

# How do you protect your marketplace, 24/7?

Guillaume Bouchard, CEO, Checkstep  
Digital Marketplace Association Conference,  
Prague, May 8th 2025

# The Hidden Epidemic of Marketplace Scams in Europe

**€100M+ in consumer losses annually**

**Widespread exploitation of unmoderated marketplaces**

- Facebook Marketplace: 34% of sampled listings were fraudulent
- Car scams on social platforms: Losses averaging nearly £1,000 per victim
- Seller scams increasing: Fake bank confirmations and spoofed IDs





# Content Moderation– the complexities



## Manual Review Overload

Compliance teams manually check thousands of listings, messages, and new profiles daily. Reputation and consumer safety at risk.



## GenAI Makes it Worse

A toy might breach safety rules in Germany, fail chemical guidelines in France, and have inadequate labelling for Italy.



## Complex Regulations

A toy might breach safety rules in Germany, fail chemical guidelines in France, and have inadequate labelling for Italy. Millions in fines.

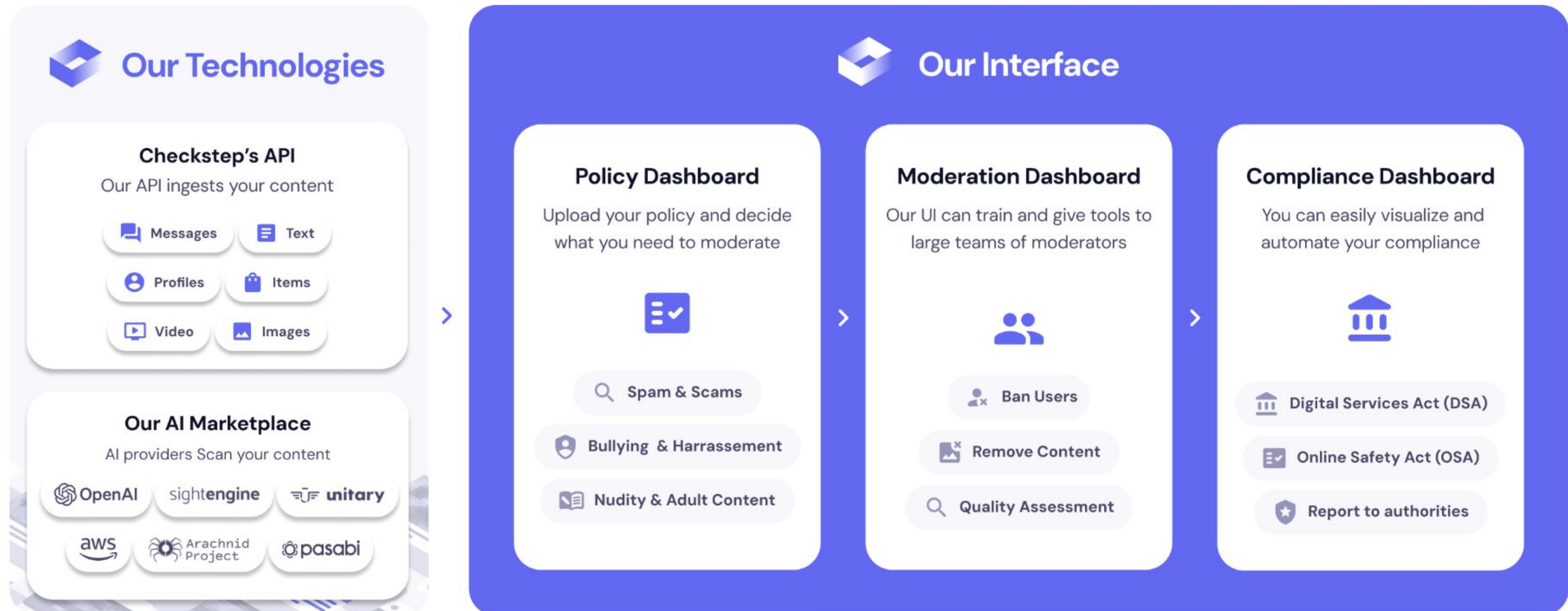
# The Compliance Maze

Market	Safety certificate / testing	Restricted chemicals	Age-grading & small parts	Connected-toy privacy / security	Child advertising	United Kingdom		CAP Code §5 – flag copy using “nag mum” or similar pester-power wording.	Toys Reg. Sch. 2 – flag images without age-warning icon or UK importer address.	GPSR 2005 / Safety Gate mirror – auto-flag SKU matching OPSS recall list.
United Kingdom	Toys (Safety) Regulations 2011 – flag listing with no UKCA/CE mark image or missing EN 71 certificate number. <a href="#">GOV.UK</a>	UK REACH / EN 71-3 – flag PVC toy if no lead/phthalate-free statement.	EN 71-1 – flag listings that show detachable ≤31 mm parts but claim “suitable under 3 yrs”.	UK GDPR & Children’s Code – flag smart toy that lacks a child-friendly privacy notice or says it shares data for ads.	CAP Code using “naïf pester-po					
United States	ASTM F963-23 & CPSIA §106 – flag if listing lacks CPSC tracking label or 3-part conformity certificate. <a href="#">ASTM International</a>   <a href="#">ASTM</a>	CPSIA §108 / 16 CFR 1307– flag “soft vinyl” toys with no phthalate-free note. <a href="#">U.S. Consumer Product Safety Commission</a>	16 CFR 1501 – flag photo that shows <31 mm pieces while listing says “ages 0 +”.	COPPA – flag connected toy page missing parental-consent flow or privacy-policy link.	FTC Act § ad copy li buy now v <13.	United States		FTC Act §5 & CARU – flag ad copy like “Only kids who buy now win!” pitched at <13.	FHSA label + ASTM projectile note – flag missing CHOKING WARNING or projectile-energy data.	CPSIA §15(b)– delist products matching CPSC recall feed immediately.
Germany	SpielzeugV / EU 2009/48 – flag CE mark absent or test report not German. <a href="#">Internal Market and SMEs</a>	REACH Annex XVII– flag PVC toys without “phthalate-free” claim.	EN 71-1 magnets & small parts – flag magnets shown yet “0-3 yrs” labelled.	GDPR + BSI IoT Label – flag smart toy page lacking security-update info.	UWG §5a or quizzes prompt ch					
France	Code de la Consommation Art. L523-1 – flag if Declaration-of-Conformity not uploaded in FR.	REACH + Arrêté cadmium – flag painted toys with no heavy-metal data.	Magnet flux decree – flag listings of magnetic toys missing ≤0.5 mT spec.	GDPR + CNIL guidelines– flag doll with mic but no DPIA summary.	ARPP En using imp tout de su	Germany		UWG §5a – flag ad games or quizzes that directly prompt child purchase.	German-language labelling duty – flag warnings not in German or too small (<1.2 mm x-height).	RAPEX/Safety Gate – cross-match EANs; flag if on current alert list. <a href="#">European Commission</a>
