

Uncovering discrimination in public fraud detection systems

Willy Tadema

Winterkongress der Digitalen Gesellschaft, 2025

What if...

Introducing myself:

- **Willy Tadema**
- **AI Governance & AI Ethics Lead** at Dutch National ICT Guild, Ministry of the Interior and Kingdom Relations
- **Standardization expert** in CEN-CENELEC's Joint Technical Committee on AI
- **Member of the Ethics Committee Data & Technology** in my home town Groningen
- **Speaking today in a personal capacity**
- Happy to connect on LinkedIn!





Dutch scandal serves as a warning for Europe over risks of using algorithms

The Dutch tax authority ruined thousands of lives after using an algorithm to spot suspected benefits fraud — and critics say there is little stopping it from happening again.

MARCH 29, 2022 6:14 PM CET
BY MELISSA HEIKKILÄ

NL#TIMES

TUESDAY, 12 NOVEMBER 2024 - 07:00

Government to refund over 10,000 students over discriminatory DUO fraud checks

In a landmark decision, the Dutch government has pledged to refund over 10,000 students who were unjustly flagged for student finance fraud by an algorithm developed by the Education Executive Agency (DUO).

Government in Netherlands Resigns After Benefit Scandal

A parliamentary report concluded that tax authorities unfairly targeted poor families over child care benefits. Prime Minister Mark Rutte and his entire cabinet stepped down.

The Guardian

Artificial intelligence
(AI)

Welfare surveillance system violates human rights, Dutch court rules

Government told to halt use of AI to detect fraud in decision hailed by privacy campaigners

BY MATT BURGESS, EVALINE SCHOT, AND GABRIEL GEIGER SECURITY MAR 6, 2023 7:00 AM

This Algorithm Could Ruin Your Life

A system used by the Dutch city of Rotterdam attempted to rank people based on their risk of fraud. The results were troubling.

Lighthouse Reports @LHreports

Serious impact in Sweden where the welfare agency has been called before parliament to answer questions over discriminatory algorithms used to profile benefits recipients

Post vertalen



Efter SvD:s granskning – MP kräver AI-garantier

Van svd.se

9:38 a.m. · 30 nov. 2024 · 986 Weergaven

The Guardian

The innocent have paid a high price for the Post Office scandal. The guilty have not

Marina Hyde



It was the word of hundreds of Post Office workers against a faulty computer system. Guess who was believed?



November 12, 2024

Denmark: AI-powered welfare system fuels mass surveillance and risks discriminating against marginalized groups – report

voxeurop
English

Serbia, a laboratory for algorithmic discrimination

The use of algorithms to optimise the allocation of social benefits is commonplace in Europe. In Serbia, data collected from the population is used to target certain groups, including the Roma. But voices are emerging in civil society to denounce this system.

Published on 15 February 2023 at 15:54

Federico Caruso - Osservatorio Balcani e Caucaso Transeuropa (Trento)

WIRED

MORGAN MEAKER

BUSINESS OCT 16, 2024 9:37 AM

Algorithms Policed Welfare Systems For Years. Now They're Under Fire for Bias

Human rights groups have launched a new legal challenge against the use of algorithms to detect error and fraud in France's welfare system, amid claims that single mothers are disproportionately affected.

BBC

Robodebt: Illegal Australian welfare hunt drove people to despair

7 July 2023

Frances Mao
BBC News



Recurrent patterns:

- Political push for a hardline approach
- Emphasis on cost efficiency
- Key decisions made by tech teams or consultants
- Lack of stakeholder engagement
- Ethical, design, and implementation blind spots
- AI mystification blocking critical reflection
- Affected individuals lacking information and recourse



