

Trust & Safety Assessment Tool

This tool has been designed to enable stakeholders to assess current practices within a tech company against the Key Recommendations in Chapter One of this report. Tech companies can use it to understand where the policies, processes and systems meet proposed best standards and where they fall short. It can also be used to re-evaluate practices if changes are made in the future. T&S professionals may find it helpful as a prompt to evaluate the extent to which their right and responsibility to safeguard society, citizens and users are supported within the organisation they work for. Policy makers and regulators can use the assessment tool to encourage greater transparency and openness from tech companies.

Mandate and authority to act	NO	Partially	YES
We recognise that T&S professionals' overriding duty is to ensure the safety of society, citizens and users — even when this conflicts with commercial priorities and interests. Our systems, processes and written policies reflect this.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Our T&S team has authority to proactively surface and measure risk of harm across all aspects of our service.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Our T&S team's remit includes oversight of product design and AI systems.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Our T&S team has the authority to modify, delay, halt, or withdraw services, products, features, and functionalities and AI systems if it determines that safety risks have not been sufficiently mitigated.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Our T&S team has the freedom to collect, share, and record safety-related data and metrics it deems necessary to carry out its role.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Safety targets and metrics are included in business-wide goals.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

<i>Mandate and authority to act</i>	NO	Partially	YES
We are transparent about the systems, processes and governance practices we have put in place to support T&S's work. We share data about the efficacy of these measures in our Transparency report.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
We require other teams to prioritise requests from T&S that relate to safety over requests from non-safety teams.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Governance systems and processes	NO	Partially	YES
We do not make decisions that impact safety unless T&S has provided advice on risk to our CEO (or equivalent).	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
If advice from T&S is not followed, or only partially followed, by our CEO (or equivalent), we document the advice from all participants in the discussion and the reasons why T&S's advice was rejected.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
When T&S team members engage with the work of legal and compliance teams, those teams provide clear advice upfront about their legal rights and duties.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Dissenting views from T&S on legal and compliance assessments are recorded and disclosed to regulators.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
T&S team members are not required support the company's defence of legal and regulatory proceedings.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
We offer T&S professionals differentiated but equivalent remuneration and incentivisation schemes designed to protect their independence.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
We recognise and support the four principles of the Right to Warn ¹ and provide support and protections to anyone raising safety concerns.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Our procedures and policies apply to all T&S professionals working on our products and services including those working via sub-contractors and third-party vendors.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

1. Open letter, (signed 4 June 2024), [A Right to Warn about Advanced Artificial Intelligence](#)

Governance systems and processes	NO	Partially	YES
Managers are trained to recognise, record, and respond to safety concerns raised by T&S professionals.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
We take a proactive approach to identifying and responding to concerns raised by T&S professionals. We do not require T&S professionals to log formal complaints as a pre-condition of taking action.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
We have procedures in place to protect the anonymity of T&S professionals raising complaints.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
We actively promote a culture of challenge. We identify and address explicit and implied inequality of power between T&S professionals and the teams they collaborate with (e.g. legal, public relations and public affairs) and those whose work they oversee (e.g. product).	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
We don't require T&S professionals to act as spokespeople or to sign off on communications or marketing materials relating to safety that they consider to be misleading or incomplete.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
To safeguard their independence, we have a written policy on the rights and responsibilities of T&S professionals.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The T&S policy includes managing real or perceived conflict of interests; advising leadership, legal and compliance teams; representing the company externally; raising complaints and remuneration including participation in share schemes.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
T&S policies are accessible to all T&S professionals, and we do not monitor or keep records of who has accessed the policies.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Safe and sustainable working practices	NO	Partially	YES
--	----	-----------	-----

We allocate resources (financial and human) to safety teams based on the assessment of needs and safe operating capacity by T&S leadership.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
---	-----------------------	-----------------------	-----------------------

Safe and sustainable working practices

NO

Partially

YES

Requests for additional resources are considered by our CEO (or equivalent). Decisions to refuse such requests are documented.



Decisions to include T&S teams in redundancy rounds are made by our CEO (or equivalent) and reported to the Board.



We have enhanced redundancy procedures when making staffing cuts that could impact user safety. These include conducting risk assessments, and a requirement to notify relevant regulatory authorities.



We have comprehensive health and safety policies, systems, and processes in place to protect the mental² and physical health of T&S professionals.



We recognise our responsibility to safeguard all those working in our T&S teams, irrespective of whether they are employed directly by us or via a third-party contractor.



We have full visibility on the health and safety policies, systems, and processes of all our T&S third-party contractors. We don't work with third parties who fail to meet our standards.



We regularly conduct anonymous surveys to understand whether T&S professionals feel supported, empowered, and able to raise concerns. Results are analysed by market, roles, seniority, and contract type.



Our executive leadership reviews the surveys and is responsible for agreeing on and implementing further action (if required).



We ensure transparency in job postings and interviews about the specific requirements of the role, so that candidates can make an informed choice about whether they are suited to working in T&S or ask for additional support they may need to perform the role.



2. Mental health is better supported by giving T&S operatives providing authority, autonomy and agency to pursue safety objectives than by offering wellbeing products such as VR headsets and apps, or employee wellness programmes. Monitoring of mental health through surveillance of T&S's operatives activities creates privacy risks that far outweigh any protective effects.

Professional standards and training

NO

Partially

YES

We have policies in place that describe the skills and qualifications needed to perform different roles within T&S and ensure that those appointed meet these requirements. If an employee doesn't meet such requirements, we provide additional training and support.



We provide support and training to lateral hires from outside the profession and to internal candidates who transfer into T&S. We also support T&S professionals when they take on additional responsibilities.



We have strategies in place to support continued professional development of T&S professionals at all stages of their career.



We support and encourage T&S professionals to convene and discuss the future of their profession.



We provide resources, time, and money to facilitate these discussions.



We do not place restrictions on the ability of T&S professionals to discuss their profession (including challenges) publicly.



We support the efforts of T&S professionals to develop professional standards, oversight and qualifications.



We do not penalise T&S professionals seeking to unionise or pursuing other forms of collective action.



We recognise the right of T&S professionals to withdraw labour.

