

Model Data Provenance and Attribute Exposure Assessment

Title VII of the Civil Rights Act, the ADA, and the ADEA

Document reference	OBC018FE8B6C
Operator	Northwind Financial Group
Regime assessed against	Title VII of the Civil Rights Act, the ADA, and the ADEA
Jurisdiction	United States
Statutory citation	42 USC 2000e-2, 42 USC 12112, 29 USC 623
System assessed	workforce-classifier, version 3
Assessment date	2026-07-28
Compared against	version 2, 2026-07-27
Generated	2026-07-29 08:07 UTC

1. System assessed

The features listed are those the deployed model actually consumed, read from the model registry rather than from documentation. Applicable decision context under this regime: Hiring, shortlisting, promotion, assignment, discipline and termination decisions about a worker or a candidate.

Registered model	workforce-classifier
Version in Production	3
Feature table	analytics_marts.workforce_features
Training run	b8dd3cb342454d16a2bbb00863070b3b
Feature count	7
Reported accuracy	0.7310

Features consumed

age, education_level_code, marital_status_code, household_size, census_region, vehicles_available, public_coverage_flag

2. Change since the previous assessment

Features added: public_coverage_flag

Reported model accuracy moved from 0.6955 to 0.7310, an increase of 0.0355. This is recorded because it establishes what performance monitoring would have shown at the time, and therefore whether any conclusion in this document could have been reached from performance monitoring alone.

3. Method

- For each attribute the regime restricts and this warehouse can express, a classifier is trained to recover that attribute from the feature set the deployed model consumes. The attribute under test is excluded from its own inputs, so the score is not the model reading the answer back to itself.
- Each measurement is repeated across three random seeds and reported with its spread, so that a change can be distinguished from a refit.
- A change counts as real when it exceeds the larger of 0.02 and 4.0 times the pooled spread of the two measurements. The threshold is therefore derived from observed variance and is not selected.

- Attributes the measurement cannot handle are reported in section 5 with the reason, and are never dropped silently.
- Column provenance is established by walking column level lineage in the metadata catalog. It is a property of the recorded graph, not an inference.

4. Measurements performed

Every measurement carried out in this assessment appears below, including those that found nothing. A record listing only the measurements that produced a finding cannot be distinguished from one where the remainder were never carried out.

Attribute	Group compared	Previous	This run	Change	Result
race	Black alone vs White alone	0.7105	0.7114	+0.0009	within noise
race	American Indian alone vs White alone	0.6543	0.6622	+0.0079	within noise
race	Asian alone vs White alone	0.6820	0.6827	+0.0007	within noise
race	Some other race vs White alone	0.7416	0.7460	+0.0044	within noise
race	Two or more races vs White alone	0.6217	0.6316	+0.0100	within noise
sex	Female vs Male	0.6025	0.6077	+0.0052	within noise
disability	With a disability vs Without a disability	0.7489	0.7883	+0.0394	EXCEEDED
citizenship	Born in a US territory vs Born in the United States	0.5978	0.5960	-0.0017	within noise
citizenship	Born abroad to American parents vs Born in the United States	0.5784	0.5847	+0.0063	within noise
citizenship	Naturalised citizen vs Born in the United States	0.7642	0.7643	+0.0001	within noise
citizenship	Not a citizen vs Born in the United States	0.8055	0.8059	+0.0004	within noise
nativity	Foreign born vs Native born	0.7614	0.7619	+0.0005	within noise
marital_status	Widowed vs Married	0.8378	0.8516	+0.0138	within noise
marital_status	Divorced vs Married	0.7755	0.7875	+0.0119	within noise
marital_status	Separated vs Married	0.7823	0.7911	+0.0088	within noise
marital_status	Never married vs Married	0.9088	0.9125	+0.0036	within noise

16 measurements carried out. 1 exceeded the threshold. 15 did not.

5. Restricted attributes not measurable

These attributes are restricted under this regime and were not measured. They are stated because an omitted attribute is otherwise indistinguishable from one that came back clean.

Attribute	Column	Why not measured
place_of_birth	place_of_birth_code	no categories declared, so there is no reference group to compare against
ancestry	ancestry_code	no categories declared, so there is no reference group to compare against
age	age	continuous, and the measurement compares two groups

Governance over these attributes rests on catalog tags alone for this assessment.