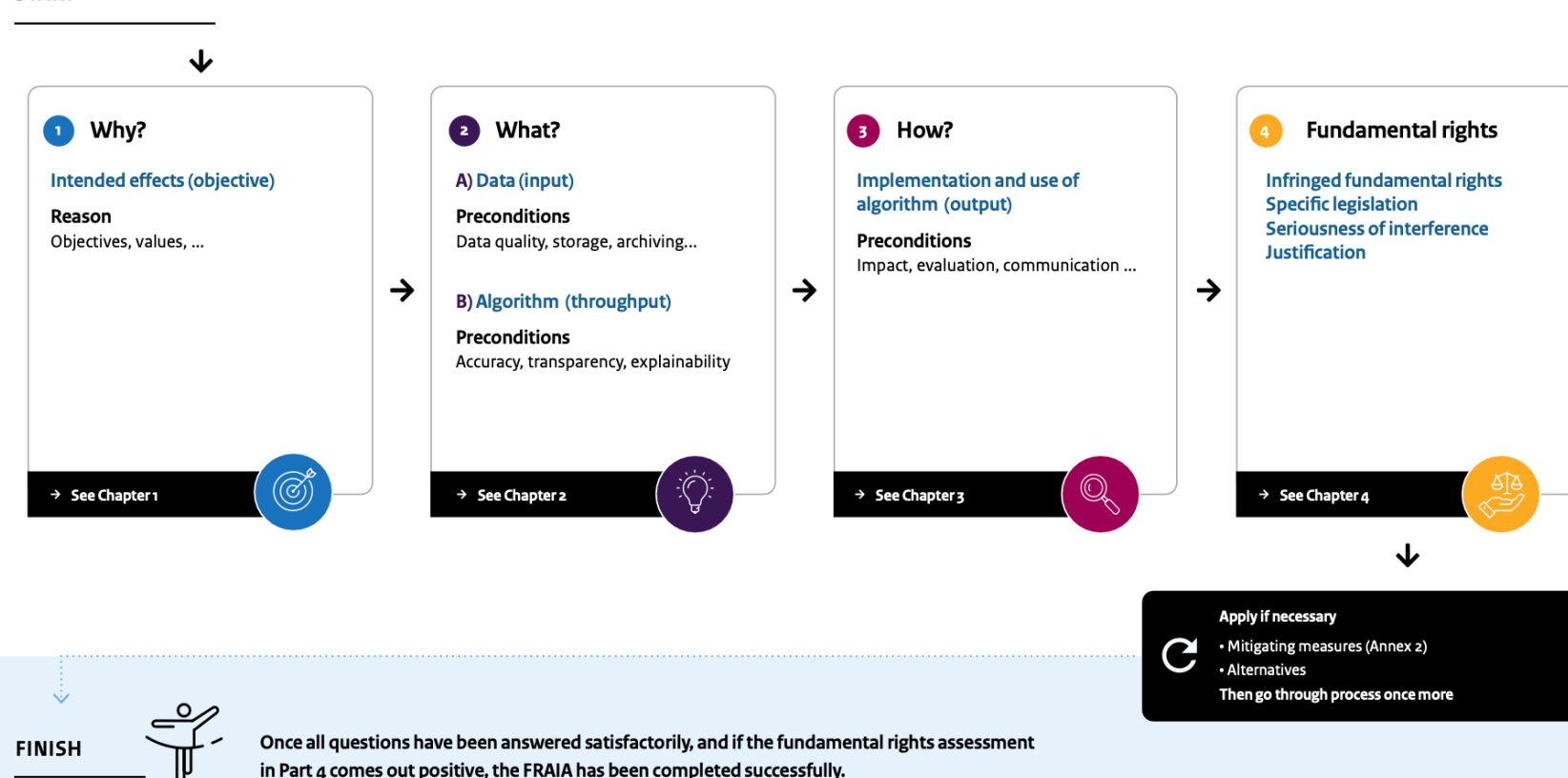
Impact Assessment

Fundamental rights and algorithms



FRAIA flow chart

START





AI regulation: Federal Council to ratify Council of Europe Convention

Bern, 12.02.2025 - Switzerland intends to ratify the Council of Europe Convention on Artificial Intelligence (AI) and to make the necessary amendments to Swiss law.

