How are Interactive Mini Apps kept safe and secure?

Interactive mini apps · https://cogniti.ai/docs/how-are-interactives-kept-safe-and-secure/

Interactive mini apps are learning activities that run in a student's browser, created by you with the help of an AI builder. They are designed with multiple layers of protection to keep you and your students safe and prevent misuse.

- **Runs in a secure sandbox:** Every mini app runs inside an isolated container in your browser. It cannot access your other tabs, navigate you away from the page, or read your cookies or login session. Even if the AI generates unexpected code, it stays contained.
- **Users cannot edit code:** The AI harness builds mini app code, and users cannot edit the mini app code directly.
- **Only pre-approved libraries:** Unlike other vibe-coding platforms, Cogniti mini apps only draw from a set of pre-approved 'libraries' or plugins to power its functionality, preventing the mini app from pulling in untrusted scripts.
- **Network restrictions**: Code inside a mini app can only communicate with the Cogniti server. It cannot send your work, your answers, or anything else to another website in the background: requests to other sites are blocked, and so is quietly redirecting you to one.
- **Links to other websites:** A mini app can offer you a link — to a reference, a source, or one of its own files — but it cannot follow one on your behalf. When a mini app wants to open something, Cogniti asks you first and shows you where you would be going. Once you choose to open a link, you have left Cogniti and the protections on this page no longer apply, exactly as when you follow a link from any other website.
- **Permission controls**: Educators control exactly who can find, use, edit, and clone each mini app. New mini apps default to the most restrictive settings. Educators can widen access as needed, and grant co-owner access to colleagues.
- **AI helper safeguards**: When a mini app includes AI-powered features (like hints or answer checking), each student session has a usage budget. Once the budget is reached, further AI requests are blocked. Session tokens are single-use and expire automatically.

- **Content sanitisation**: Any text authored by educators (such as instructions) is cleaned to remove potentially harmful content before it is shown to students.
- **Content safety:** AI-generated content and user inputs are automatically screened for harmful or inappropriate material. This applies to both the AI code generation process and any AI-powered helper functions within a mini app. If a violation is detected, the request is blocked and the incident is recorded for educator and administrator review.
- **Safe cloning**: When a mini app is cloned, the copy starts with restricted permissions and is unpublished by default. Only resources and AI agents the cloning user has access to are carried over.
- **File security**: Uploaded resources (images, data files) are stored securely, scanned for malware, and served with unique access keys.
- **Monitoring & audit**: All significant actions — creating, editing, deleting, and AI usage — are logged for accountability and troubleshooting.