

Package ‘AIGovernance’

May 27, 2026

Type Package

Title Statistical Auditing and Governance Reporting for Employment AI Systems

Version 0.1.0

Description Provides statistical auditing, risk documentation, and reporting tools to support AI governance workflows for employment and hiring decision systems. Implements the EEOC four-fifths adverse impact rule (Equal Employment Opportunity Commission, 1978, <<https://www.ecfr.gov/current/title-29/subtitle-B/chapter-XIV/part-1607>>), NYC Local Law 144 bias audit requirements (New York City, 2023, <<https://www.nyc.gov/site/dca/about/automated-employment-decision-tools.page>>), and the AI Risk Management Framework checklist items from the National Institute of Standards and Technology (2023, <[doi:10.6028/NIST.AI.100-1](https://doi.org/10.6028/NIST.AI.100-1)>). Optionally supports EU AI Act high-risk classification (European Parliament and Council, 2024, <<https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32024R1689>>). The package does not provide legal advice or certify legal compliance; it is a statistical and documentation support tool.

License MIT + file LICENSE

Encoding UTF-8

LazyData true

RoxygenNote 7.3.3

Depends R (>= 4.1.0)

Imports cli, rlang, stats, tibble, dplyr

Suggests ggplot2, knitr, rmarkdown, testthat (>= 3.0.0)

VignetteBuilder knitr

Language en-US

URL <https://github.com/causalfragility-lab/AIGovernance>

BugReports <https://github.com/causalfragility-lab/AIGovernance/issues>

NeedsCompilation no

Author Subir Hait [aut, cre] (ORCID: <<https://orcid.org/0009-0004-9871-9677>>)

Maintainer Subir Hait <haitsubi@msu.edu>

Repository CRAN

Date/Publication 2026-05-27 09:50:07 UTC

Contents

aigov_adverse_impact	2
aigov_audit_nist	3
aigov_audit_nyc	4
aigov_build	6
aigov_checklist	7
aigov_classify	8
aigov_report	9
aigov_scope	10
hiring_sim	11
Index	12

aigov_adverse_impact *EEOC Adverse Impact Analysis (4/5ths Rule)*

Description

Computes selection rates by group and applies the EEOC Uniform Guidelines 4/5ths (80%) rule to assess adverse impact in employment selection procedures (EEOC, 1978).

The **Adverse Impact Ratio (AIR)** for each group is:

$$AIR_g = \frac{\text{selection rate}_g}{\text{selection rate}_{\text{ref}}}$$

A group is flagged for adverse impact when $AIR < 0.80$. The function also reports the two-proportion Z-test and Fisher's exact test p-values as supplementary statistics.

Note: The 4/5ths rule is a rule of thumb, not a bright-line legal standard. Small sample sizes reduce reliability. This function does not provide legal advice.

Usage

```
aigov_adverse_impact(gov, min_n = 30L)
```

Arguments

gov	An aigov object from aigov_build .
min_n	Integer. Minimum group sample size for a reliable AIR estimate. Groups below this threshold are flagged with a warning. Default 30.

Value

The input gov object with gov\$results\$adverse_impact appended, a tibble containing:

group Group label.
n Total applicants in group.
n_selected Number selected.
selection_rate Proportion selected.
AIR Adverse Impact Ratio vs reference group.
fourfifths_flag Logical: TRUE if AIR < 0.80.
z_stat Two-proportion Z statistic.
p_value Two-sided p-value from Z test.
fisher_p Fisher's exact test p-value.
small_n_flag Logical: TRUE if n < min_n.

References

Equal Employment Opportunity Commission (1978). Uniform guidelines on employee selection procedures. *Federal Register*, 43(166), 38295–38309.

Examples

```
data(hiring_sim)
gov <- aigov_build(hiring_sim, selected, race_ethnicity, ref_group = "White")
gov <- aigov_adverse_impact(gov)
gov$results$adverse_impact
```

aigov_audit_nist

NIST AI Risk Management Framework (AI RMF 1.0) Audit

Description

Implements a structured checklist audit aligned with the NIST AI Risk Management Framework (AI RMF 1.0, NIST, 2023). The RMF organises AI risk management into four core functions: **GOVERN**, **MAP**, **MEASURE**, and **MANAGE**.

This function presents the checklist items most relevant to employment AI systems and records user-supplied responses (or defaults to NA for items that cannot be verified from data alone).

Usage

```
aigov_audit_nist(gov, responses = list())
```

Arguments

gov	An aigov object from <code>aigov_build</code> .
responses	An optional named list of logical values (TRUE/ FALSE) for checklist items. Use <code>aigov_checklist(gov, "NIST_RMF")</code> to see item names. Items not supplied default to NA.

Value

The input gov object with `gov$results$nist_rmf` appended, containing:

`checklist` Named list with TRUE/FALSE/NA per item.

`scores` Per-function completion scores (proportion of confirmed items).

`overall_score` Overall confirmed proportion.

`verdict` One of "GREEN" (≥ 0.75), "AMBER" (0.50–0.74), or "RED" (< 0.50).