The **FRAIA** as implementation of the **HUDERIA**,
the Council of Europe's human rights impact assessment

Council of Europe adopts Turing-developed human rights risk and impact assessment for AI systems

Thursday 05 Dec 2024

The Algorithm Register of the Dutch government

Find one of the 773 algorithms

Search by word, theme or organisation

Search

Choose a organisation type

| | |
|---|-----|
| Municipality | 517 |
| Organizational unit | 52 |
| Agency | 46 |
| Independent administrative body | 43 |
| Province | 31 |
| Ministry | 28 |
| Regional cooperation body | 24 |

Show more

Choose a publicationcategory

| | |
|--------------------------------------|-----|
| Other algorithms | 378 |
| Impactful algorithms | 370 |
| High-Risk AI-system | 27 |

Choose a impact assessment

| | |
|---|-----|
| Geen | 545 |
| Data Protection Impact Assessment (DPIA) | 174 |
| Overig | 95 |
| Impact Assessment Mensenrechten Algoritmes (IAMA) | 18 |

775 results found

< 1 2 3 4 ... 78 >

Download all algorithms

Anonymise

Municipality of Veenendaal

This algorithm helps employees of the Municipality of Veenendaal identify and anonymise privacy-sensitive information in documents.

[Read more](#)

Last change on 24th of February 2025, at 13:57 (CET) | Publication Standard 1.0

Publication category

Other algorithms

Impact assessment

DPIA

Status

In use

Virtual municipality assistant Gem

Municipality of West Betuwe

The virtual assistant Gem is a digital help on West Betuwe's website for answering civic and business questions. Using the chat button on our websites, you can talk to Gem.

[Read more](#)

Last change on 20th of February 2025, at 13:29 (CET) | Publication Standard 1.0

Publication category

Other algorithms

Impact assessment

Field not filled in.

Status

Out of use

The 5 ★ model for transparent algorithms

★ **Published** Publish online what the algorithm should do

★ ★ **Explained** Describe what has been done to ensure that the algorithm does what it is supposed to do

★ ★ ★ **Verified** Share results showing that the algorithm does what it is supposed to do

★ ★ ★ ★ **Testable** Make it possible to test the algorithm

★ ★ ★ ★ ★ **Open** Open up the algorithm completely





Public standard
profiling algorithms

Public standard profiling algorithms

Responsible profiling algorithms through

- **quantitative safe guards**,
- **qualitative safe guards** and
- **organizational control measures**

based on Dutch government use cases.



Public standard profiling algorithms

Step 1

Contextualise and document proposed profiling method

Data points



+

Contextualize and document

- > Problem analysis + legitimize aim pursued
- > Identify vulnerable groups
- > Specify performance requirements

Step 2

Qualitative assessment
profiling characteristics



Use existing
sources

- Jurisprudence
- Algo-prudence
- Regulatory guidance

or



Ineligible

Grounds for exclusion, e.g.,

- Legally forbidden
- No relation with legitimate aim pursued
- Proxy characteristic
- Subjective
- Subject to change

Decide together
with diverse
stakeholders

Step 3

Quantitative assessment
profiling characteristics

Candidate
profiling characteristics



Random sample + Hypothesis test

+ Bias measurement with
internally available variables



Ineligible



Conditionally
eligible

Candidate profiling characteristics

Step 4

Profile composition



Composed by subject
matter experts

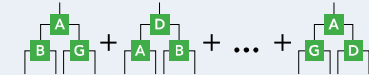


Profile



Profile

Expert
check



If explanation requirements
can be satisfied

Step 5

Implementation

- > Apply risk profiling, complemented with random sampling and signal-driven selection
- > Establish complaint and appeal procedure

Human-driven decision-making

- > Formal decision made by subject matter experts
- > Draft work instructions for manual review carefully

