ATLASSIAN

Our No BS Guide to Responsible Technology Reviews

Playing responsibly, as a team

Table of contents

Atlassian's mission is to unleash the potential of every team.

As a company focused on teamwork and collaboration, we strongly believe AI and emerging technologies have the power to supercharge teams and lead to better outcomes for our communities. But with this exciting opportunity comes the serious responsibility to develop these technologies thoughtfully and deploy them with care.

We also believe this responsibility is a shared one: responsible technology is impossible alone.

This is partly because the current landscape can be overwhelming. Shiny new AI tools are appearing everywhere. Meanwhile, between existing regulatory frameworks and the coming wave of reform proposals, emerging standards, and guidance, it can be hard to navigate the rules of the game.

Atlassian knows this all too well. We've gone through the challenging process of translating the outcomes we're aiming for into the guardrails, processes, and practices needed to achieve them – then embedding them across an entire organization.

That's why we are taking an open, collaborative and, above all, iterative approach to AI governance. Atlassian has a rich tradition of sharing our own practices in the name of unleashing other teams' potential. Now, we're opening up our Responsible Technology Principles and the Responsible Technology Review Template we use to hold ourselves accountable.

Both of these resources are available at
www.atlassian.com/trust/responsible-tech-principles

We know these resources won't suit everyone perfectly "right out of the box." Our ways of working can be unique and our use cases for technology are likely to be different (and have different risk profiles) from others'. And to be clear, these resources also don't replace technical standards and guidance from recognized authorities and expert bodies. They don't substitute for thoughtful regulation and enforcement, which we believe remains necessary to build trust and help foster the responsible development and use of technology.

That said, we've learned *a lot*. We hope the lessons we've taken from this journey so far will resonate beyond the tech industry and spark inspiration for teams everywhere – that's where this guide comes in. It's both a template that aims to help other companies integrate these practices faster and a deep dive into the "whys" behind them.

We are profoundly grateful to all the experts, industry peers, and fellow Atlassians who have shaped these practices. In the spirit of our "Open company, no bullsh*t" value, it should be said that we'll almost certainly continue to evolve them as the AI landscape changes. And that's ok. Unlike a Jira ticket, responsible technology never moves to the "done" column.

Let's dive in!

Our high level takeaways

To start, here are the three biggest lessons we've learned and embedded in our ways of working:

1. **Connection is fundamental to collaboration.** This means making cross-functional connections within and between teams and stakeholders. But we also mean "connection" in another sense: being able to connect your work directly to the goals of your organization. Translating your organization's mission to its values, to its responsible technology principles, and through to your work – that's what makes the value of your work crystal clear.

2. **Perfection is the enemy of progress.** When it comes to responsible technology, the journey *is* the destination: there's always more to be done. We could easily have spent hours, days or months refining our processes, adding more detail and perfecting our choice of words. But by shipping early and iterating often, we hit the ground running while also welcoming and integrating feedback along the way.

3. **Design for practice, not (just) process.** We SMEs all had ideas on what the "best" review and assessment processes should look like. But we consciously put that aside in favor of meeting our teams where they are. We focused just as much on the design and form of our templates and processes as we did on their content. We emphasized elements that encourage curiosity and openness while minimizing aspects that could nudge teams towards check-the-box responses.

Embedding responsible technology at Atlassian

Strong technology governance is built both from the top down and the bottom up.

Taking it from the top, we first established our Responsible Technology Principles as our North Star. These principles guide all of our work and help us to take accountability for using technologies like AI responsibly and in line with our company values.

Our Responsible Technology Principles were heavily informed by, and designed to align with, a number of similar principles embedded in frameworks globally. But they also feel uniquely Atlassian. We drew on our company mission and values to articulate a clear perspective as well as our commitments to our customers, employees, and community stakeholders.

Of course, principles are nothing without action.

From the bottom up, we focused on practices – or "plays" as we call them. Atlassian teams use dozens of plays as templates for addressing common challenges, starting important conversations, or just getting sh!t done day-to-day.

Building on this, we're embedding responsible technology reviews (with our Responsible Technology Review Template at the heart) as standard practice across all our teams that develop or make decisions about technology. This is how we integrate responsible technology into everyday operations.