Australia	Consumer Product Safety Standard (Toys) 2020 – flag if AS/NZS 8124 cert absent.	ACCC Button-Battery Standard – flag toys with batteries visible in photos but no compliance note.	Toys ≤36 months Std – flag small balls shown without correct age warning.	Privacy Act 1988 Kids Code (draft) – flag smart plushies missing opt-in consent banner.	AANA Code for Marketing to Children – flag “collect all now!” copy.	Mandatory ‘0-3 yrs’ symbol ≥5 mm – flag thumbnails too small.	Product Safety Australia recall list – auto-flag matches; suspend seller.	unlicensed Marvel figures; require proof of brand rights.	“was-now” savings where “was” price used < reasonable period.	Product Stewardship Act – flag “recyclable plastic” claims without ARL label.
Canada	CCPSA Toy Regs SOR/2011-17 – flag listing lacking bilingual test cert.	Phthalates Regs SOR/2016-188 – flag PVC toys with no chemical disclosure.	Toy Regs small-parts gauge – flag items with small parts marked “safe for babies”.	PIPEDA / future CPPA – flag connected toy missing explicit parent consent message.	CAP Code + Québec ad ban– flag ads to under-13 in French markets.	Bilingual warning labels – flag “WARNING” only in English.	Health Canada recall feed – auto-flag matching SKUs.	Copyright Act notice-and-notice– flag anime figures that rights-holder flagged.	Competition Act price accuracy – flag misleading savings claims.	Plastics EPR (Ontario) – flag “recyclable” with no producer responsibility number.
India	BIS IS 9873-2019 – flag listing without BIS licence number.	Lead Paint Rules 2021 – flag coloured toys lacking lead statement.	IS 9873 small-parts spec – flag toys with small parts but “3 yrs +” claim.	DPDP Act 2023 – flag smart toy with no age-gate or guardian consent.	ASCI Children’s Ad Code – flag listings making unrealistic “increase IQ” claims.	Legal Metrology Rules – flag package images missing MRP & importer address.	BIS recall notice – auto-flag if toy model in BIS defective list.	Copyright Act & DPIIT raids – flag Hot Wheels clones or unbranded Lego bricks.	GST price display – flag HS-code mismatch or advertised GST-exclusive price.	E-waste & Plastic Waste Rules – flag electric ride-on labeled “eco” with no EPR number.
Japan	ST-Mark / CPSA – flag no ST-Mark or foreign cert only.	Food Sanitation Act – flag mouth-contact toys without migration test mention.	ST-2016 small-parts – flag gacha capsules labelled for <3 yrs.	APPI – flag robot pet storing images with no JP privacy link.	Premiums & Representations Act – flag gacha “guaranteed win” claims.	Household Goods Labelling Act – flag warnings not in kana for ≤3 yrs toys.	NITE recall list – auto-flag matched JAN codes.	Trademark Act – flag counterfeit anime figurines; match Customs watch-list.	Specified Commercial Transactions Act – flag missing tax-inclusive price.	Recycling of Small Appliances Act – flag lithium battery toys w/o disposal note.
Brazil	INMETRO Ordinance 302/2021 – flag listing w/o INMETRO seal photo.	ANVISA heavy-metal rules – flag painted toys no test data.	ABNT small-parts spec – flag magnets or small parts w/ “+6 m” claim.	LGPD – flag connected toy privacy policy not in Portuguese.	CONAR Code Criança – flag copy urging “peça já para seus pais!”.	Portuguese labelling & CNPJ – flag importer info missing.	SNVS recall system – auto-flag listed toys.	IP Law 9279 – flag knock-off mini-figures; seller must prove rights.	Consumer Protection Code pricing – flag unrealistic “60% off” anchor.	National Solid-Waste Policy – flag “green toy” claims lacking reverse-log proof.
China	CCC + GB 6675 – flag no CCC mark in images or no GB test report.	GB 6675.4 – flag soft plastic toys with no phthalate statement.	GB 6675.2 small parts – flag sharp-point toys marketed to toddlers.	PIPL 2021– flag smart doll listing lacking on-shore data storage note.	Advertising Law 2015 – flag “best toy ever” superlatives aimed at children.	GB 5296.5 labelling – flag warnings not in simplified Chinese.	SAMR recall rules 2024 – auto-flag recalled model IDs.	Trademark Law + Customs IP– flag “LOgo bricks” infringing brand.	Price Law – flag false discounts or missing VAT-inclusive price.	Circular 97 plastics pilot– flag “fully degradable” claim without GB test reference.