Step 6

Evaluate and monitor

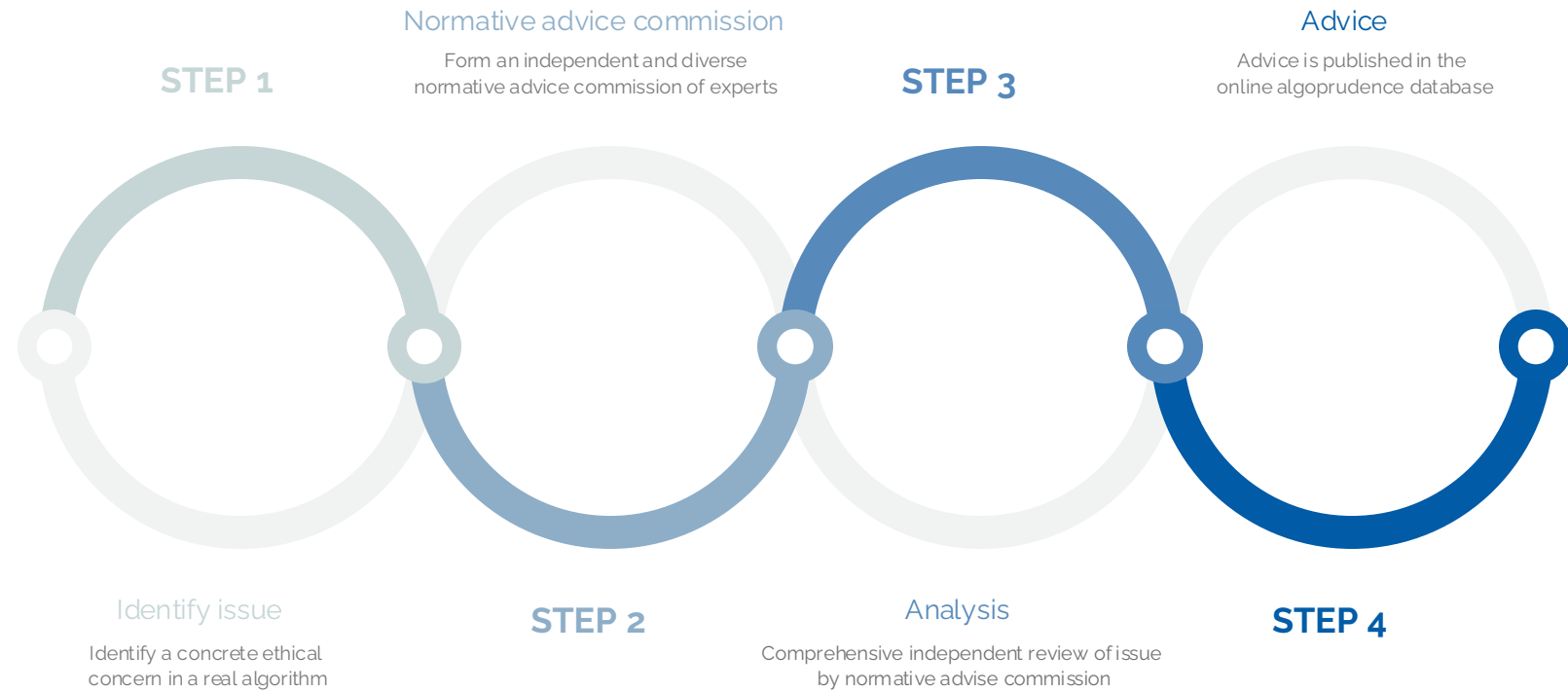
- > Periodic re-examination, e.g., performance requirements
- > Determine periodically (dis)parate impact on (vulnerable) groups with help of protected labels
- > Test conditional significance in applied model



AlgoPrudence: Jurisprudence for algorithms

AlgoPrudence

bottom-up, case-based, normative advice for algorithms



Example of **algotrudence**

Case: risk profiling for social welfare re-examination

Eligible criteria

| | | |
|--|---|---|
| Age | |  |
| No show at appointment with municipality |  |  |
| Reminders for providing information |  |  |
| Participation in trajectory to work (training, workplace, social duty) |  |   |
| Type of living (cohabitation, living together) |  |  |
| Cost sharing |  |  |

Legend

| | |
|--|---|
|  Unlawful |  Subject to change |
|  Proxy discrimination |  Subjective |
|  Connection with aim pursued |  Unclear definition |
|  No connection with aim pursued |  Manageable risks |