The full template is available to download separately but is also linked at the end of this guide. We designed it to be suitable for any team and to support consistent, efficient, and scalable practices. It's our hope that anyone can pick it up and run with it.

We also have a lot to share about how we got the practice of responsible technology reviews to where it is today – which is where we'll turn our attention next.

Interested in learning more about plays?
We open-sourced the Atlassian Team Playbook! **Check it out here**

Setting the scene for responsible technology reviews

❶ An upfront promise: no surprises

Our template is first grounded in our Responsible Technology Principles.

We then explain upfront that this template is for all teams thinking about how to build, deploy, and use a wide range of new technologies across Atlassian. The principles and our template apply to, but aren't limited to, AI (which is why we talk about responsible *technology*, not just responsible AI). They're used across all aspects of our organization and our teams' work, not just in our customer-facing products, but across all of our internal and external activities that relate to building, deploying, and using new technologies.

We were intentional in describing the template as a way for teams to reach a shared understanding. Its goal is to prompt and promote discussion and internal alignment (vs. reaching the "right" answers). This is critical when it comes to topics that can feel fuzzy and may not have any right answers.

This "no surprises" approach is also why our template strives as much as possible to include all of the information teams will need to complete the template in one place (though not necessarily in one sitting). While external resources can be incredibly useful – internal documents like company policies will always need to be consulted – we didn't want to create a situation of endless open tabs and confusing cross-references. Because ease of use matters.

② How to have the important conversations

We mentioned that our responsible technology review aligns with our team practices – or, plays.

Our plays are designed to be self-serve so teams don't always need an expert facilitator. To ensure that our reviews can scale across our company, our template works the same way.

We ask teams to work through the template together, as a form of play that can be completed either synchronously or asynchronously. The "directly responsible individual" for the project, product, or tool will be accountable for making sure that the template is completed, but we encourage wide participation across teams.

We also reinforce that the template is intended to be a living document. It needs a regular check to make sure it stays accurate and complete, and to track the actions the project team commits to.

Internally, and based on feedback we received since we first developed the template, we added a version control dialogue to reinforce and help teams record when and how they revisit their reviews.

A | **Introduction** Project Summary Principle 1 Principle 2 Principle 3 Principle 4 Principle 5 Key Takeaways

Introduction

This template is based on **Atlassian's Responsible Technology Principles** and helps teams build a shared understanding of how their project affects employees, customers, and community stakeholders. By being thoughtful in the use and development of technology, we can move fast and *not* break things.

"How do I use this thing?"

1. Step through this template as a project team and don't expect to work through the whole thing in one go. You may need to split off to gather information or talk to stakeholders, then come back together.

2. Make sure the project's lead is involved throughout. This person might be called the owner, approver, directly responsible individual (DRI), or some other clever name your company has invented.

3. Tip: you don't have to complete a responsible technology review before beginning development work. But you'll be more objective in your thinking (and end up with less re-work) if you do.

4. Treat this as a living document. Come back and reassess if the project pivots direction, the user persona changes, you go from beta to GA, etc.

5. Tip: use this template to assess the technology you're building or buying for internal use, too.

A note about alignment ratings

Each section focuses on a different guiding principle. Along with specific questions to consider, you'll also rate how well the project aligns with that principle. We've found a simple red/yellow/green system works well. Here's what each rating means.

FEELS GOOD

You're confident this use of technology aligns with the relevant principle.

NEEDS WORK

You've uncovered risks or potential issues. That's OK! Reassess your plans and adapt as needed.

NOT ALIGNED

You've identified ethical issues and/or problems that might emerge later.

Importantly, because we ask teams to self-serve the template, we ask them to self-rate their responses both overall and in each sub-section.

Those ratings are subjective, and so they can't be definitive. This is why, although we use simple red/yellow/green "traffic light" indicators, we don't label them using straightforward ratings like high/medium/low. We instead use labels like "feels good" or "needs work."

In this way, we use this as an important mechanism for teams to pause, think through how they feel about the alignment between their work and our principles, and take accountability for their answers.

