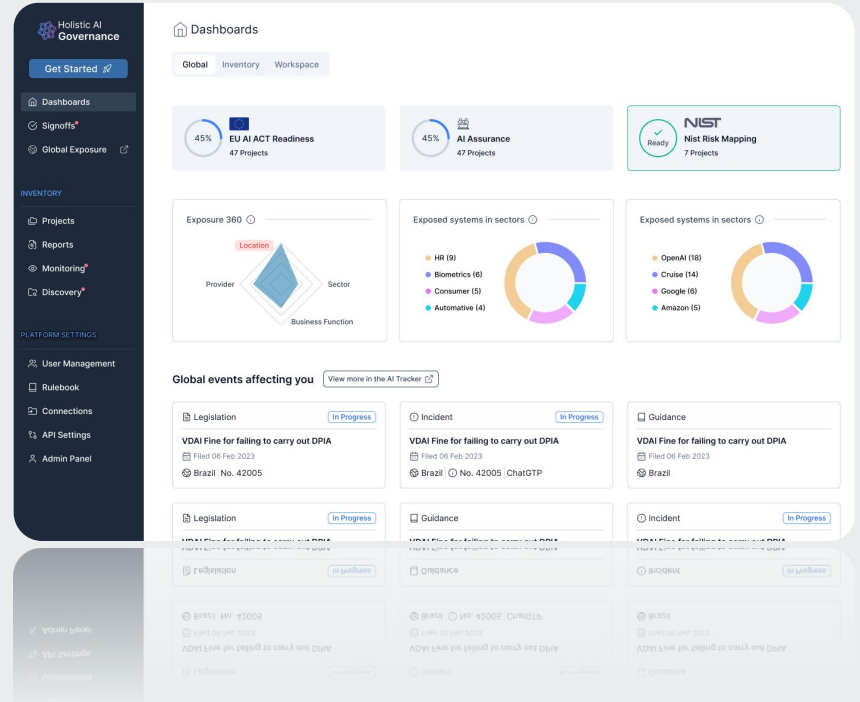


Holistic AI

Designing Ethical AI at Scale

Design Systems & Product Leadership

by Sim Deol



Outline

- Context and goals
- The problem
- My role
- Approach
- Scaling design
- Products and deliverables
- Rollout, impact and reflection
- Further exploration



Context

Context and goals

Company - Holistic AI - enterprise platform for AI governance, risk and compliance

Initial State - One legacy Angular platform, custom components, poor scalability, no unified UX, inconsistent UI

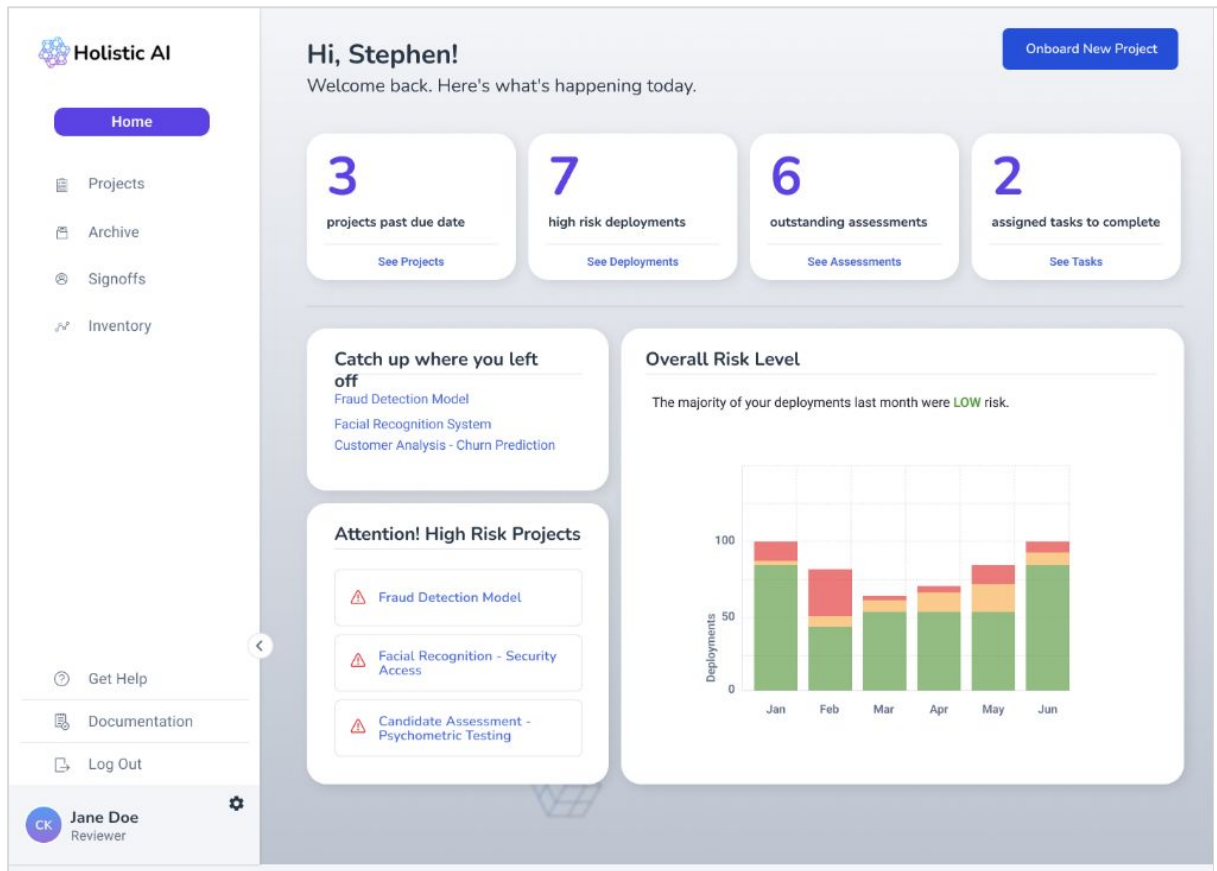
Goal - Modernise the core platform and enable rapid expansion into new products

Mission - Redesign infrastructure, scale UX, accelerate delivery



The Governance Platform - Centralised Ethical AI Management

The original core platform
was robust but rigid.