Ineligible criteria

| | |
|---|---|
| ZIP code, city district |  |
| Sex, gender |  |
| Reason for appointment with municipality (annual meeting, intake) |  |
| Type of contact (mail, phone, text, post) |  |
| Literacy rate |  |
| ADHD |  |
| Mental health services |  |
| Number of children |  |
| Sectoral (work) experience (hospitality, construction, logistics) |  |
| Assertiveness |   |
| Professional appearance |  |

■ NIEUWS

Rechtspraak wil driekwart van de vonnissen online publiceren

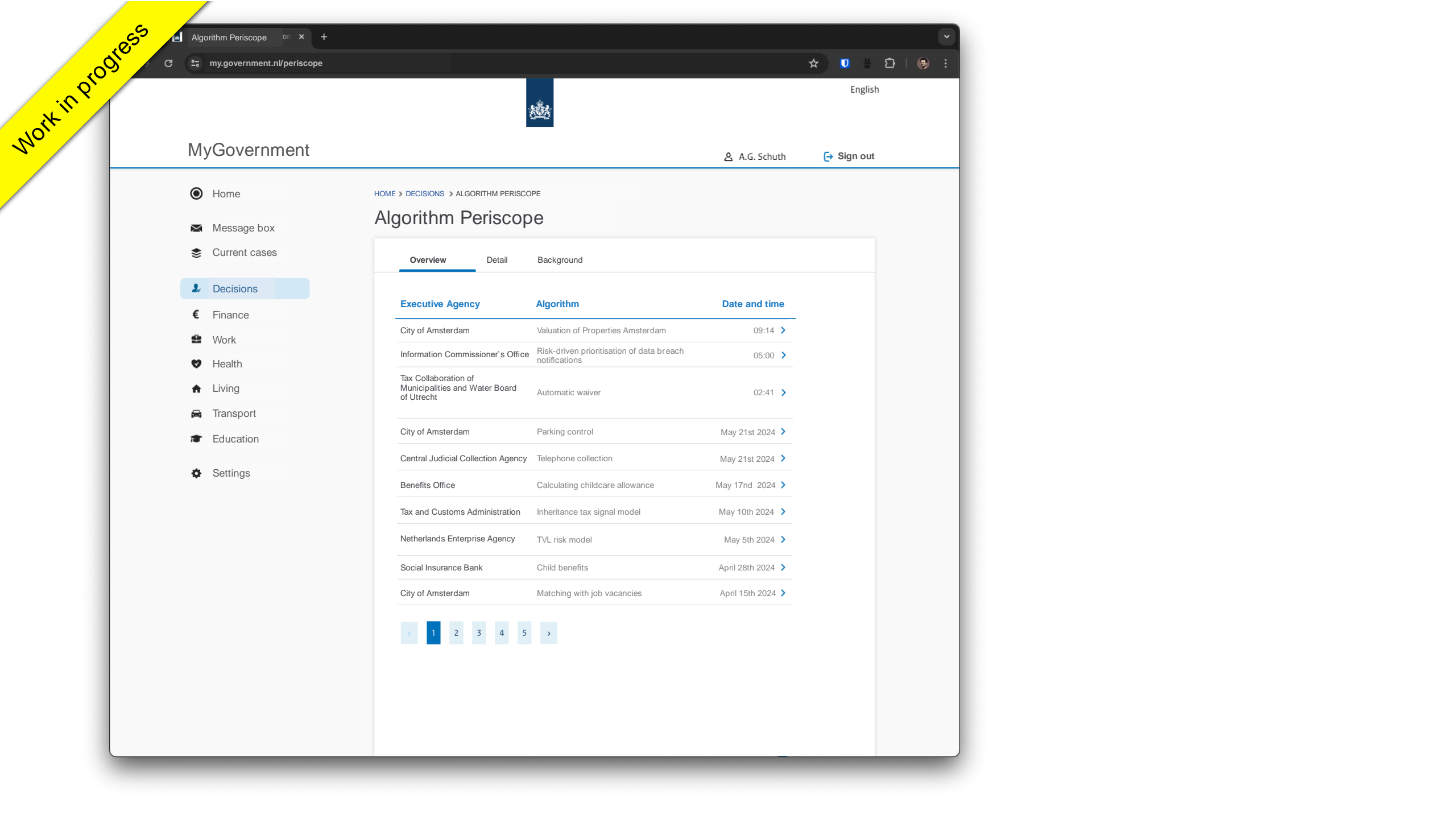
Raad voor de Rechtspraak Nog geen 5 procent van de vonnissen wordt op dit moment online gepubliceerd. Dat moet naar zo'n 75 procent, zegt de voorzitter van de Raad voor de Rechtspraak.