Similarly, we also build in consistent reminders that the aim of this template isn't to "pass" or fly through with all green responses. The purpose is for teams to be thoughtful, accurate, and complete in their review.

③ Use design to emphasize substance over form

Our responsible tech review follows a consistent format. We break down our Responsible Technology Principles one by one, then ask teams to consider and respond to a series of open-ended questions that relate to different aspects of that principle.

- Explain why we're asking teams to think about and respond to it

- Ask teams to self-rate their alignment with the principles in their responses (see above)

- Require teams to set out what they have already done to improve that alignment, and document their plans and actions to keep making those improvements moving forward

Because the template is used by teams across Atlassian, we use the somewhat imperfect term "tech" as a placeholder for the product, feature, tool, or project that the team is describing throughout their review.

Why add the "why"?

Each section of our template includes two lines at the top, after we identify the principle to which the section relates:

- An overarching question designed to prompt teams to think about the purpose of that section and introduce the more specific questions asked under it – e.g., *"What are the known and potential uses?"*

- A brief, 1-2 sentence statement that explains to teams *why* we're asking those questions – e.g., *"It's our responsibility to control for bad outcomes to the best of our ability and drive toward good outcomes. That starts with ensuring this tech is fit for purpose."*

These weren't in the earliest versions of the template. But we added them because it became clear that the design of the template needed to do a lot of work to walk teams through their self-serve, potentially asynchronous review.

We know that responsible technology concepts might be new to a lot of readers and team members, and that they can also be quite broad and a little fuzzy. This opens these concepts and questions up to lots of different interpretations.

By also describing to teams why we ask each series of questions, we aim to set out not just the letter of the requirements but also to articulate their spirit.

④ Think about what comes next

The design choices we make in the template also feed into how we treat completed reviews.

All teams working on relevant projects and tools are required to complete the template. However, they are not doing so in order to seek any form of responsible technology "approval" for their project.

This is why the template is not designed as a checklist that teams must "pass". As we described above, it is designed as a play – and it is used as a learning tool. In guiding teams through their reviews, the template not only helps us to have the confidence that we have fully thought through the tech, its benefits and its impacts. It also helps us to raise awareness across Atlassian of the ways to identify these issues.

So, what happens after the review is complete?

First, our Responsible Technology Working Group members will review the completed template. Because teams remain accountable for their own projects and decisions, this review is undertaken with an eye to making sure that the template is "complete" more than that the decisions made within it (or about it) are "right". Our working group often engages directly with project teams to suggest ways to make sure that their template is a thorough record of the benefits and impacts of the project.

Then, the completed template can be used as an input for all of the usual review and approval processes within our company, including legal, privacy, and procurement checks. This allows us to work cross-functionally and ensure that the template – while not a standalone approval process – feeds into those existing processes and helps to inform their outcomes.

As part of these processes, we also take a second look at how those templates have been rated. Recall that teams' self-ratings aren't definitive. So these processes have to account for templates with a lot of amber, or any red ratings without clear improvement plans. But they also compare these against objective thresholds for reviews, and might even prompt extra attention where templates are all green.

What are the trade-offs?

When it comes to responsible technology, trade-offs are inevitable.

They are directly mentioned in the template – and also inherent in this process.

Our choice to approach responsible technology reviews as a play with a focus on learning meant that the template was intentionally subjective. The template isn't scored, and we try to discourage check-the-box approaches.

In doing so, we made a conscious trade-off between the effectiveness of our responsible technology reviews and the resource intensiveness of subsequent reviews. Right now, the review that our Responsible Technology Working Group members complete is manual: they need to take in the completed page and consider it against the criteria we refer to above.

But we know that this approach won't be scalable over time. We're actively working to address this issue right now, and look forward to sharing our progress with you.

Finally, as we mentioned above, we expect teams to revisit and revise their review at regular intervals. This includes scenarios and trigger events like:

- Products preparing for a general availability launch, a wide internal release of product capabilities, or any public / third-party release

- A non-trivial change to the use case(s)

- Any major changes to the information on which the review was based

- At the time(s) specified in the completed template for action/ improvement items

- When the team has answers to any unknowns they've also documented

With all that in mind, it's time to consider how we connect our principles to their practice.