The Challenge

- Outdated tech (Angular + bespoke components) hindered dev velocity
- Difficult for engineers to onboard and iterate quickly
- No reusable design patterns, inconsistent UI, high design/development overhead
- Needed to scale into two new products while redesigning the original

My role



Sim Deol

Sr Product Designer

- Sole Product Designer across all product streams
- Led UI audits, design system strategy, and component library development
- Partnered closely with engineers and product managers to align on priorities, technical frameworks and implementation approach
- Assisted in roadmapping, sprint leadership
- Conducted research and collated insights from users
- Led design reviews & stand-ups,



Approach



Strategic & Technical Approach



- Audited UI and front-end workflows to identify friction points, as well as a comprehensive review of existing UI and UX patterns across the product.
- Worked with engineering to implement a React + Tailwind CSS stack, leveraging shadcn/ui prebuilt components to accelerate delivery despite limited front-end capacity.
- Designed and built a tokenised component library in Figma to mirror implementation structure
- Introduced UX conventions, naming logic, and documentation to ensure ease of adoption

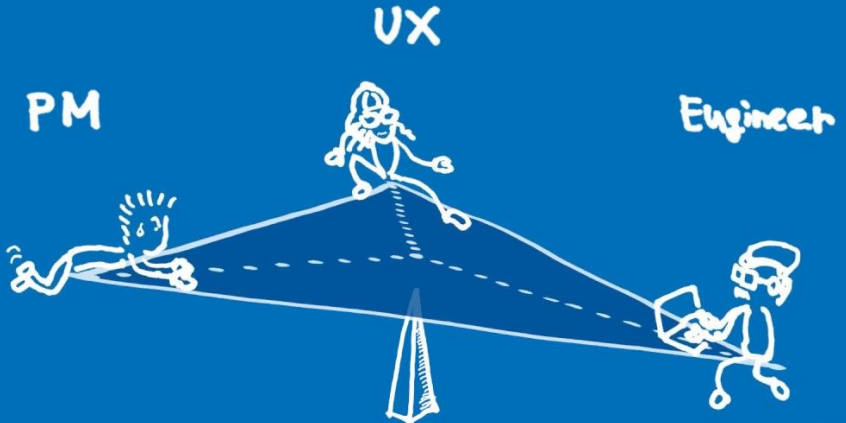
Cross-functional collaboration



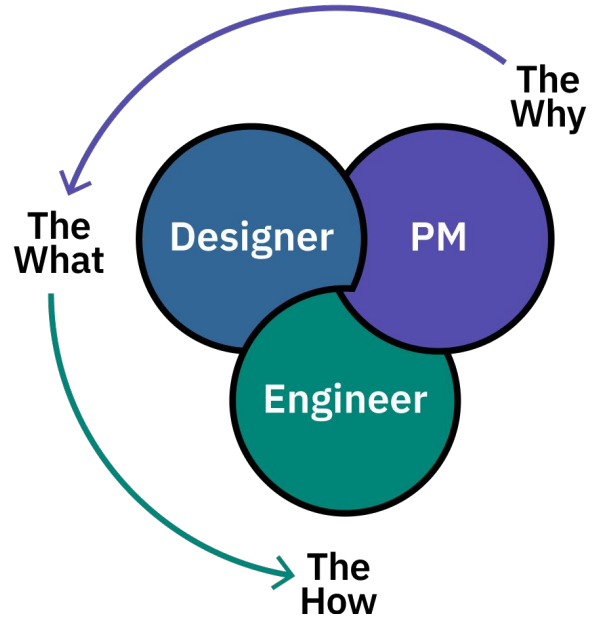
- Aligned with PMs on scope and client priorities (e.g., Unilever retention, Mapfre onboarding).
- Partnered with engineers, ensuring design intent met technical feasibility.
- Collaborated with legal, policy, and data science teams to embed compliance and trust into UX.

Start-up/Scale-up vs Large structures

with product roadmap and engineering feasibility.



Start-up



Larger structures

Work-in-progress



Research snippet

ADMIN

Who they are

Senior-level employees or leadership team

ROLES

Chief Data Officers (CDOs)

AI Governance Leads

Compliance Officers

other executives.

Overview

- These users oversee the AI systems and are responsible for ensuring compliance, reducing risks, and improving the overall AI quality.
- They may or may not have technical and regulatory knowledge
- Familiarity with organizational goals, risk management frameworks, and business strategies.
- Strong decision-making skills with the ability to balance technical complexity and business needs.

Goals

- Ensure the organization's AI systems align with regulatory and ethical standards.
- Oversee AI governance frameworks and ensure compliance across all projects.
- Mitigate risks associated with AI (e.g., bias, data security, model robustness).
- Continuously improve the quality, transparency, and accountability of AI systems across the enterprise.
- Efficiently manage resources and team capacity for AI-related projects.

Needs

- High-level dashboards with key metrics and insights on AI system performance, risks, and compliance.
- Regular reports that show trends, potential risks, and strategic recommendations.
- Clear visibility into the status of ongoing AI projects, audit logs, and compliance checks.
- Tools to define and adjust governance policies and risk frameworks at an organization-wide level.
- Ability to delegate tasks and monitor team progress efficiently.

Pain Points

- Difficulty overseeing multiple AI projects and teams at scale.
- Challenges in keeping up with evolving regulations and ensuring compliance across diverse AI models.
- Limited visibility into the root cause of AI failures or bias across projects.
- Lack of alignment between technical data and business impact when making decisions.
- Ensuring effective collaboration between cross-functional teams (AI engineers, legal, compliance, etc.).

Tasks

- Reviewing AI governance reports and compliance dashboards.
- Adjusting risk and governance frameworks as needed.
- Making decisions based on AI performance insights and compliance outcomes.
- Monitoring the progress of AI projects and ensuring resources are allocated properly.
- Approving key governance policies or escalating issues to senior leadership.
- Overseeing audits, model evaluations, and critical incident responses.

MEMBER (USER)

Who they are

Mid-level to junior-level employees

ROLES

Data Scientists

AI Engineers

Compliance Specialists

AI Project Managers

Overview

- These users are responsible for implementing and maintaining AI models within the organization and ensuring they adhere to governance and compliance rules.
- More focused on day-to-day operational tasks and project execution.
- Familiar with specific AI tools and frameworks, but less concerned with overall organizational strategy.
- Some understanding of compliance and regulatory standards as they relate to their projects, but may not be experts in governance.

Goals

- Successfully complete AI projects while ensuring they meet governance and compliance standards.
- Improve the performance, fairness, and reliability of AI models they work on.
- Identify and address risks, including bias and model drift, in their daily work.
- Collaborate with other team members to execute tasks efficiently.

Needs

- Clear and actionable tasks that directly relate to their AI projects.
- Detailed insights on their AI models' performance, fairness, and compliance.
- Notifications and alerts when their models exhibit potential risks or non-compliance issues.
- Guidance on how to resolve governance issues or optimize their AI systems.
- Access to project-specific documentation, guidelines, and best practices.