# Centralize your Moderation Operations

Policy, Detection, Moderation, Compliance and Analytics in one place





# Checkstep – how it works

## Regulatory Integration

Comprehensive database of country-specific regulations.

## Flexible Updates

On-the-fly policy changes and instant AI recalibration.



## AI Analysis

Instant scanning of products against relevant rules.

## Precise Flagging

Clear identification of breaches with specific reasons per country.

[Policy Dashboard Video](#)

## Detection

Easily detect bad actors.

### Bad Actor repository

Detect bad actors using playbooks at scale.



### Network & scoring

Identify fraud rings & investigate with our graph.



## Moderation

Moderate content in a few seconds.

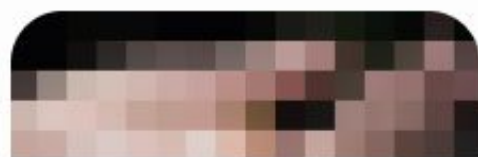
### Workflows

Easily block, report, suspend or remove content or users.



### Moderator wellness

Gray, blur or accelerate harmful content.



## Analysis

Make better informed decisions.

### Moderator training

A quality score to easily evaluate training needs.



### Analytics & insights

Get insights into trends for violations flagging & more.



## Compliance

Manage and enforce efficiently policies.

### Policy management

Manage policies, define violation thresholds.



### Compliance hub

Ensure your compliance with global regulations.



# The Turkish Metal Detector Ban

## Sudden Regulatory Change

Following an earthquake, Turkey banned online metal detector sales to protect archaeological sites.

## Rapid AI Adaptation

Within hours, Checkstep's AI updated to flag all metal detector listings for Turkish consumers.

## Clear Communication

Sellers received specific removal reasons, enabling quick compliance and protecting the platform.