Translating our principles into reviews

Our responsible technology principles underpin all of our work, including our reviews. They are:

- Open communication, no bullsh*t

- Build for trust

- Accountability is a team sport

- Empower all humans

- Unleash potential (not inequity)

Our template maps teams' responsible technology review to each principle, though in a different order to the way that we published them, and set out above. This deviation is intentional because the sequencing is equally important in both contexts.

That is, many of our external readers will encounter our principles in practice first from the perspective of our efforts at open communication (which is why that principle is front and center). In contrast, we want our internal teams to start with – and keep front of mind – the humans who will ultimately benefit from (and be affected by) our products and our work.

Let's look at how each principle is carried forward into the template.
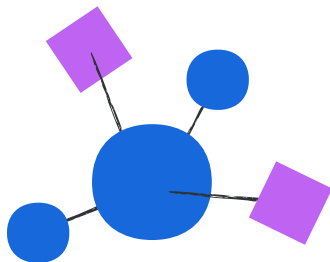
① Empower all humans

Each section starts with a broad series of open-ended questions to the team, and then focuses in on areas of potential concern.

This is particularly important when it comes to this first principle: we need to make sure that teams are thinking about every person who may come into contact with their project (or its outputs and outcomes). Critically, teams must also consider how people's identities, experiences, and perspectives could shape or be shaped by these interactions with technology.

In asking this series of questions, we have two aims.

1. Guide teams step-by-step through ways to understand and identify the circumstances in which unfair and unjust outcomes might arise, both for those who use the tech directly and for those who may have it used *on* or *about* them.

2. Prompt teams to take stock of the perspectives that they had (and will have) access to throughout the course of their project, and how broadening their stakeholder community could improve those outcomes.

At Atlassian, we want our company and our technologies to be open, inclusive, fair and just: to reflect the human-centric values and fundamental human rights that we all share. Our journey to build responsibly reflects this aim.

As we've worked with teams on their reviews, we've found that this section of the template has not only helped teams think about unfair or unjust outcomes but has also surfaced positive outcomes for certain groups and individuals. For example, tools designed to help users by providing them with standardized text suggestions and responses can be disproportionately useful for non-native English speakers, by leveling the playing field with their native English-speaking colleagues.

We know that for each of the areas covered in the template, the difference between the alignment ratings can be a matter of degree and shouldn't be considered in isolation from the rest of the review. The difference between a **red** and **amber** result could depend on a number of factors:

- Can issues can be fixed by taking the actions documented to improve the result?

- Might shortcomings be addressed in other areas (such as through increased transparency)?

- Does the team need more information to understand whether the result can be improved?
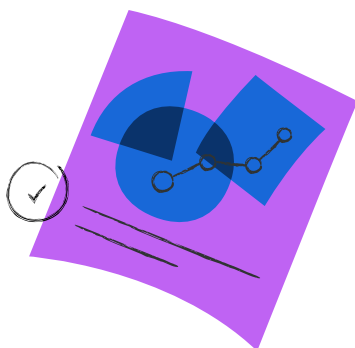
What does this look like in practice?

One example of a potential **red** and **amber** result, which was relevant to our products and teams, could occur where **we're thinking about building new features to help highlight and map frequent collaborators across teams,** and auto-suggest that those collaborators be added to new projects.

- This might contribute to exclusionary or cliquey behavior in the workplace, or even discourage including new joiners. However, we don't know whether excluded individuals are part of a marginalized group.

- There may also be simple actions we can take (such as introducing "noise" into results) to mitigate this potential impact.

2 Unleash potential, not inequity

In this section, we keep our users and stakeholders front and center while bringing in the basics of the technology that they will interact with and be affected by.

Use the right tool for the job

Before teams commence their review in earnest, we ask them to summarize their (sometimes complex) projects down to a short, plain language description. Importantly, we also ask them to clearly set forth how they intend it to be used.

Then, when we reach this section of the template, we ask teams to explain how the technology that they have chosen is fit for that intended purpose. In this way, the intended purpose becomes a touchstone for teams and any later reviewers in assessing the suitability of their chosen tool.