Pain Points

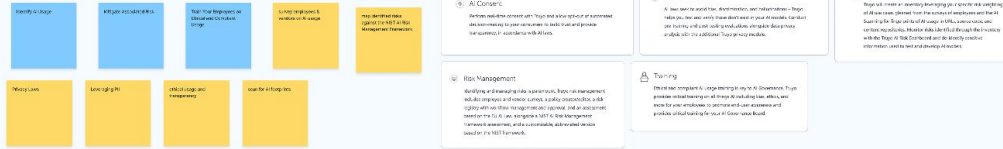
- Frustration with unclear or changing compliance requirements.
- Difficulty managing multiple AI models and projects while ensuring adherence to governance policies.
- Overwhelmed by manual compliance checks or model evaluations.

Tasks

- Running bias detection and mitigation tests on their models.
- Ensuring compliance with governance policies.
- Responding to alerts about non-compliance, bias, or performance degradation in models.
- Collaborating with other members on AI project execution and reporting.
- Attending regular check-ins to provide updates on project status and risks.

Research Snippet

Unique Value Proposition What makes this company unique?



truy

AI Governance - Help you understand how to use AI/ML responsibly, transparently, and ethically.

credo ai

AI Governance - Help you understand how to use AI/ML responsibly, transparently, and ethically.

By Use Case

- Generative AI Governance
- Vendor Risk Assessment
- AI Adoption Tracking
- Regulation Compliance
- Scalable AI Governance
- Audit AI Effects

By Regulations & Standards

- Colorado 5529-150
- DUI AI Act Readiness
- ISO/IEC 42001
- NIST AI RMF
- NYC Local Law No. 144

EU AI Act

Regulate your AI Systems

Align with AI Act Requirements

Monitor the AI Act's progress

GenAI

Track every GenAI use case

Apply GenAI Policy Packs

Align GenAI use case to your risk

third party vendors

Track third party vendors to ensure ethical supply

Define vendor risk categories

Monitor vendor risk scores

Manage vendor risk scores and alerts

DataRobot

AI Governance - Help you understand how to use AI/ML responsibly, transparently, and ethically.



Workshops snippet



findings from the meeting

painpoint: end user comes to their flow to submit all answers, and they need to do it again on IAI for assurance and EU AI act

needs: ask all information upfront, additional questions are asked if when it is necessary

painpoint: they have to come to IAI to archive so many projects cause there are no activities.

whenever there is an AI inventory, they always come to answer the questions in the Microsoft form

questions are not for assurance but for internal governance

the intention behind this is they don't people come to IAI before ideation is done, other designs are PoC, pilot and Live

they need to make sure they have enterprise architecture to align with their strategy

PRA ID?

also ask for legal team has been involved or not

insights: with questions that not get answered, people may come back to answer them, or drop.

painpoint: they have to come to IAI to archive so many projects cause there are no activities.

assurance stage is manually updated for reporting

comments area to update the status

Gen AI greenlist to see if a tool is approved by U, R is a first step, still need AI assurance

ideation

avoid answer same questions repeatedly

clearly state that the project needs to be PoC or after to process

build a good way for them to communicate with enterprise architecture and legal people

nice to have a comment/note feature to add extra info/tags

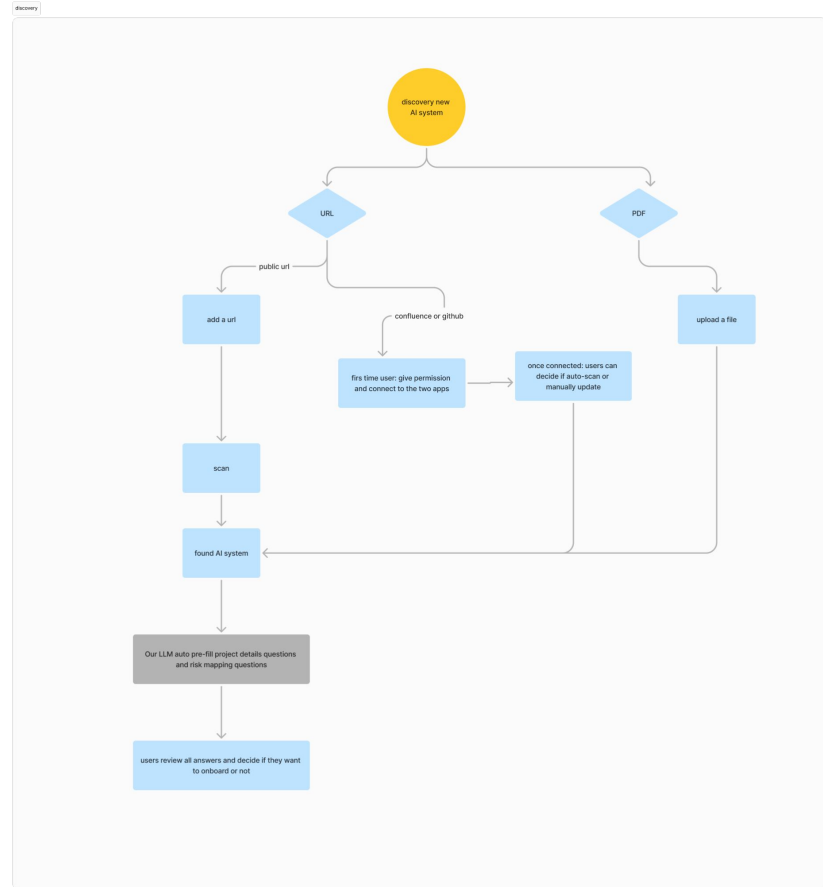
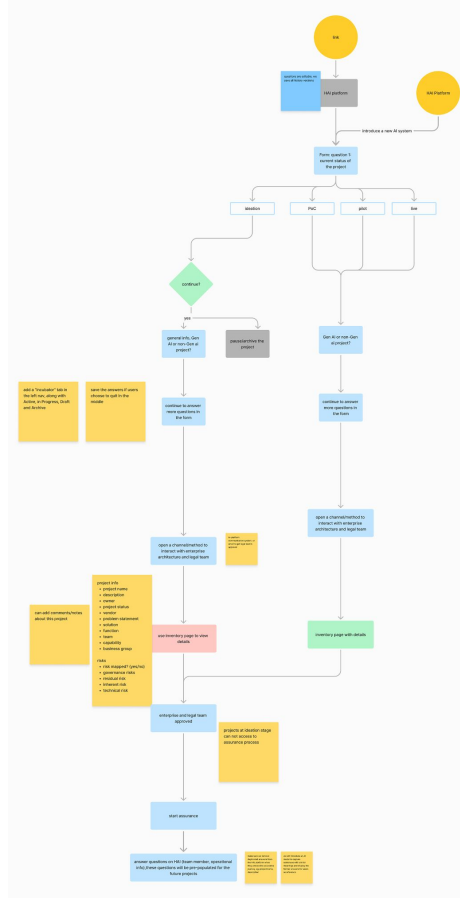
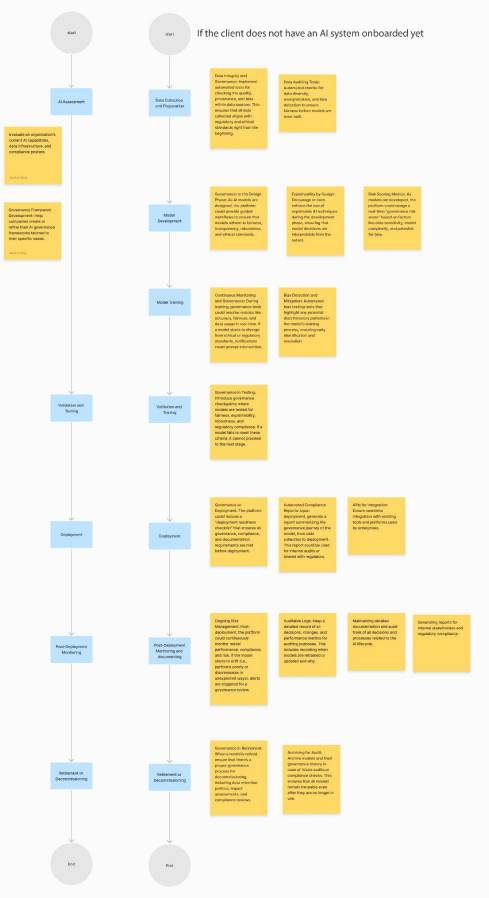
only ask necessary questions, limit the number of questions