References

National Institute of Standards and Technology (2023). *Artificial Intelligence Risk Management Framework (AI RMF 1.0)*. NIST AI 100-1. [doi:10.6028/NIST.AI.1001](https://doi.org/10.6028/NIST.AI.1001)

Examples

```
data(hiring_sim)
gov <- aigov_build(hiring_sim, selected, race_ethnicity, ref_group = "White")
gov <- aigov_audit_nist(gov, responses = list(
  GOVERN_1_1 = TRUE,
  GOVERN_1_2 = TRUE,
  MAP_1_1    = TRUE
))
```

aigov_audit_nyc

NYC Local Law 144 Bias Audit Module

Description

Implements the statistical components of a NYC Local Law 144 (2023) bias audit for Automated Employment Decision Tools (AEDTs). The law requires employers using AEDTs to:

1. Conduct an annual independent bias audit.
2. Publish a summary of audit results on their website.
3. Provide advance notice to candidates/employees.

This function computes the required **impact ratio** statistics for each race/ethnicity and sex category and generates a publication-ready summary table matching the format expected in public disclosures.

The NYC LL144 impact ratio is:

$$IR_g = \frac{\text{selection rate}_g}{\text{selection rate of most-selected category}}$$

Note: LL144 uses the *most-selected category* (not a user-specified reference group) as the denominator – this differs from the EEOC approach. This function implements both.

Usage

```
aigov_audit_nyc(gov, use_most_selected = TRUE)
```

Arguments

`gov` An aigov object from `aigov_build`.

`use_most_selected` Logical. If TRUE (default), the denominator is the group with the highest selection rate (NYC LL144 standard). If FALSE, uses `gov$ref_group` (EEOC standard).

Value

The input `gov` object with `gov$results$nyc_ll144` appended, containing:

`impact_table` Tibble with selection rates and impact ratios.

`most_selected_group` The reference category used.

`disclosure_table` Formatted table for public disclosure.

`checklist` LL144 procedural checklist (named logical vector).

`verdict` Statistical verdict: "PASS" or "REVIEW".

References

New York City Local Law 144 of 2021 (effective January 1, 2023). NYC Department of Consumer and Worker Protection (DCWP). <https://www.nyc.gov/site/dca/about/automated-employment-decision-tools.page>

Examples

```
data(hiring_sim)
gov <- aigov_build(hiring_sim, selected, race_ethnicity, ref_group = "White",
                  frameworks = c("NYC_LL144"))
gov <- aigov_audit_nyc(gov)
```

aigov_build

*Build an AIGovernance audit object***Description**

Constructs an aigov object from employment decision data. This is the entry point for all auditing, classification, and reporting functions in the **AIGovernance** package.

Disclaimer: **AIGovernance** provides statistical and documentation support tools only. It does not provide legal advice and does not certify compliance with any law or regulation.

Usage

```
aigov_build(
  data,
  outcome,
  group,
  ref_group,
  frameworks = c("EEOC", "NYC_LL144", "NIST_RMF"),
  org_name = NULL,
  system_name = NULL,
  audit_date = NULL
)
```

Arguments

data	A data frame containing the employment decision records.
outcome	Unquoted column name of the binary decision variable (1 = selected / hired / advanced; 0 = not selected).
group	Unquoted column name of the protected-class variable (e.g., race/ethnicity, gender).
ref_group	Character string identifying the reference group (typically the highest-selection-rate group, e.g. "White" or "Male").
frameworks	Character vector of governance frameworks to activate. One or more of "EEOC", "NYC_LL144", "NIST_RMF", "EU_AI_Act". Default is c("EEOC", "NYC_LL144", "NIST_RMF").
org_name	Optional character string — organisation name for reports.
system_name	Optional character string — name of the AI system being audited (e.g., "Resume screening tool v2.1").
audit_date	Optional Date or character string (ISO format). Defaults to Sys.Date().

Value

An object of class "aigov" containing:

data The input data frame.

outcome Name of the outcome column (character).
group Name of the group column (character).
ref_group Reference group label.
frameworks Active frameworks.
org_name Organisation name.
system_name AI system name.
audit_date Audit date.
group_levels All observed group labels.
n_total Total number of records.

Examples

```
data(hiring_sim)
gov <- aigov_build(
  data      = hiring_sim,
  outcome   = selected,
  group     = race_ethnicity,
  ref_group = "White",
  frameworks = c("EEOC", "NYC_LL144", "NIST_RMF"),
  org_name  = "Acme Corp",
  system_name = "Resume Screening Tool v1.0"
)
print(gov)
```

aigov_checklist

Display Governance Checklist for a Framework

Description

Returns all checklist item names and descriptions for a given framework. Use the returned item names as keys in the responses argument of [aigov_audit_nist](#).

Usage

```
aigov_checklist(gov, framework)
```

Arguments

gov An aigov object (used only for class checking).
framework Character. One of "NYC_LL144", "EEOC", "NIST_RMF", or "EU_AI_Act".

