: the risk pointers (Annex A)

The guidelines include a practical annex of suggested disclosure pointers — ideas, not requirements. For each risk, IMDA frames disclosure around four questions: the **risk addressed**, the **safeguards** in place, their **effectiveness** (how the deployer knows they work, and any residual limits), and recommended **user precautions**. Importantly, effectiveness can be expressed **qualitatively** — a description of the testing done or an assurance statement — and deployers need not publish numerical scores to meet the bar.

Risk category	Most relevant to	Illustrative disclosures (risk, safeguards, effectiveness, precautions)
Harmful or inappropriate content	Most chatbots	Content types addressed (sexual, violent, self-harm); moderation measures; how you know they work; how to flag or report
Risks to younger users	Any chatbot reachable by under-18s	Age assurance, restricted features, stricter moderation; session limits and usage nudges; parental controls; how to report age-inappropriate content
Content safety in high-risk use	General-purpose chatbots	How health, legal or financial queries are handled — disclaimers, redirects to qualified professionals; when to seek expert human advice
Emotional safety and healthy usage	Companion chatbots	Measures against emotional over-reliance; how the bot responds to distress or self-harm ideation; usage nudges; where to seek support
Reliability in professional contexts	Domain-specific enterprise chatbots	The scope where accuracy is improved; anti-hallucination measures; when outputs should be independently verified

The annex also details the **data** and **reporting** pointers. On data, deployers can disclose what is collected (conversation history, metadata, usage patterns, device information), whether the chatbot reaches data from connected apps, how data is stored and for how long, who has access and any third-party sharing, whether interactions train or fine-tune models, and the controls users have (opt-out of training, deletion, export). On reporting, deployers can give a **support channel** and the types of issues it accepts, and describe the **response process** — how and when reports are acknowledged, how they are investigated and resolved, and how they feed back into improving the chatbot.

To show what this looks like in practice, the guidelines include a worked sample based on a fictional general-purpose chatbot, 'TalkPAL' — an overview card, a plain-language statement of what the bot does and its prohibited uses, and safety and reliability measures for harmful content, younger users, high-risk queries and healthy usage. It is offered as a reference format, not a mandatory template; deployers can reorganise sections or add interactive elements to suit their audience.

7. What this means for a Singapore business

If you deploy a customer-facing chatbot — and thousands of Singapore businesses now embed a support widget or an in-app assistant — the guidelines are a low-cost way to raise trust and get ahead of expectations that will only tighten. Assuming the underlying safety work is already done, the transparency task itself is mostly disclosure you can assemble from documentation you already hold. A pragmatic first pass:

- **Classify your chatbot and its users.** Is it general-purpose, a companion, or domain-specific? Can minors reach it? Does it touch health, legal or financial topics? The answers decide which risks and disclosures are most relevant to you.
- **Draft one info card that answers the four questions** — what it can and cannot do, how reliable and safe it is, how data is handled, and how to report issues — with at least one concrete, actionable

disclosure in each. Resist the generic reassurance; write what a user could act on.

- **Surface it where users will see it.** Show a short, substantive safety statement with a clear link to the card at first use, and keep a persistent access point (an information icon) inside the chat thereafter. Host the full card on a linked web page if that is easier.
- **Write for a non-technical reader.** Use plain language, explain terms like 'hallucination', and lay the card out for scanning with headers, bullets and layered detail. A comparison of what the bot does and does not do reliably is worth more than paragraphs.
- **Make it current and keep it current.** Publish before or at launch (including beta), add a last-updated date and version, and set a trigger list — model swaps, guardrail changes, new capabilities — plus an annual review, so the card never drifts from the product.
- **Protect what you should.** You need not reveal proprietary safeguards or system internals; describe what protection exists and how effective it has been, and disclose only what is reasonably available to you if you build on a third-party model.

8. The bigger picture: where this fits

The chatbot guidelines are one piece of a wider, deliberately voluntary-first approach to AI governance in Singapore. They build on IMDA's AI Verify testing framework and its Model AI Governance Framework for Generative AI, and sit alongside the National AI Strategy. The through-line is accountability by disclosure rather than prescriptive rules — and its practical edge is that a deployer's published commitments become a benchmark others can act on: regulators can hold a chatbot to its stated standard, enterprise buyers can compare vendors on it, and users can choose between one bot and another on more than marketing.

IMDA has been explicit that this is a first step. It will monitor how deployers implement info cards and may refine the guidelines; it expects sectoral regulators to build on them for applications in their domains, from education to financial and health services; and it will keep working with international organisations and standards bodies to broaden the global conversation from model transparency to application transparency. Coverage of the launch reported that organisations including Google, Meta, DBS, OCBC, Singapore Airlines and the healthcare technology agency Synapxe intend to use the guidelines as a reference in strengthening their public-facing chatbot disclosures over the following 6–12 months — a signal that, voluntary or not, the info card is becoming a norm rather than an option.