separate projects which are at ideation stage and the stages after

automate some process, eg. greenlist/redlist apps, if redlist, provide explanations and usage instructions

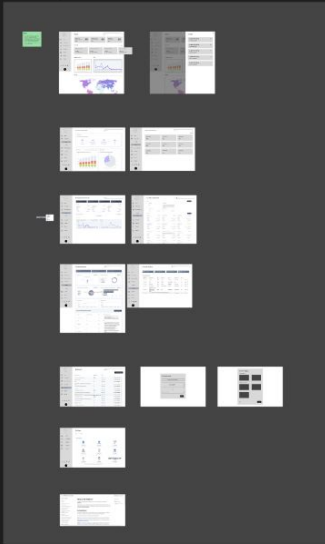
questions in the form are editable, and should be easy to edit. We should save history version for them.

Refining complex user flows

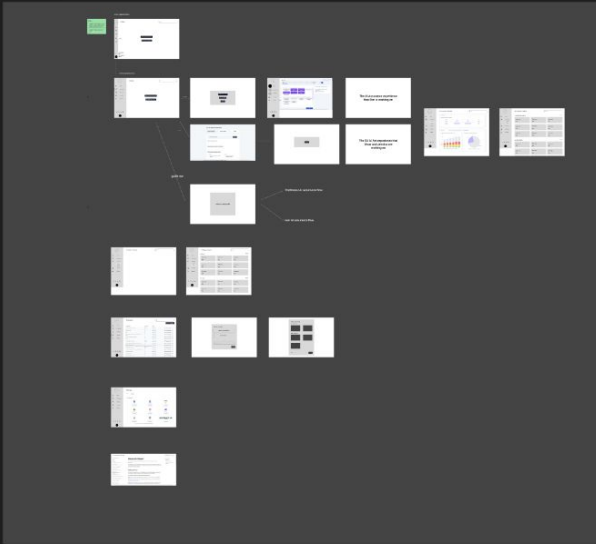


A breakdown of the phased approach

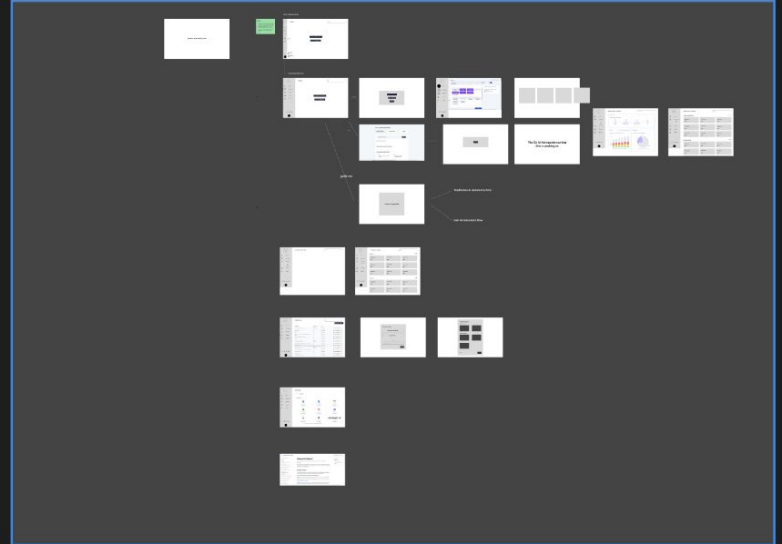
phase 1



phase 2



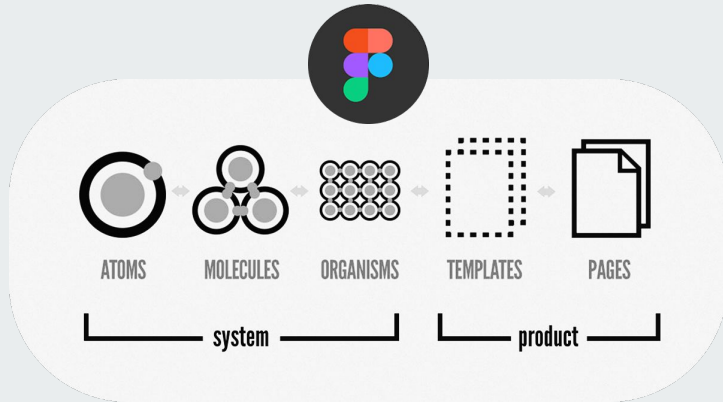
phase 3





The Design System

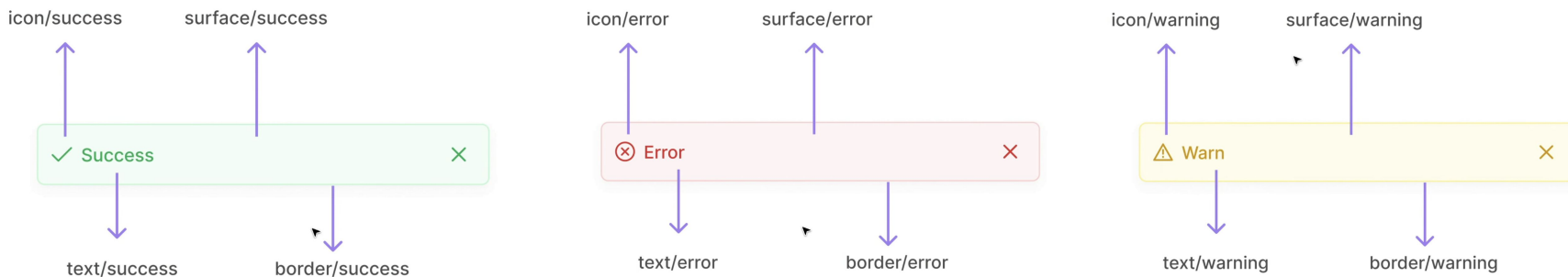
Figma Ninja



- **Designers:** Need scalable, reusable components and clear UX patterns to move faster.
- **Front-end Engineers:** Seek alignment with design and robust specs to reduce back-and-forth.
- **Product Managers:** Want consistency and velocity in shipping features.
- **Leadership:** scalability, governance, and elevating design quality across org.
- **End Users:** Need a product that flows seamlessly and purposefully