We warn teams that a **red** and **amber** "fitness for purpose" result could occur where the tech that they're proposing to use needs to rely on too many substitutes or proxies for what it's really trying to measure.

To drill this down to its basics, we bring things back to our company mission: to unleash the potential of every team. We ask teams to explain how their tech fits not only the purpose they've described but also our values and drivers of teamwork and collaboration. It's a real confidence boost when teams can clearly say that their tech makes collaboration easier for everyone and describe how it does so.

We know that behind every great human achievement, there is a team. We also believe that new technologies can help empower those teams to achieve even more. If we use these technologies (like AI) responsibly and intentionally, then we can supercharge this vision and contribute to better outcomes across our communities.

What does this look like in practice?

There are lots of ways in which tech could have a **red** and **amber** rating here. For example, teams should carefully investigate and think about how to assess **tools that try to rank the productivity of team members.**

- While there are several proxies for how productive a team member is (e.g. how quickly they respond to and resolve tickets, or how often they work on and are mentioned in pages or tickets), productivity itself can be very difficult to measure.

- The more closely we can tie metrics to desired outcomes, and to the context in which they are measured, the more likely it is that the use case is fit for purpose.

Risks and impacts

There are lots of potential ways in which tech and tools can be used
– or *mis*used.

As the Human Technology Institute explained in their report on the
State of AI Governance in Australia, misuse of new technologies like AI
can arise from a number of scenarios:

- **Failures** – when the tech fails "to operate in the way, or to the level of quality, required"

- **Malicious use** – when the tech is deployed for malicious purposes or used in misleading ways

- **Overuse** – when the tech is over-used, used inappropriately, or deployed recklessly

These forms of misuse can result in any number of potential harms. However, we know that some types of harm may not be immediately apparent or recognizable to teams, especially where those teams consider themselves to be strongly user-centric and values-aligned.

This is why we first introduce these types of misuse and harm in our template as "unintended consequences" that teams should consider. Although we know that the impact is the same whether or not harm was intended or unintended, we specifically want to prompt teams to come up with impacts and consequences that they might otherwise not account for if they are *only* thinking about their intended purposes, uses and consequences as set out above.

Ultimately, this means that there's a lot of ground to cover in a fairly short review.

So we wanted to boil this down to a simple question that anyone can respond to, regardless of whether they've thought about this potential for misuse before: *How might a supervillain use this?*

Thinking through the worst-case scenario of a tool (and how to control for it) in this way can also help teams consider ways to upgrade their provisional rating from **red** to **amber** to **green**.

But taking a truly human-centric approach means that we also don't want to lose sight of the best case. To maximize the benefit and value that we all derive from these projects, we ask teams to work through their use case and design their corresponding actions and improvements, with both best and worst cases in mind.

Throughout this section, we also start to introduce more reminders:

- Think about how these matters might change and evolve over the entire lifecycle of the tech

- Obtain, refer to, and document evidence of the ways in which the teams have been able to test and verify the information they have set out in the template
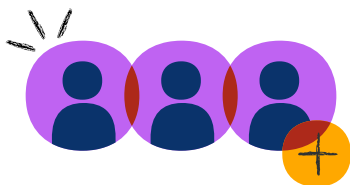
What does this look like in practice?

One example of a potential **red** or **amber** result in this area could arise in relation to **systems that describe or show people, cultures, or society.**

- This could give rise to risks that certain groups could be stereotyped or even erased (e.g. tools that generate stock images of "business leaders" as older men and "recruiters" as younger women, and underrepresent all people of color).

- Teams should carefully consider the documentation available about how these systems perform, their propensity for reflecting harmful stereotypes like this, and on that basis assess whether it's worth the risk to build or procure this tech.

3 Build for trust

Data, choice, and control

Our template is clearly no substitute for detailed privacy and legal advice, analysis, and assessment.

But it's a useful place to remind teams about the ways in which company policies, commitments, and public statements will have already influenced users' and stakeholders' expectations about how we'll safeguard their data.