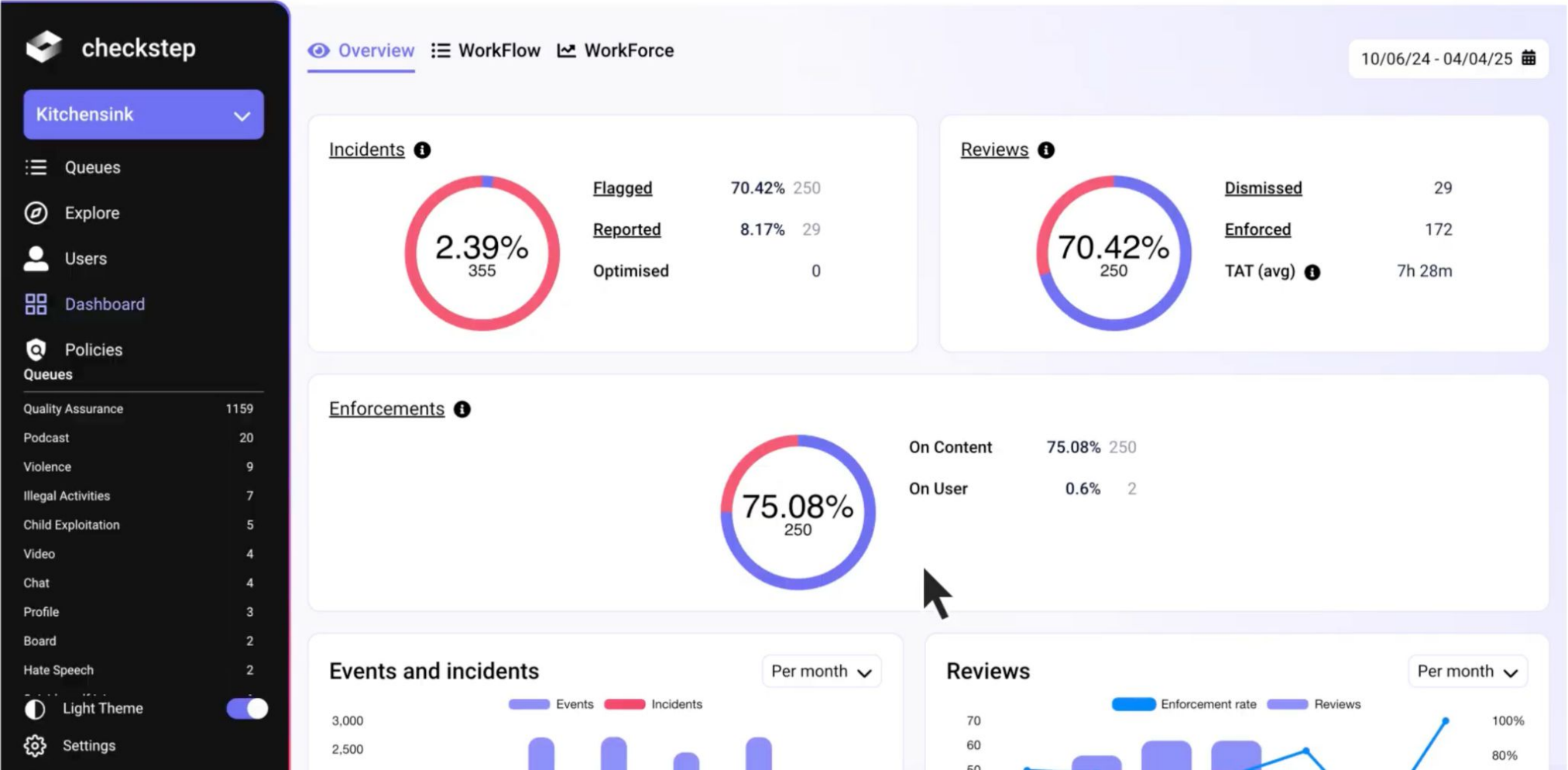


# A safer internet

[guillaume@checkstep.com](mailto:guillaume@checkstep.com)