Design Tokens

- 100+ components created and rolled into scalable tokenised system
- Accelerated dev workflows and improved design consistency
- Enabled us to launch new products with minimal UI rework
- Adopted across all three platforms



Example: A breakdown of the 'success' colour tokens

Brand

Name	Value
green	
100	BAFFB9
200	81F57F
300	7ED37C
400	53C351
450	53C351
500	11820F
600	067504
700	055603



Alias

Name	Value
success	
100	green/100
200	green/200
300	green/300
400	green/400
500	green/500
600	green/600
700	green/700



Mapped

Name	Value
text	
success	success/500
icon	
success	success/500
border	
success	success/600
surface	
success	success/100

Example: multi-brand and theming

Alias			
All variables	53	Name	brand 1
primary		100	purple/100
success		200	purple/200
error		300	purple/300
neutral		400	purple/400
information		500	purple/500
warning		600	purple/600
border width		700	purple/700
border radius		700	purple/700

Mapped			
All variables	42	Name	Light
text		page	neutral/white
border		primary	primary/100
icon		success	success/100
surface		disabled	neutral/100
		on-disabled	neutral/200
		error	error/100
		warning	warning/100
		information	information/100
		action	primary/500
		action-hover	primary/600
		action-hover-2	primary/400

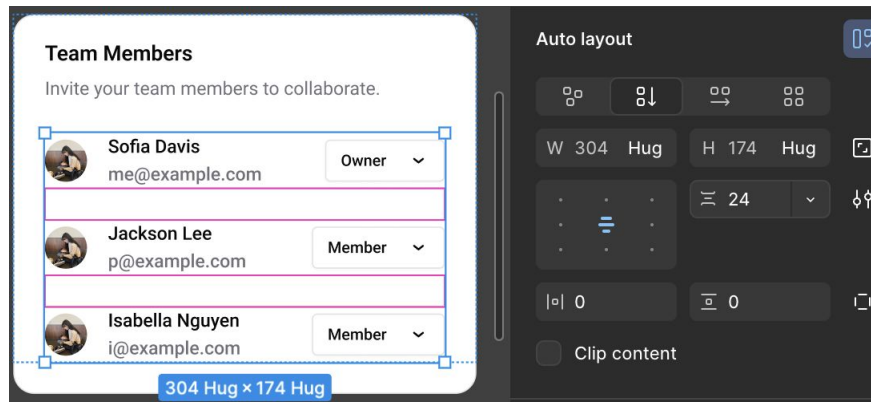
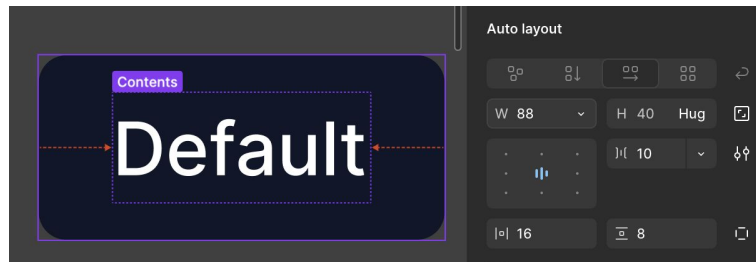
Why does this matter?



- Ensures wider team of designers in future have a consistent pattern to follow
- Much less likelihood of an inconsistent UI across teams
- Removes guesswork on components and styles
- Far easier to make system wide changes

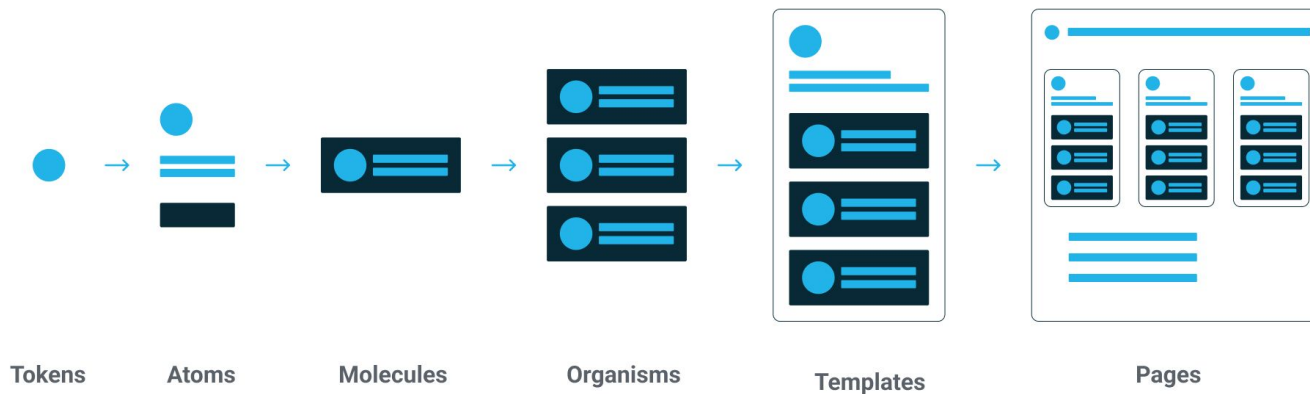
Auto-Layout

- Responsive by Design
- Scalable Component Libraries
- Mimics the flexbox/grid logic used by front-end engineers



Atomic design system

- 100+ components created and rolled into scalable tokenised system
- Accelerated dev workflows and improved design consistency
- Enabled us to launch new products with minimal UI rework
- Adopted across all three platforms



When to go Atomic



- Large organisations with complex design systems
- Teams that need a shared language for scalability
- Creating reusable components for different contexts



- Smaller teams or projects with limited scope
- Teams that need a shared language for multi-product projects
- When a simple, practical component library is sufficient

Build from scratch or build off an existing DS?