It's a truism that trust is hard-won and easily lost. This is just one of the reasons why teams should be thoughtful about not only what they *can* do within the bounds of the law, but what they *should* do in light of what users and stakeholders have come to expect.

We also use this section to signal to teams that the choices they make here can also enhance trust, particularly when it comes to giving users and affected individuals choice and control. In particular, thinking carefully about the types of data used, how much data is used, and the technical controls that teams can implement can all contribute to reducing the risk of a **red** or **amber** rating.

Trust is at the heart of our work and our products:
if someone doesn't trust our company, they won't use our products or want to work here. This extends to the technologies that underpin and power our products and our work.

We know that trust is not just about ensuring the security and privacy of our products, but is also earned and kept through our actions and commitments to reliability and performance.

Quality and performance

We then remind teams about concepts already discussed in the review, and consider how they contribute towards trust in the technologies that we use.

This includes establishing trust through processes like testing and verification, and the ways we can make sure that we're using the right tech for our intended purposes.

To do this, we ask teams to delve a bit deeper into the datasets, deployment environment and performance of the technology that underlies the tool or project the team is seeking to implement.

These questions try to prompt teams to think about some fairly complex matters:

- Understanding the suitability of the datasets that were used in developing and training the tech

- Thinking about whether the outputs that the tech provides are straightforward or not

- Thinking about how the overall environment in which the tech is deployed might impact how it evolves and changes over the course of its lifecycle

We know that this is a lot to get into in the context of this type of review and that teams may not always be equipped to answer these questions. We have aimed to provide them with some initial, indicative guidance (see page 26) and expect this area to be one that develops and is added to over time.

We are particularly grateful to external and industry experts for the work that has already been done and will continue to be done in this space. This includes initiatives like **Datasheets for Datasets,** and others outlined at the end of this guide.

What does this look like in practice?

In this case, it's helpful to consult available research, evidence and other documentation about the tech and its underlying datasets and consider what these resources reveal about the purposes for which they were intended to be used. For example:

- Teams might have confidence in a **green** rating where there is clear evidence that the tool has succeeded in the same environment and conditions in which the team intends to use it.

- Alternatively, **red** ratings are more likely to result where a team can only find basic (or no) research or information, and no evidence of use in similar circumstances to the use case.

This is also an area where teams can find ways to upgrade their potential ratings by discussing, confirming and documenting the defined timeframes within which they will re-review system performance in complex environments.

④ Open communication, no bullsh*t

In this section, we ask teams to think through not only *what* might be important to explain to customers and users, but also *when* and *how* best to provide those explanations.

There are two reasons why we do this.

The first is because we know that transparency needs to be meaningful to be effective. Disclosures that are invasive, annoying, or excessive in the circumstances could just as easily cause notification fatigue. This leads users to ignore notifications and miss out on the potential benefits that transparency could otherwise give them. To illustrate, just think about how many pop-up consent boxes you might have checked, just to make them disappear…

The second reason is that we know that transparency alone isn't enough. With new technologies like AI, just telling users that a feature is AI-powered won't mean much unless we are also able to equip the user with an understanding of what it means when the AI tool is in use, how best to interact with it, and how to respond to (and if need be, challenge) its outputs.

As a result, this becomes one area where the actions that the project team agrees to and follows-through on can easily help upgrade a response from **red** to **amber** to **green**.

Openness is foundational to Atlassian – one of our core values is Open Company, No Bullsh*t. It's important that anyone who wants to make the most of new technologies is equipped with the right information to do so.

What does this look like in practice?

For example, if a team wants to deploy **a chatbot that responds to customers using natural (and conversational) language.**

- This could be classified as **red** if we don't tell the customer that it's not a human.

- Ensuring that we disclose this fact could make it **amber** (e.g. if we make the disclosure but it's not clear) or **green** (if the disclosure is clearly shown to the customer at the right time and in a helpful way, such as when they begin to use it).

- There are likely to be several different appropriate forms of disclosure depending on the context, and these may also change over time as more similar chatbots proliferate. For example, a simple icon or logo that identifies the chatbot as being AI-powered could be enough.

⑤ Accountability is a team sport

Human oversight and involvement: going beyond "humans in the loop"

In this area, we start by defining and ensuring our own areas of accountability.