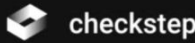


# Reporting Dashboard





# Policy Dashboard



Kitchensink

Queues

Explore

Users

Dashboard

Policies

Queues

Quality Assurance1159

Podcast20

Violence9

Illegal Activities7

Child Exploitation5

Video4

Chat4

Profile3

Board2

Hate Speech2

Light Theme

Settings

New policy

Hate Speech

Harassment

Child exploitation

Terrorism

Promotes suicide, self injury or eating disorders

Violence

Sexual Content

Illegal activities

Spam

Private information

Impersonation

Policy

Hate Speech - HTE

Archive

Description

Internal guidelines

Rules

Edit

History

Content that creates circumstances which lead to harassment, oppression and violence aimed towards a person or a group of people due to their affiliation to a protected category would count as Hate Speech on our platform. Our aim is to prevent users from having such unpleasant experiences just because they belong to a particular group/category. We also encourage users to criticize and raise awareness about Hate Speech on our platform in cases where their intent is clear.

We Checkstep do **NOT** allow the following:

- Violence & Violent speech - Targeting people based on their affiliation to a protected group
  - Statements that incite violence towards the protected group in any form including statements that are in the favor of causing death, harm or illness towards the protected category
  - Negative stereotypes about protected groups in both visual and written format
  - Comparison, generalization or harmful unfounded beliefs about a particular category of people in both visual and written format
  - Mocking/trivializing protected groups, people who were casualties of Hate crimes and sensitive events
  - Statements denying existence of protected group
- Generalizations of inferiority - Judging people who are a part of a protected group basis both in visual and written format
  - Physical capacity



# Policy Violation Confidence Levels

New policy

Hate Speech

Harassment

Child exploitation

Terrorism

Promotes suicide, self injury or eating disorders

Violence

Sexual Content

Illegal activities

Spam

Private information

Impersonation

Policy

Hate Speech - HTE

Archive

Description

Internal guidelines

Rules













History

New rule

Non violating

Send to review

Remove/Enforce

Label	Threshold	Restricted to
  Nazi Party	<div><div>&lt;50%</div><div>50-95%</div><div>&gt;95%</div></div>	▼
  White Supremacy	<div><div>&lt;50%</div><div>50-95%</div><div>&gt;95%</div></div>	▼
  Middle Finger	<div><div>&lt;50%</div><div>50-100%</div></div>	▼
  confederate_flag	<div><div>&lt;99%</div><div>99-100%</div></div>	▼
  pepe_frog	<div><div>&lt;99%</div><div>99-100%</div></div>	▼
  nazi_swastika	<div><div>&lt;99%</div><div>99-100%</div></div>	▼





# Moderation Dashboard

checkstep

Kitchensink

Queues

Explore

Users

Dashboard

Policies

Queues

Quality Assurance1159

Podcast20

Violence9

Illegal Activities7

Child Exploitation5

Video4

Chat4

Profile3

Board2

Hate Speech2

Light Theme

Settings

## Queues

Sort by: Please select

## Appeals

Open Appeals	0 12 hours	0 12-48 hours	1 48+ hours	: 1 Total	31 Mar 2025, 15:16	Review
--------------	---------------	------------------	----------------	--------------	--------------------	--------

## Escalations

Escalations	0 12 hours	0 12-48 hours	1 48+ hours	: 1 Total	16 Jan 2025, 11:40	Review
-------------	---------------	------------------	----------------	--------------	--------------------	--------


## Inquiries

Inquiries	0 12 hours	0 12-48 hours	1 48+ hours	: 1 Total	12 Mar 2025, 20:02	Review
-----------	---------------	------------------	----------------	--------------	--------------------	--------

## Posts

Suicide, self-injury	0 12 hours	0 12-48 hours	1 48+ hours	: 1 Total	10 Feb 2025, 13:24	Review
Harassment	0 12 hours	0 12-48 hours	1 48+ hours	: 1 Total	24 Feb 2025, 08:58	Review

# Content Timeline

 **checkstep**

Kitchensink

Queues

Explore

Users

Dashboard

Policies

Queues

Quality Assurance1159

Podcast20

Violence9

Illegal Activities7

Child Exploitation5

Video4

Chat4

Profile3

Board2

Hate Speech2

Light Theme

Settings

cm6z6fsde001ioxjudexf2c6t

Illegal Activities

Blur:  Gr

In Review

Timeline

Potential V

2

► ◀ ▶ ◂ ◃ 1x 0:00/0:00

Show translation

User

clhhsrdhi0000oxgn824g443r

178 Strikes

178 Major 0 Minor

!