From scratch:

- Unique product domain or UI patterns
- When there's inconsistency or fragmentation across the product
- Targeting multiple squads or regions, needing consistency at scale

Existing DS:

- Startup or MVP and speed matters more than scale or 'beauty'
- Design needs are fairly standard
- Working inside a platform that already uses one

In the case of Holistic AI:



- Design system was built using shadcn as a base (standard functionality, less focus on UI flair more focus on functionality. We can add custom components on top: imaging viewer, workflow builder, etc.)
- Atomic design (scalable, thinking of product suite in future) if resources allow
- Tokenised library



The new platform(s)

The Core platform redesign



Centralised Ethical AI Management - The original core platform was robust but rigid. I evolved it for enterprise scalability. Just some of the improvements:

- Proposed and designed a new intake flow that automated onboarding via file uploads and integrations with Confluence, GitHub
- Designed the workflow and UX around key frameworks (e.g., EU AI Act), giving organizations a transparent view of their compliance and risk posture.

Result: significant decrease in manual effort and user frustration (Unilever reported a 60% increase in monthly AI system onboarding), helping retain Unilever as a long-term client and winning two new clients.

Before and after

Before

The 'Before' dashboard features a clean, modern design with a dark sidebar on the left. The main content area is light gray and includes a welcome message, four key metrics cards, a 'Catch up where you left off' section, an 'Attention! High Risk Projects' section, and an 'Overall Risk Level' section with a bar chart.

Hi, Stephen!
Welcome back. Here's what's happening today.

3 projects past due date
[See Projects](#)

7 high risk deployments
[See Deployments](#)

6 outstanding assessments
[See Assessments](#)

2 assigned tasks to complete
[See Tasks](#)

Catch up where you left off
Fraud Detection Model
Facial Recognition System
Customer Analysis - Churn Prediction

Attention! High Risk Projects

- Fraud Detection Model
- Facial Recognition - Security Access
- Candidate Assessment - Psychometric Testing

Overall Risk Level
The majority of your deployments last month were **LOW** risk.

Month	Low Risk	Medium Risk	High Risk
Jan	80	10	10
Feb	40	30	30
Mar	50	10	5
Apr	50	10	5
May	50	10	10
Jun	80	10	10

After

The 'After' dashboard is more data-rich and organized into a grid. It includes a 'Dashboards' section with three overview cards, a 'Global events affecting you' section with multiple event cards, and a dark sidebar on the left with more navigation options.

Dashboards

- 45%** EU AI ACT Readiness (47 Projects)
- 45%** AI Assurance (47 Projects)
- Ready** NIST Nist Risk Mapping (7 Projects)

Exposure 360

Exposed systems in sectors

- HR (9)
- Biometrics (6)
- Consumer (5)
- Automotive (4)

Exposed systems in sectors


- OpenAI (18)
- Cruise (14)
- Google (6)
- Amazon (5)


Global events affecting you [View more in the AI Tracker](#)





- Legislation** [In Progress](#)
VDAI Fine for failing to carry out DPIA
Filed 06 Feb 2023
Brazil No. 42005
- Incident** [In Progress](#)
VDAI Fine for failing to carry out DPIA
Filed 06 Feb 2023
Brazil No. 42005 ChatGTP
- Guidance**
VDAI Fine for failing to carry out DPIA
Filed 06 Feb 2023
Brazil
- Legislation** [In Progress](#)
VDAI Fine for failing to carry out DPIA
Filed 06 Feb 2023
Brazil No. 42005
- Guidance**
VDAI Fine for failing to carry out DPIA
Filed 06 Feb 2023
Brazil
- Incident** [In Progress](#)
VDAI Fine for failing to carry out DPIA
Filed 06 Feb 2023
Brazil No. 42005 ChatGTP

Navigation: Home, Projects, Archive, Signoffs, Inventory, Get Started, Dashboards, Signoffs, Global Exposure, INVENTORY, Reports, Monitoring, Discovery, PLATFORM SETTINGS, User Management, Rulebook, Connections, API Settings, Admin Panel, Help, Docs, Logout, User: Slim Deal Admin





The 'Get Started' screen

 Holistic AI Governance






Get Started 




-  Dashboards
-  Signoffs ^{*}
-  Global Exposure 



INVENTORY

-  Projects
-  Reports
-  Monitoring ^{*}
-  Discovery ^{*}


PLATFORM SETTINGS

-  User Management
-  Rulebook
-  Connections
-  API Settings
-  Admin Panel

 Help  Docs  Logout

 SD Sim Deol
Admin 

 **Get Started** - Get started by using featured responsible AI solution frameworks or modules


 **Intake** | Complete your onboarding, triage and assessment for your AI system. Catalogue project

- Frameworks**
- Compliance
- Regulations
- Risk


AI Assurance
Audit and monitor your LLMs to ensure they are risk free


EU AI Act Readiness
Get your AI System ready to comply with the EU AI Act with our AI Assurance Journey


NIST Risk Mapping
Fulfill key requirements of the NIST AI Risk Management Framework


NYC Bias Audit
Get your AEDT bias audited in accordance with NYC Local Law 144

Details

Select a solution to see more details.

Run a **quick** assessment on your system to determine what solution(s) are of relevance.

Assess system

Simplifying system onboarding

Holistic AI Governance

Get Started ↗

Dashboards

Signoffs*

Global Exposure ↗

INVENTORY

Projects

Reports

Monitoring*

Discovery*

PLATFORM SETTINGS

User Management

Rulebook

Connections

API Settings

Admin Panel

Help Docs Logout

SD Sim Deol Admin

< Cancel Register new project

Create Project 📄

Is the Solution for a new or existing project?

Next >






How would you like to onboard this project?

- New Project**
You will be required to create a new project in order to use this solution.

- Existing**
Search for and select an existing project to assign this solution to

Search projects

Recently discovered projects

↕ Date	System name	Data origin	Provider	Inherent Risk	
06 Feb 2023	System name example	 Azure	OpenAI	High Risk	Select project
06 Feb 2023	System name example	 AWS	OpenAI	High Risk	Select project
06 Feb 2023	System name example	Manual	OpenAI	High Risk	Select project
06 Feb 2023	System name example	Manual	OpenAI	High Risk	Select project
06 Feb 2023	System name example	 Confluence	OpenAI	High Risk	Select project
06 Feb 2023	System name example	 Confluence	OpenAI	High Risk	Select project
06 Feb 2023	System name example	 Confluence	OpenAI	High Risk	Select project

Collaborative view

Intake

New Intake

Search Intake...