We believe that it's critical to understand and define how humans use and govern new technology over time – and decide when and how to use it. This is because new technologies like AI are not just standalone, technical systems that operate in isolation: these systems (and the data that power them) reflect the values and behaviors already present in our societies, and they are also continually influenced by how we interact with and are affected by them.

In our view, this isn't just about having a "human in the loop", but in thinking through what that human involvement looks like.

In general, the less human involvement with and eyes on a tool, the more likely we are to miss problems with how it's working and the more risk we could introduce.

But this isn't a set-and-forget rule. Even in circumstances where humans are still the ultimate decision-makers – such as when they are expected to act on an AI-generated recommendation – evidence shows that humans often over-rely on recommendations made by these sorts of tools. This will impact how much human involvement is able to "fix" the potential issues with the tech.

This section is also important from the perspective of following through on earlier parts of the template. Oversight of *how* and *when* to deploy the tech will also help to ensure teams can adhere to any limitations we uncovered in the transparency analysis above.

At Atlassian, we know a thing or two about collaboration and teamwork. Our products are powered by our own people, upon the foundational technologies that we use to deliver them – and, of course, by how our customers' teams choose to use them.

While we take ownership over our technologies, true accountability is a team sport.

What does this look like in practice?

One example of a **green** rating in this area could arise where a team wants to deploy an **AI-powered security monitoring tool that generates alerts for human review.** Green is the likely result because individuals will be reviewing those alerts and making decisions based on them, as well as the types of alerts and decisions involved can be easily investigated and verified.

Third parties and our place in the supply chain

In all cases, we own and take accountability for the tech we put out into the world.

But we also know that we are often only one link in the supply chain that makes up the technology ecosystem because we often use third-party vendors and partners to provide the underlying technology, or a part of it (including datasets) that powers our tech.

In those cases, we might have less ability to directly control its outputs, and so we need to think carefully before we proceed. In addition to the matters raised above in relation to testing, validation, and verification, we ask teams to consider the vendor or partner that they're dealing with.

- If we know (and have verified) that the third party is aligned with our views on responsible technology, we can be more confident in our ability to address any issues with their tech as they arise.

- The third party might be more advanced than we are on their responsible technology journey. In that case, we would love to learn from them so teams should think about how we can formalize ways to do so.
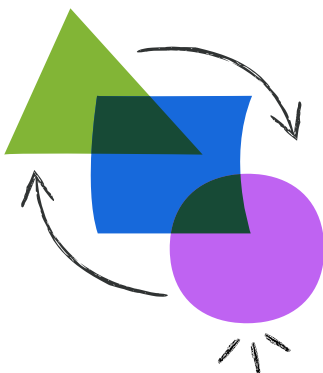
Feedback and continuous improvement

We are committed to having processes and mechanisms in place to help us obtain feedback from our stakeholders and take guidance from experts, internally and externally, on an ongoing basis.

We were also intentional in asking about feedback and continuous improvement at the very end of the template. This is because even after teams have worked through all of the previous sections, they are likely to still have questions, hypotheses that haven't yet been proven, and other unknowns.

With these unknowns fresh in their minds, the team can:

- Consider how to supplement any preventative controls and improvements with processes for reporting, identifying, and responding to any outputs that manage to slip past those controls.

- Identify ways to use user, stakeholder, and community feedback to bolster and prove (or disprove) their tentative hypotheses, as well as identify further issues they may not have considered.

- Set themselves up for success in the governance of their project moving forward.

This section also helps to remind teams that the template is meant to be a living document, with the capacity to track and monitor these issues on an ongoing basis.
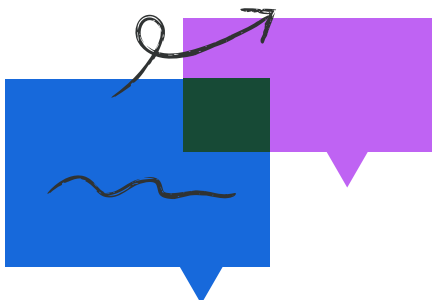
Lessons from reviews in the wild

① Case study: Atlassian

Recall that we ask teams across Atlassian working on relevant projects and tools to participate in responsible technology reviews. This includes our product teams working on **Atlassian Intelligence** features.