10 Feb 2025, 14:57 - Incident open

Evaluation

Illegal activities

■

10 Feb 2025, 14:57 - New content version analysed

Post

2

► ◀ ▶ ◂ ◃ 1x 0:00/0:00

Show translation

Illegal activities

firearm98.52%







checkstep

# Moderation Dashboard

checkstep





# Compliance Dashboard

checkstep



# Reporting Dashboard

checkstep



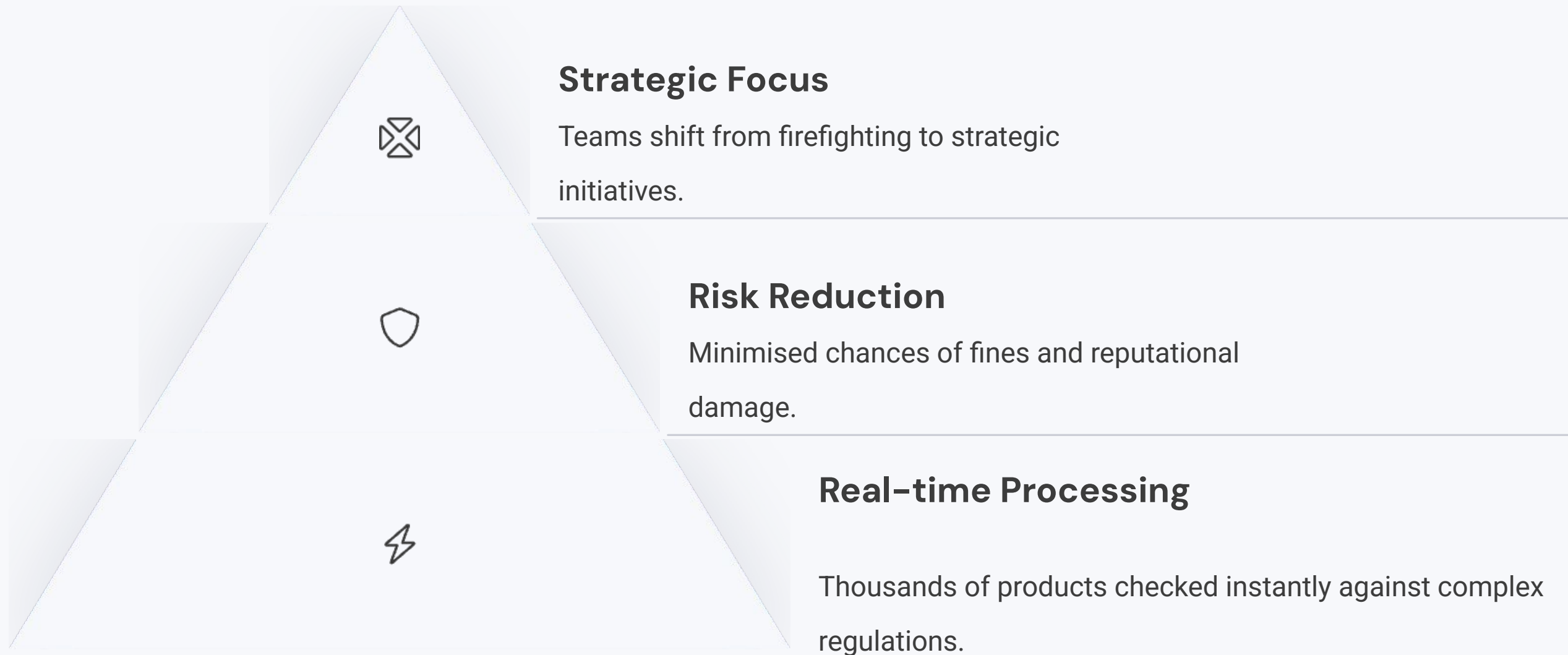
# Reporting Dashboard

checkstep






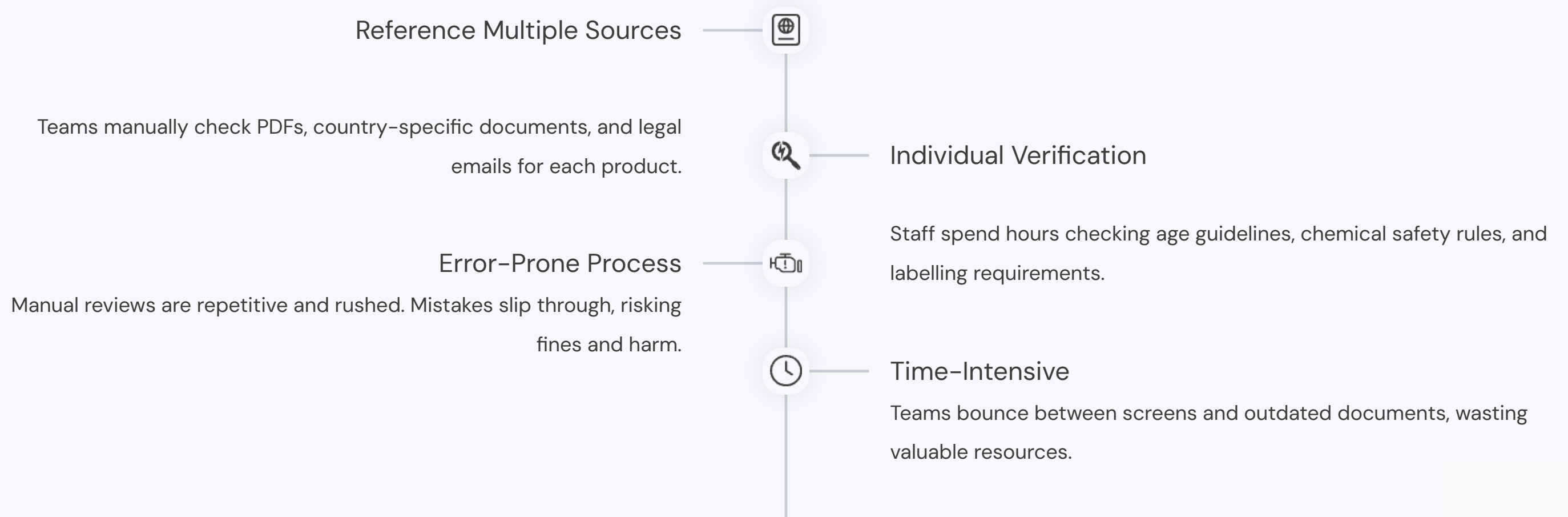