Tags

Clear all tags



Ideation

16 | 8

Intake_item

last updated 30 jun 2022

Intake Tag + add tag

Adriano Koshiyama

16 | 8

Intake_item

last updated 30 jun 2022

Intake Tag + add tag

Adriano Koshiyama

16 | 8

Intake_item

last updated 30 jun 2022

Intake Tag + add tag

Adriano Koshiyama

16 | 8

Blocked

16 | 8

Intake_item

last updated 30 jun 2022

Intake Tag + add tag

Adriano Koshiyama

16 | 8

In Progress

16 | 8

Intake_item

last updated 30 jun 2022

Intake Tag + add tag

Adriano Koshiyama

16 | 8

Intake_item

last updated 30 jun 2022

Intake Tag + add tag

Adriano Koshiyama

16 | 8

Intake_item

last updated 30 jun 2022

Intake Tag + add tag

Adriano Koshiyama

16 | 8

Review

16 | 8

Intake_item

last updated 30 jun 2022

Intake Tag + add tag

Adriano Koshiyama

16 | 8

Intake_item

last updated 30 jun 2022

Intake Tag + add tag

Adriano Koshiyama

16 | 8

Intake_item

last updated 30 jun 2022

Intake Tag + add tag

Adriano Koshiyama

16 | 8

Collaborative view

Messages

Latest comments

Intake

Projects

Search projects or intake...

Intake_item

last updated 24 Sep 2024

Active Tag + add tag

Sim Deol

16

3

Project_01

last updated 30 Jun 2024

Active Tag + add tag

Adriano K

8

3

Project_02

last updated 08 Apr 2023

Active Tag + add tag

Nigel Kingsman

9

3

Project_03

last updated 24 Mar 2023

Active Tag + add tag

Sim Deol

2

3

Project_04

last updated 30 Jun 2022

SD

Sim Deol **REVIEWER**

24 Sep 2024 at 13:01

Commented on: Intake_item (Stage 2)

There seems to be a mistake on the project details. Please contact nigel.kingsman@holistical.com.

KY

Kasem Yassine **ASSIGNEE**

24 Sep 2024 at 18:00

Commented on: Intake_item (Stage 5)

I still need to amend 3 fields on the form. Can I get back to this after I sort some other things out?

SD

Sim Deol **REVIEWER**

24 Sep 2024 at 18:32

Commented on: Intake_item (Stage 2)

Okay but I also added Nigel into the loop so he can take a look while you work on those

Type comment

Submit

Team Details

Add user

Team Members

SD

Sim Deol **REVIEWER**

sim.deol@holistical.com

KY

Kasem Yassine **ASSIGNEE**

kasem.yassine@holistical.com

Intake Name

Intake_item

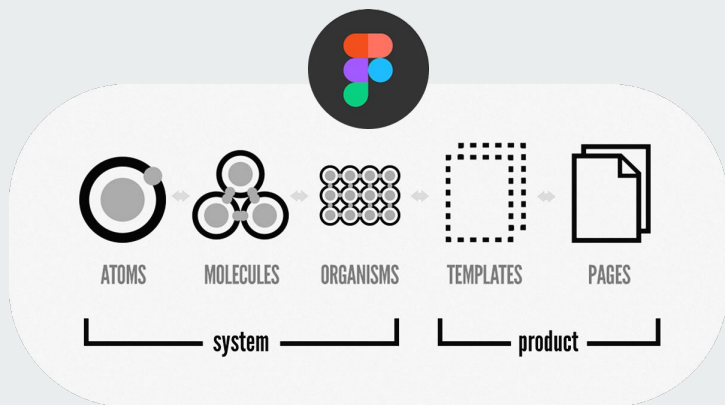
Last Updated

24 Sep 2024 at 18:32

Potential Risk

Low Risk

Scaling Design Across New Products



- With the new system in place, we expanded into two products in parallel:
 - **AI Tracker:** Policy monitoring for legal and compliance teams globally
 - **AI Safeguard:** LLM prompt auditing for risk, PII, hallucination, toxicity and much more
- System allowed immediate reuse and adaptation of components
- Designed with consistency, speed, and accessibility in mind

8 - The AI Tracker



Making Global AI Regulation Navigable - A real-time intelligence tool for monitoring AI legislation, policy, lawsuits, and more.

- Translated dense, multi-region AI law data into digestible UX
- Created dynamic filtering, search, and visualisation tools to help users digest large, complex data
- Introduced dynamic filters, visual maps, and intelligent grouping
- Designed for mixed audiences: legal users and operational decision-makers

Result: successful launch of a brand new product with over 1k sign-ups in one month.

Country / Region Clear

Search countries/regions

Events Clear Selection

124 Legislations

12 Guidance

76 Standards

35 Legal Actions

15 Incidents

12 Penalties

14 Regulations

Advanced Filters 0 Selected

Show Advanced Filters



Show Data 84 Events

AI Assistant

News

Blogs

World

Asia

Africa

Asia

Asia

Sectors

Subheading

Subheading

Technical

Subheading

Subheading

Risk Management

Subheading

Subheading

Explainer

Subheading

Subheading

Re-usable components built using Radix UI and Tailwind CSS

News Anti-discrimination



Nanda Smith, Sim Deol

07 Jun 2023

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud



Re-usable components built using Radix UI and Tailwind CSS

News Anti-discrimination



Nanda Smith, Sim Deol

07 Jun 2023

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud

Re-usable components built using Radix UI and Tailwind CSS

Blogs Tag



Nanda Smith, Sim Deol

07 Jun 2023

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud

Re-usable components built using Radix UI and Tailwind CSS

Blogs Tag



Nanda Smith, Sim Deol

07 Jun 2023

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute...

Hot Topics

OpenAI

14 Articles



ChatGPT

14 Articles



Google

14 Articles



Holistic AI Expert Community

Tracked Experts >

Open Source Resources



GitHub



Holistic AI Library

Upcoming Events

24h
Feb 2023

Event Title
Event detail copy to describe the event taking place

24h
Feb 2023

Event Title
Event detail copy to describe the event taking place

24h
Feb 2023

Event Title
Event detail copy to describe the event taking place

Get Started

Home

Inventory

Discovery 23 Nov

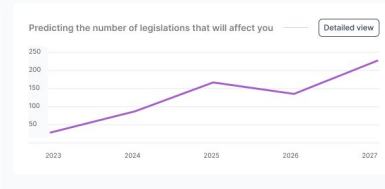
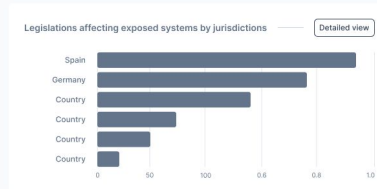
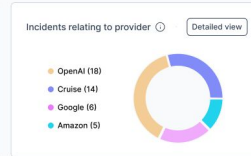
Integrations

Policy Center

Exposure Overview

New Exposure Assessment >

102 Systems | 47 Exposed Systems | 15 Events affecting you | \$3025 Total Cost



Events affecting your systems

All tracked systems >

8 Legislations affecting this system

5 Incidents

2 Penalties

5 Regulations

Event Name	Date Filed	Provider	Sector	Business Function	Region	
VDAl Fine for failing to carry out DPIA	06 Feb 2023	Provider	Sector	Business Funct	USA	+ Start Tracking
VDAl Fine for failing to carry out DPIA	06 Feb 2023	Provider	Sector	Business Funct	USA	+ Start Tracking
VDAl Fine for failing to carry out DPIA	06 Feb 2023	Provider	Sector	Business Funct	USA	+ Start Tracking
VDAl Fine for failing to carry out DPIA	06 Feb 2023	Provider	Sector	Business Funct	USA	+ Start Tracking