One of these teams is Central AI. As part of the process of launching smart answers in Confluence search, the team was one of the first to use the template and experience the play, so we asked them to outline their experiences with it. **Izzy Kohout, Product Manager, laid it out for us:**

"The template was already suited to the way we work: a lot of our work already happens in Confluence, so that made it easy to pick up and get started. Our team is also distributed across both Sydney and the US, so we collaborated on the template asynchronously to ensure that the perspectives of all crafts were represented, including engineering, product and design.

The review prompted us to think about the potential worst-case scenario of our feature, as well as how we could mitigate existing issues with large language models: like what could happen if our AI-powered search presented inaccurate or misleading information to users. We were able to consider these risks early on, and mitigate them by designing our communications with users thoughtfully to help educate them and manage expectations. We also expedited our work to reduce the frequency of hallucinations in answers, and made significant improvements as a result.

As one of the first teams to use this play, we worked with the Responsible Technology Working Group to iterate on and improve the template based on our experience. This allowed us to hone in on the most important questions, and more easily embed responsible development into our process."

② Case study: Human Tech Institute

Atlassian partners closely with the **Human Technology Institute** (HTI) at the University of Technology Sydney (UTS), to help us put fairness, accuracy, and accountability at the heart of how AI is designed and used.

Our work with HTI included collaboration on the template over the course of its initial development and iteration. **Nicholas Davis, Co-Founder and Director of HTI, discussed our work together on the template:**

"HTI is incredibly proud to have Atlassian as one of its core advisory partners in advancing responsible and human-centered technology. Our collaboration on Atlassian's Responsible Technology Review Template is emblematic of an incredibly useful exchange of insights and practices at the frontier of tech governance.

In particular, as Atlassian built out the template, our team were proud to be able to add expertise from HTI's AI corporate governance program. This included highlighting the range of current practices that organizations have adopted for assessing emerging tech investments before they are launched, as well as gaps that existed at the project and organizational levels. We were very pleased to be able to contribute specific guidance around how the template can align with emerging international standards in AI risk management and governance, which we hope enhances its robustness and applicability across industries and countries.

When compared to other instruments, Atlassian's template is noteworthy for its balance and practicality. Its engaging format supports relatively detailed project documentation, risk assessment and team reflection while remaining user-friendly. We particularly appreciate how it prompts users to identify stakeholder impacts and balances technological opportunities against risks.

Inspired by the template's user-centric design and collaborative format, HTI now uses Atlassian's Confluence platform for its own tech review template development and prototyping. Moreover, the template's adaptability and ease of use have influenced our approaches to supporting organizations keen to embed human-centered approaches to AI development and deployment. This has been particularly useful to us as HTI focuses on increasing the awareness and usage of structured AI governance strategies and approaches across Australia and other markets."

Resources for taking a review to the next level

We know that there are loads of resources out there to help organizations bring principles like ours to life.

These can range from detailed technical tools, to standards and guidelines, to games and plays.

Any shortlist of these resources will inevitably leave out a lot of important and inspiring work that has been done in this space. But when speaking with our teams, it became clear that they were seeking some additional, specific resources to help them further explore parts of their reviews.

We found the following resources helpful in those scenarios:

- Microsoft's **Judgment Call** game, to work through stakeholder perspectives

- Google's **People + AI Guidebook**, on designing with AI

- The **Ethics Canvas**

There are also a lot of Atlassian team plays that can be adapted to those questions, like:

- **Pre-Mortem**

- **Empathy Mapping**

- **5 Whys**

- **Disruptive Brainstorming**

Conclusion

We are committed to using our principles alongside the practices, plays, and processes that underlie them to guide our work, decision-making, and communications on the use of responsible technology.

We are also committed to continuing to learn, iterate, improve – then share. We have learned not to let perfection be the enemy of progress, and we also believe that no one company can solve this challenge alone. We will keep being open about our journey, and inviting your feedback.

Find your Responsible Technology Review template here