Value

A tibble with columns item_id, function_area, and description.

Examples

```
data(hiring_sim)
gov <- aigov_build(hiring_sim, selected, race_ethnicity, ref_group = "White")
aigov_checklist(gov, "NYC_LL144")
aigov_checklist(gov, "NIST_RMF")
```

aigov_classify	<i>Classify AI System Risk Level</i>
----------------	--------------------------------------

Description

Classifies the AI system under two frameworks:

1. **EU AI Act (2024)** – Assigns one of four risk tiers: Unacceptable, High, Limited, or Minimal risk. Employment/worker-management AI is listed in Annex III as *High Risk*.
2. **NIST AI RMF** – Assigns a risk tier (1–4) based on impact on individuals’ rights and opportunities.

Usage

```
aigov_classify(
  gov,
  domain = "employment",
  makes_final_decision = TRUE,
  human_oversight = NA
)
```

Arguments

gov	An aigov object from aigov_build .
domain	Character. Application domain. One of "employment", "education", "credit", "housing", "healthcare", "law_enforcement", "other". Default "employment".
makes_final_decision	Logical. Does the AI system make or substantially influence a final employment decision? Default TRUE.
human_oversight	Logical. Is meaningful human review in place before the AI decision takes effect? Default NA (unknown).

Value

The input gov object with gov\$results\$risk_class appended, containing EU AI Act and NIST risk classifications with explanatory text.

References

European Parliament and Council (2024). Regulation (EU) 2024/1689 (EU AI Act). <https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai>

NIST (2023). *AI RMF 1.0*. doi:10.6028/NIST.AI.1001

Examples

```
data(hiring_sim)
gov <- aigov_build(hiring_sim, selected, race_ethnicity, ref_group = "White")
gov <- aigov_classify(gov, domain = "employment",
                    makes_final_decision = TRUE,
                    human_oversight = FALSE)
```

aigov_report

Generate an AI Governance Audit Report

Description

Produces a self-contained HTML (or plain-text) governance audit report from a completed aigov object. The report is suitable for internal documentation, legal review, or public disclosure (e.g., NYC LL144 website posting requirement).

Usage

```
aigov_report(gov, format = "html", output_file = NULL, open = TRUE)
```

Arguments

gov	A completed aigov object (at least one audit module run).
format	Character. Output format: "html" (default) or "text".
output_file	Optional character. File path for the output. If NULL, a temporary file is used and opened in the browser (HTML) or printed to console (text).
open	Logical. If TRUE (default) and format = "html", attempt to open the report in the default browser.

Value

Invisibly returns the path to the generated file.

Examples

```
data(hiring_sim)
gov <- aigov_build(hiring_sim, selected, race_ethnicity, ref_group = "White")
gov <- aigov_adverse_impact(gov)
gov <- aigov_audit_nyc(gov)
gov <- aigov_audit_nist(gov)
aigov_report(gov, format = "html")
```

aigov_scope

Determine Applicable Governance Frameworks

Description

Returns a summary of which governance frameworks apply to the AI system described in the aigov object, based on domain and jurisdiction.

Usage

```
aigov_scope(gov, domain = "employment", us_state = NULL)
```

Arguments

gov	An aigov object from aigov_build .
domain	Character. Application domain (default "employment").
us_state	Optional character. US state, e.g. "NY", "CO", "IL" for state-specific law flags.

Value

A tibble with columns framework, applies, jurisdiction, and note.

Examples

```
data(hiring_sim)
gov <- aigov_build(hiring_sim, selected, race_ethnicity, ref_group = "White")
aigov_scope(gov, domain = "employment", us_state = "NY")
```

`hiring_sim`*Simulated Employment Screening Dataset*

Description

A synthetic dataset of 500 job applicants processed by a hypothetical automated resume-screening tool. Generated for illustrative and testing purposes only. All individuals are fictional.

Usage`hiring_sim`**Format**

A data frame with 500 rows and 6 variables:

`applicant_id` Integer applicant identifier.

`race_ethnicity` Character. One of "White", "Black", "Hispanic", "Asian", "Other".

`gender` Character. "Male" or "Female".

`years_experience` Numeric. Years of relevant experience.

`score` Numeric. AI screening score (0–100).

`selected` Integer (0/1). Whether the applicant was advanced to the next stage (1 = selected, 0 = not selected).

Details

Selection probabilities were set to produce a realistic adverse impact pattern across race/ethnicity groups, consistent with published empirical ranges. The data are purely synthetic and do not represent any real organisation or hiring process.

Source

Simulated by the package authors for illustration purposes.

Examples

```
data(hiring_sim)
head(hiring_sim)
table(hiring_sim$race_ethnicity, hiring_sim$selected)
```

Index

* datasets

hiring_sim, 11

aigov_adverse_impact, 2

aigov_audit_nist, 3, 7

aigov_audit_nyc, 4

aigov_build, 2, 4, 5, 6, 8, 10

aigov_checklist, 7

aigov_classify, 8

aigov_report, 9

aigov_scope, 10

hiring_sim, 11