Steven Derix

30 mei 2021 • Leestijd 1 minuut



In 2021, **less than 5 percent** of court rulings were published online. The goal is to increase this to 75 percent.



Work in progress

Algorithm Periscope

my.government.nl/periscope

English

MyGovernment

A.G. Schuth Sign out

Home

Message box

Current cases

Decisions

Finance

Work

Health

Living

Transport

Education

Settings

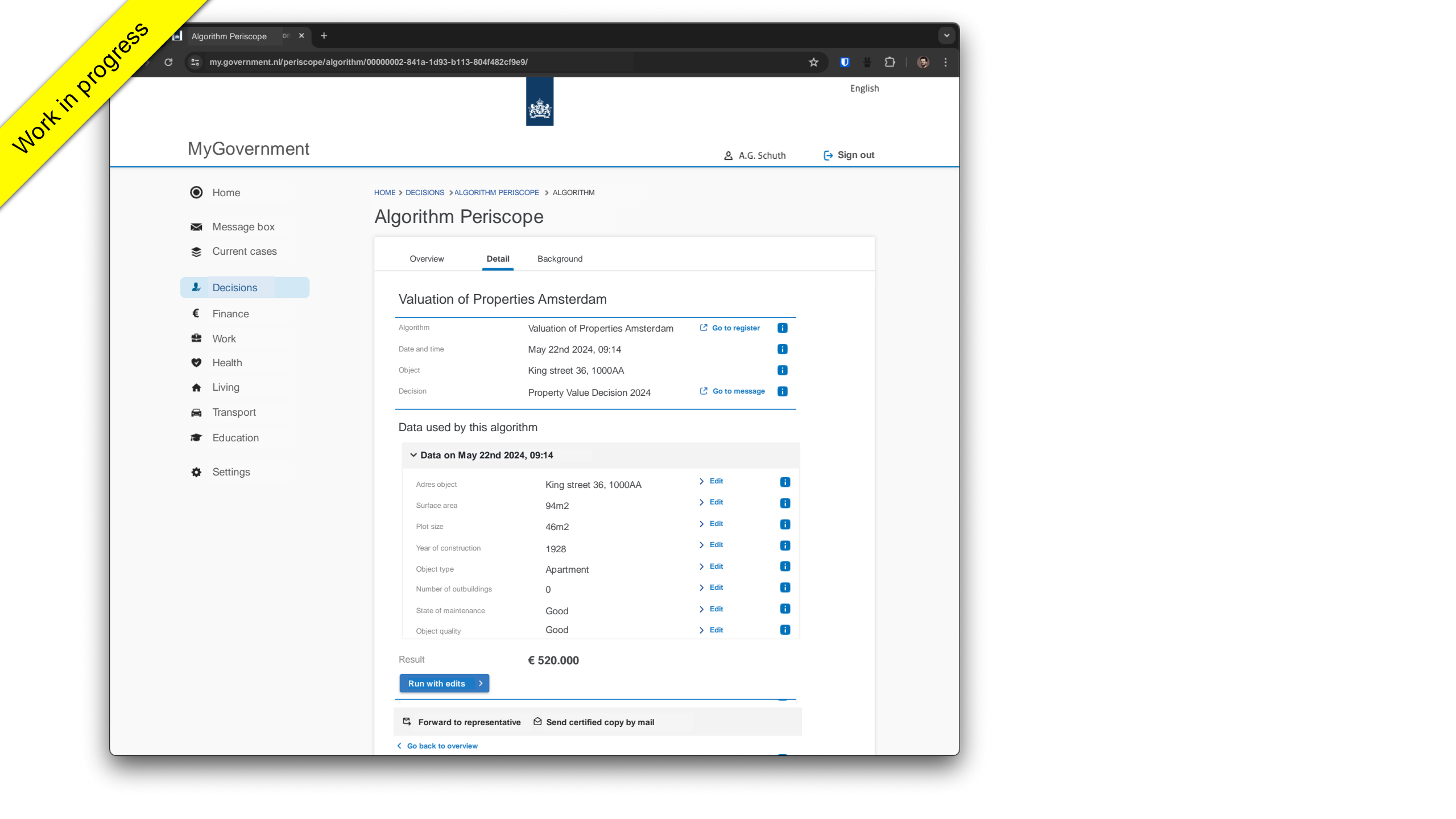
HOME > DECISIONS > ALGORITHM PERISCOPE

Algorithm Periscope

Overview Detail Background

| Executive Agency | Algorithm | Date and time |
|--|---|-------------------|
| City of Amsterdam | Valuation of Properties Amsterdam | 09:14 > |
| Information Commissioner's Office | Risk-driven prioritisation of data breach notifications | 05:00 > |
| Tax Collaboration of Municipalities and Water Board of Utrecht | Automatic waiver | 02:41 > |
| City of Amsterdam | Parking control | May 21st 2024 > |
| Central Judicial Collection Agency | Telephone collection | May 21st 2024 > |
| Benefits Office | Calculating childcare allowance | May 17nd 2024 > |
| Tax and Customs Administration | Inheritance tax signal model | May 10th 2024 > |
| Netherlands Enterprise Agency | TVL risk model | May 5th 2024 > |
| Social Insurance Bank | Child benefits | April 28th 2024 > |
| City of Amsterdam | Matching with job vacancies | April 15th 2024 > |

< 1 2 3 4 5 >



- Home
- Message box
- Current cases
- Decisions
- Finance
- Work
- Health
- Living
- Transport
- Education
- Settings

HOME > DECISIONS > ALGORITHM PERISCOPE > ALGORITHM

Algorithm Periscope

Overview Detail Background

Valuation of Properties Amsterdam