AI Assistant

- Home
- Saved Searches 14
- Tracked 12
- Saved Articles
- Policies

Saved Searches (14 updates)

Only show items with updates

Saved Search	Updates	Event Types	Country/Regions	Other filters	Email Alerts	
Legislations and Incidents in Spain	12 new events	Legislations, Incidents	EU, Spain	HR, Enacted	<input checked="" type="checkbox"/>	Remove
Incidents in UK	None	Incidents	EU, UK	HR, Enacted	<input checked="" type="checkbox"/>	Remove
Legislations and Incidents in Spain	12 new events	Legislations, Incidents	EU, Spain	HR, Enacted	<input checked="" type="checkbox"/>	Remove
Legislations and Incidents in Spain	12 new events	Legislations, Incidents	EU, Spain	HR, Enacted	<input checked="" type="checkbox"/>	Remove
Legislations and Incidents in Spain	12 new events	Legislations, Incidents	EU, Spain	HR, Enacted	<input checked="" type="checkbox"/>	Remove
Incidents in UK	None	Incidents	EU, UK	HR, Enacted	<input checked="" type="checkbox"/>	Remove
Incidents in UK	None	Incidents	EU, UK	HR, Enacted	<input checked="" type="checkbox"/>	Remove
Incidents in UK	None	Incidents	EU, UK	HR, Enacted	<input checked="" type="checkbox"/>	Remove
Incidents in UK	None	Incidents	EU, UK	HR, Enacted	<input checked="" type="checkbox"/>	Remove
Incidents in UK	None	Incidents	EU, UK	HR, Enacted	<input checked="" type="checkbox"/>	Remove
Incidents in UK	None	Incidents	EU, UK	HR, Enacted	<input checked="" type="checkbox"/>	Remove
Incidents in UK	None	Incidents	EU, UK	HR, Enacted	<input checked="" type="checkbox"/>	Remove
Incidents in UK	None	Incidents	EU, UK	HR, Enacted	<input checked="" type="checkbox"/>	Remove
Incidents in UK	None	Incidents	EU, UK	HR, Enacted	<input checked="" type="checkbox"/>	Remove
Incidents in UK	None	Incidents	EU, UK	HR, Enacted	<input checked="" type="checkbox"/>	Remove
Incidents in UK	None	Incidents	EU, UK	HR, Enacted	<input checked="" type="checkbox"/>	Remove

The AI Safeguard



Oversight for LLMs. A platform for monitoring and auditing the usage of Large Language Models (LLMs) within enterprises.

- Created real-time interface for AI prompt testing (toxicity, PII, hallucinations)
- Designed dashboards to visualise model outputs and flag risks
- Built flows for automated test creation, result reporting, and compliance tracking

Result: successful launch of a new product, with 2 major clients adopting the product.

Get Started

Dashboard

Audits

ROI

Monitoring

Chatbot

Governance

MONITORING

Settings

User Management

ADMIN

Settings

User Management

Dashboard

Overview Risks Usage Leaderboard

33

Audits

15

APIs

36

Users

\$7k

Cost Savings (ROI)

ROI by task type



Costs/Savings



Costs by model



Audits passed/failed by category



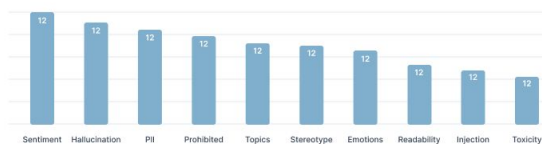
Audits overview



Exceptions over time



Exceptions overview



Top scoring LLMs (Global benchmarks)

LLM	Overall	Efficacy	Privacy	Robustness	Bias
GPT-4-turbo	0.1453	0.453	0.453	0.453	0.453

Get Started

Dashboard

Audits

ROI

Monitoring

Chatbot

Governance

MONITORING

Settings

User Management

ADMIN

Settings

User Management

Help

Docs

Logout

SD Sim Deol
Admin

< Cancel New Audit

Create



Tests

< Prev

Frameworks Use Cases Test Modules

Legal

Assesses the model's ability to execute legal reasoning, a crucial competence for AI systems engaged in legal analysis.

Test Name Test Name Test Name Test Name Test Name

Medical

Assesses the model's ability to execute medical reasoning, a crucial competence for AI systems engaged in medical analysis.

Test Name Test Name Test Name Test Name Test Name

RAG

Assesses the model's ability in utilizing Retrieval-Augmented Generation (RAG) techniques, important for LLM integrating external knowledge databases to enhance real-time information processing, decision-making, and response generation accuracy.

Test Name Test Name Test Name Test Name Test Name

HR

The model tests its HR benchmark (RJDB) and smaller language models to offer valuable resources for the HR domain and beyond.

Test Name Test Name Test Name Test Name Test Name

Governance and rollout strategy



- Created documentation within Figma with usage rules and naming logic
- Ran handoff walkthroughs and design-dev-pm syncs to promote adoption and prioritise which components needed developing first
- conducted system demos and shared libraries to support scale and quality
- Created a contribution model (proposal → review → build → release).
- Created clear and consistent guides, usage rules, and rationale for components and patterns.

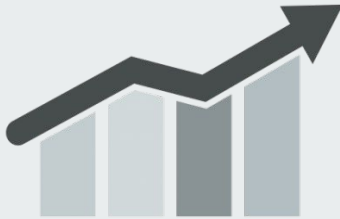
Leadership



I lead by amplifying strengths, removing roadblocks, and fostering true collaboration between design, product, and engineering for a shared goal



Measurable impact



- 60% faster onboarding via redesigned flows and automation
- Retention of Unilever as a key client
- Signed Mapfre as a major new customer
- System adopted by engineers across three frameworks
- Launched AI Tracker to over 1,000 sign-ups in the first month
- 2 major clients signing-up for the AI Safeguard

Reflection



- Wins: scaled UX, created shared language between teams, accelerated roadmap
- Challenges: legacy code, aligning stakeholders, no full-time researcher
- Next time: introduce design system governance earlier, invest in onboarding documentation sooner
- When the time is right: Introduce a better, more customised UI



Further exploration

- A singular platform with modules that allowed each organisation to tailor the platform to their needs
- AI-driven onboarding tips (e.g. GPT-powered assistant for workflows)
- Persistent shared project activity stream as a separate screen (cross-team collaboration view)
- Smart templates - creating project flows based on the project type, not just the user type



How about it? Let's work together and build the future together :)