# Automation using AI



AI fundamentally changes what's possible in compliance. Accuracy becomes consistent and immediate, eliminating error-prone manual processes.



# Content Moderation – the current approach



# Checkstep's Approach



## Build In-house

Significant investment, years of development, continuous maintenance.



## Generic Solutions

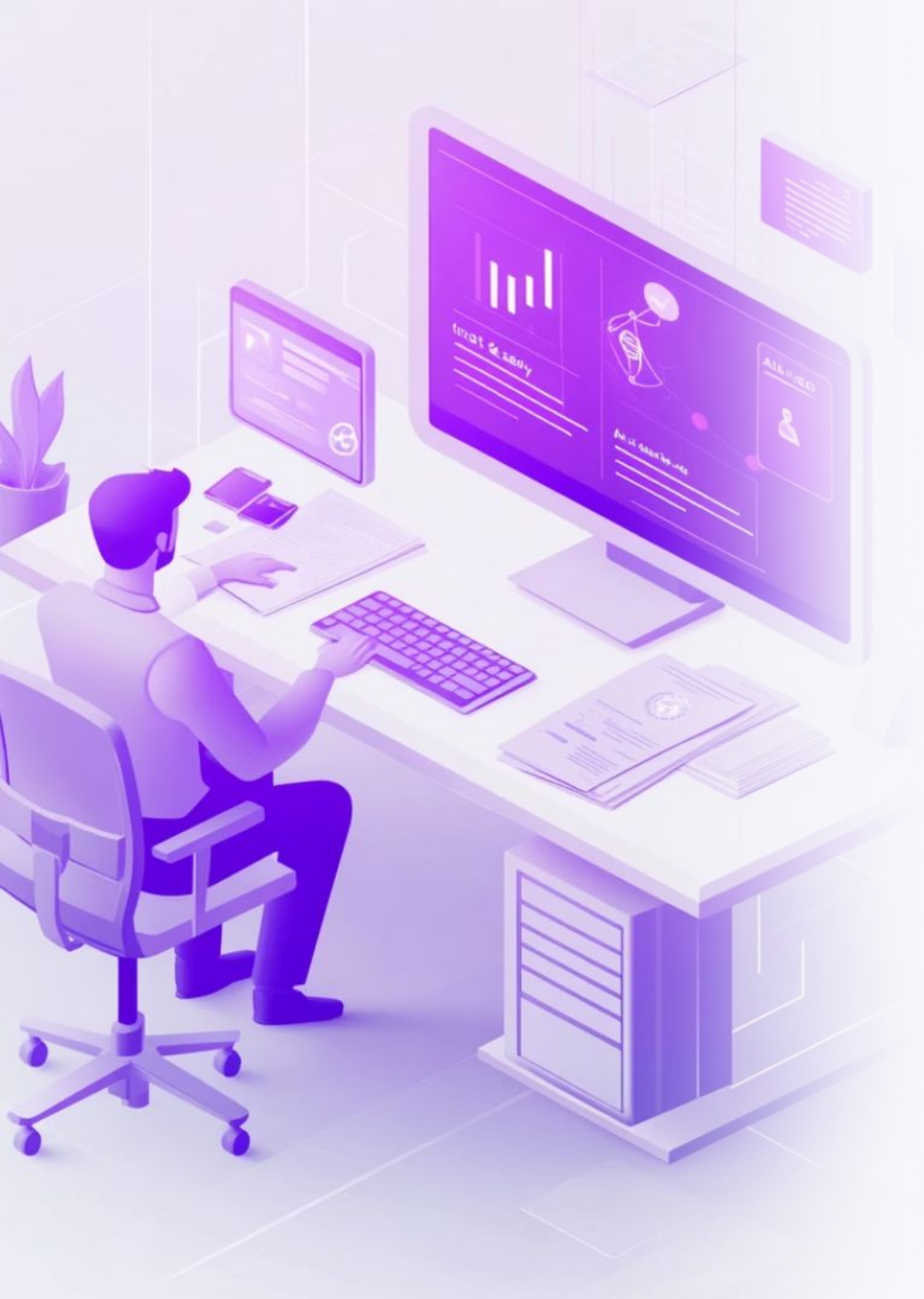
Competitors lack flexibility, speed, and regulatory specificity.