About this report

This report was researched and published by Tech Directory SG (TechDirectory.sg), a Singapore B2B technology directory of recorded company profiles. It summarises and interprets IMDA's *Transparency Guidelines for Generative AI Chatbots*, published on 20 July 2026, together with public reporting of the launch; the examples of chatbots, risk categories and disclosures are drawn from the guidelines themselves. The guidelines are voluntary and do not replace any legal or sector-specific obligation. Details are stated as at August 2026 — confirm the current position with IMDA's primary document before it informs a compliance, product or disclosure decision.

Deploying or sourcing an AI chatbot?

Tech Directory SG maintains directory profiles of Singapore's AI, automation and software vendors — the partners who build, integrate and govern customer-facing chatbots — with UENs where recorded and Profile signal scores derived from documented fields.

Frequently asked questions

What are IMDA's Transparency Guidelines for Generative AI Chatbots?

They are voluntary guidelines that IMDA issued on 20 July 2026 setting out how deployers of consumer-facing generative AI chatbots can be transparent with users. At their centre is a 'chatbot info card' — a single, plain-language reference point, likened to a medical label — that tells users what the chatbot can and cannot do, how safe and reliable it is, how their data is handled, and how to report issues. IMDA describes them as among the first such guidelines in the world.

Are the guidelines mandatory?

No. They are voluntary and impose no obligations, and they do not replace any duty under other laws or sector-specific rules — a chatbot regulated as a medical device, for example, still meets those requirements. IMDA encourages deployers to implement at least the described minimum as a baseline of meaningful transparency, then decide what more to disclose.

What is a chatbot info card?

It is a single, consolidated reference point where a consumer can find the essential facts about the chatbot they are using. It can take several forms — a disclosure document, a dedicated information page, or a pop-up behind an information icon. IMDA asks that it be relevant (the information users need), accessible (easy to understand, navigate and find) and timely (kept up to date as the chatbot changes).

What four questions should the info card answer?

What the chatbot can and cannot do; how reliable and safe it is (its risks, safeguards and the precautions to take); how user data is collected, shared, used for training and controlled; and how users can report a problem. IMDA asks for at least one substantive, actionable disclosure in each of the four areas — a concrete statement a user can act on, not a generic assurance like 'we take safety seriously'.

Which chatbots do the guidelines apply to?

External-facing generative AI chatbots used by customers or the public — grouped by IMDA into general-purpose (such as ChatGPT, Claude, Gemini), companion (such as Character.AI, Replika) and domain-specific enterprise bots (such as Foodpanda's in-App Support or Singapore Airlines' Kris). Chatbots deployed solely for internal use, such as employee productivity or internal knowledge tools, fall outside the scope.

When must a deployer update the info card?

Publish it by the time the chatbot is available to users, including in beta, and update it within a reasonable time whenever a change meaningfully affects users — a model change that alters reasoning or safety-sensitive responses, guardrail changes, new capabilities like image generation, or widespread high-risk use it was not built for. Minor model upgrades, interface tweaks and backend or infrastructure changes usually need no update. IMDA also suggests a periodic review, for example annually, plus a last-updated date and chatbot version.

Does the info card have to publish numerical safety scores?

No. Where the guidelines refer to the 'effectiveness' of safeguards, deployers may express this qualitatively — a description of the testing done or an assurance statement — as an alternative to evaluation results. The aim is to show how a conclusion about safety was reached; numerical scores are not required. Deployers also need not reveal proprietary safeguards or sensitive system details that could compromise security.

Sources

1. IMDA — Transparency Guidelines for Generative AI Chatbots (full document, PDF) · primary source
   https://www.imda.gov.sg/-/media/imda/file/emerging-technologies-and-research/artificial-intelligence/transparency-guidelines-for-generative-ai-chatbots.pdf
2. IMDA — New Transparency Guidelines to help consumers use generative AI chatbots safely and responsibly (press release, 20 July 2026) · primary source
   https://www.imda.gov.sg/resources/press-releases-factsheets-and-speeches/new-transparency-guidelines-to-help-consumers-use-generative-ai-chatbots-safely-and-responsibly
3. AI Verify Foundation — testing framework and toolkit
   https://aiverifyfoundation.sg/
4. Baker McKenzie — Singapore: New Transparency Guidelines for Generative AI Chatbots (July 2026)
   https://www.bakermckenzie.com/en/insight/publications/2026/07/singapore-new-transparency-guidelines-for-generative-ai-chatbots
5. NBC News — Man hospitalised with hallucinations after following ChatGPT advice to cut dietary salt (the case cited in the guidelines)
   https://www.nbcnews.com/tech/tech-news/man-asked-chatgpt-cutting-salt-diet-was-hospitalized-hallucinations-rcna225055
6. Thomson Reuters Institute — GenAI hallucinations are still pervasive in legal filings, but better lawyering is the cure (cited in the guidelines)
   https://www.thomsonreuters.com/en-us/posts/technology/genai-hallucinations/

This report was researched and published by Tech Directory SG (TechDirectory.sg). It summarises and interprets IMDA's Transparency Guidelines for Generative AI Chatbots (published 20 July 2026). The guidelines are voluntary and do not replace any legal or sector-specific obligation. Details are stated as at August 2026 — confirm the current position with IMDA's primary document before it informs a compliance, product or disclosure decision.