6. Findings

disability (With a disability against Without a disability)

Recoverability of this attribute from the model's feature set moved from 0.7489 to 0.7883, a change of +0.0394. That is 12 times the measured noise of 0.0033 and exceeds the threshold of 0.0200. Observed between version 2 and version 3.

Status under this regime: prohibited. **Citation:** 42 USC 12112, ADA. The ADA also restricts when disability related enquiry may occur at all, which is a stronger constraint than merely not using the answer.

Associated change: the model gained public_coverage_flag between the two assessments.

Provenance of the added column

Established by walking column level lineage in the catalog, furthest hop first. Where the column is named differently at a hop, that is the name it carries there.

Hops back	Column at that point
3	warehouse.analytics_staging.raw_person.pubcov
3	warehouse.public.raw_person.pubcov
2	warehouse.analytics_staging.stg_person.public_coverage_flag
1	warehouse.analytics_marts.dim_person.public_coverage_flag
the model reads it here analytics_marts.workforce_features.public_coverage_flag	

7. Record filed in the metadata catalog

The following incidents are open in the catalog against the feature table under this regime, read back from the catalog rather than from any local record. They are the durable form of this assessment.

Severity	Finding	Catalog reference
CRITICAL	age reaches workforce-classifier_3	urn:li:incident:ac291971-4ca5-41b1-9acb-0c58f9ed511a
HIGH	disability (With a disability) became easier to rebuild from workforce-classifier	urn:li:incident:107fdb6a-680d-45d7-a670-122736c40b50
LOW	marital_status_code reaches workforce-classifier_3	urn:li:incident:305c0844-755e-4fdb-a782-8b9c653be4da

8. Duties concerning the decision

The sections above concern which attributes may reach the model. The duties below concern what is owed to the person a decision is made about, which is a separate obligation that does not reduce to a list of restricted columns. For each, this record states what it supplies and what it does not, so that no part of a duty is treated as discharged by this document when it is not.

Records, and determination of adverse impact

Citation: 41 CFR 60-3 (Uniform Guidelines on Employee Selection Procedures), 60-3.4 and 60-3.15; 29 CFR 1602

Applies when: A selection procedure is used as a basis for an employment decision. A model that ranks or scores applicants is a selection procedure for this purpose.

The duty requires

- Maintain records of the impact the selection procedure has by race, sex and ethnic group, sufficient to determine whether it has adverse impact.
- Where adverse impact is found, either validate the procedure or otherwise justify its continued use.
- Retain documentation of any validity study relied upon, in the detail the Guidelines specify.

Supplied by this assessment

- A dated record, per model version, of what the selection procedure's inputs permit it to infer about the protected groups the Guidelines name.
- Detection of the case that is otherwise hardest to notice, where the information enters through a column that carries no protected label.
- Provenance for each input, which is part of documenting what the procedure actually consists of.

NOT supplied by this assessment

- × Impact ratios computed on actual selection outcomes. That needs the decisions made, not the model's inputs, and is the measure the Guidelines specify.
- × Any validity study.
- × The four fifths rule determination itself.

9. Scope and limitations

- This assessment concerns the data reaching a model and what that data permits the model to infer. It does not determine whether any decision was unlawful. Correlation with a protected attribute is not itself prohibited; the questions of harm, business necessity and less discriminatory alternatives are not answered here.
- Attributes not declared for this warehouse cannot be assessed. Where the regime restricts an attribute that the warehouse does not hold in a form this assessment can read, its absence from this document reflects the data, not a conclusion of compliance.
- The measurement establishes what is recoverable from the feature set. It does not establish that the model in fact relies on that information when producing an output.
- Findings rest on the metadata catalog being an accurate record of the pipeline. Lineage that was never ingested cannot be walked.
- The absolute level of recoverability is difficult to interpret in isolation and no threshold is asserted for it. What is reported is the change between two dated assessments of the same model.

10. Basis of preparation

Prepared automatically from the recorded measurement history and the metadata catalog on 29 July 2026. The statutory positions stated in this document are taken from the declared policy pack **employment_us** and were not determined by the tool. The pack should be verified against the current consolidation of the instrument before this document is relied upon.

Reviewed by

Date

Accountable owner

Date