## Checkstep

Specialized, adaptable AI with rapid deployment and immediate ROI.

Our technology analyzes hundreds of products against multiple regulatory frameworks.







# Compliance Transformed

## **Reduce Operational Costs**

Cut manual review time by up to 90%. Free your team from repetitive tasks.

## **Minimise Regulatory Risk**

Consistent, accurate compliance checks across all products and markets.

## **Scale With Confidence**

Handle increasing product volumes and regulatory complexity with ease.

Checkstep delivers an AI solution that's both specialized and highly adaptable.



# Generative AI means more Content created

User-Generated Content and Generative AI are fueling the new digital economy by making existing markets bigger and opening new opportunities

**1 billion** number of user contributions per day on TripAdvisor

**USD 4.8 trillion** global User Generated Content platform market size in 2022, CAGR + 26.24%



**100 million** unique content creators active globally across major platforms

**1 billion** messages written in dating apps per month globally



**This has become so critical that it is being regulated globally.**

Over the past decade, governments worldwide have been increasingly introducing legislation on T&S issues. As platforms increase their global scope, companies will need to equip their T&S workflows with the appropriate localization. A T&S program that proactively detects and remedies issues is less likely to face regulatory and legal investigations, fines, and potential suspensions.



## Digital Services Act (DSA)

Enforced on 17 Feb 2024

More important than GDPR\* in the EU  
Makes content Moderation mandatory



## Online Safety Act

Enforced 2025

2025 UK regulation stricter set of rules than the DSA

\* GDPR fines for privacy breach were to \$4.5b in 2023 and continue to increase, but privacy laws are hard to enforce as consent is not visible. For moderation laws, one can check compliance just by using the service and checking how breaches are enforced. Fines will be 6% (in the EU) to 10% (in the UK) of the global revenue, larger than GDPR (5%).



### Reduce costs



Using AI to partly or fully moderate content helps reduce active handle time on average by 90% (Duco).

Improving operational efficiency reduces human moderation cost and **increase user engagement & growth.**

### Improve productivity



Clients of T&S SaaS experience on average a **38% improvement in operational efficiency** (Duco).

Giving the right tools to moderators such as queue prioritization or workflows increases productivity.

### Ensure compliance



Regulations like the DSA require more user transparency & protection.

This calls for new tools and streamlined operations to manage **age verification & transparency reporting.**

### Face threats, reduce risks



New threats are emerging and evolving constantly: counterfeit, spam, scams etc.

Using advanced technology like LLMs helps increase threats' detection, moderation and **decrease platform's risks.**





# A safer internet

[guillaume@checkstep.com](mailto:guillaume@checkstep.com)