| | | | |
|---------------|-----------------------------------|--------------------------------|-------------------|
| Algorithm | Valuation of Properties Amsterdam | Go to register | i |
| Date and time | May 22nd 2024, 09:14 | | i |
| Object | King street 36, 1000AA | | i |
| Decision | Property Value Decision 2024 | Go to message | i |

Data used by this algorithm

▼ Data on May 22nd 2024, 09:14

| | | | |
|------------------------|------------------------|----------------------|-------------------|
| Adres object | King street 36, 1000AA | Edit | i |
| Surface area | 94m2 | Edit | i |
| Plot size | 46m2 | Edit | i |
| Year of construction | 1928 | Edit | i |
| Object type | Apartment | Edit | i |
| Number of outbuildings | 0 | Edit | i |
| State of maintenance | Good | Edit | i |
| Object quality | Good | Edit | i |

Result € 520.000

Run with edits

[Forward to representative](#) [Send certified copy by mail](#)

[Go back to overview](#)

Learn more

- The [Algorithm Register](#) of the Dutch Government
- [Fundamental Rights and Algorithm Impact Assessment](#) (FRAIA)
- [HUDERIA](#)
- The [5-star model](#) for transparent algorithms
- [Public standard profiling algorithms](#)
- [Algoprudence database](#)
- European harmonized AI standards for the AI Act
(not in today's presentation, but [happy to tell you all about them!](#))



Thank you for your attention.

Happy to connect! [linkedin.com/in/willytadema](https://www.linkedin.com/in/willytadema)

Slides are available for download at github.com/friesewoudloper/slidedecks

Images courtesy of Depositphotos and Shutterstock.

↓ [download the slides](#) ↓

