

Human-centred Explanation through Graph Databases

LangDev 2023

12/10/2023

Suzan Zuurmond
MDSE at the DTA (ALEF)

Human-centred explanation of rule-based decision-making systems in the legal domain

Suzan Zuurmond (4834186)

Faculty of Science, Utrecht University

Master Artificial Intelligence

Supervisors

A. Borg (Project supervisor, First examiner)

M. van Kempen (Daily supervisor)

S. Renooij (Second examiner)

R. Wieten (Second daily supervisor)

Master thesis

Credits: 30 EC

June 5th, 2023

What did I do?

Regelgroep Betaling

Regel Betaling tijdig betaald 01
geldig sinds

Een **Betaling** is **tijdig betaald** indien de **betaldatum** van de **Betaling** eerder of gelijk is aan de **einddatum betalstermijn** van de **aangifte** bij **Betaling**.

Regel Betaling betaald na 1e kwartaal kalenderjaar na ter beschikkingstelling dividend indien de **betaldatum** van de **Betaling** later of gelijk is aan de 1 april van het jaar volgend op het jaar waar beschikking is geheeld van de **aangifte** bij **betaling** van de **Betaling**.

Bronnen:
1. https://wetten.overheid.nl/nl/c1.3:c:BWBR0002320&hoofdstuk=VA&artikel=30h&lid=4&z=2022-12-08&g=2021-01-01
H VA art 30h lid 4 wet BWBR0002320
geldig sinds 2021

Een **Betaling** is betaald na 1e kwartaal kalenderjaar volgend op jaar van ter beschikkingstelling dividend indien de **betaldatum** van de **Betaling** later of gelijk is aan de 1 april van het jaar volgend op het jaar waar beschikking is geheeld van de **aangifte** bij **betaling** van de **Betaling**.

Regelgroep Belastingrente

Regel Belastingrente verschuldigd wegens te late betaling 01
geldig sinds

1. https://wetten.overheid.nl/nl/c1.3:c:BWBR0002320&hoofdstuk=VA&artikel=30h&lid=4&z=2022-12-08&g=2021-01-01
H VA art 30h lid 4 wet BWBR0002320
geldig sinds 2021

Een **Belastingrente** is belastingrente verschuldigd wegens te late betaling indien de volgende voorwaarden worden voldaan :

- de vast te stellen naheffingsaanslag van de **aangifte** voor volgend op het kalenderjaar waarin de **aangifte** is ingevuld of voor de beschikking is geheeld van de **aangifte** bij **betaling** van de **Betaling**;
- ten minste één van de betalingen van de **aangifte** is belastingrente verschuldigd wegens te late betaling.

Regel Belastingrente verschuldigd wegens te late betaling 01
geldig sinds

1. https://wetten.overheid.nl/nl/c1.3:c:BWBR0002320&hoofdstuk=VA&artikel=30h&lid=4&z=2022-12-08&g=2021-01-01
H VA art 30h lid 4 wet BWBR0002320
geldig sinds 2021

Een **Belastingrente** is belastingrente verschuldigd wegens te late betaling indien de **Betaling** aan alle volgende voorwaarden voldoet :

- de **Betaling** is niet tijdig betaald;
- de **Betaling** is verschuldigd na de eerste half van het kalenderjaar waarin de **aangifte** is ingevuld of voor de beschikking is geheeld van de **aangifte** bij **betaling** van de **Betaling**;
- de **aangifte** is niet voorafgaand aan de **Betaling**;

In geval dat **betaling** in ophoud is in jaar van de dividend plus 1 is niet betaald gelijk aan geen betaling kunnen diende betalstermijn er niet rekening houden. Geen bron hiervoor T000

Regel Belastingrente verschuldigd 01
geldig sinds 2021

Regelgroep Totalen naheffingsaanslag

Regel Totale aan grondreger te late betalingen 01
geldig sinds

je is vastgesteld na het einde van het kalenderjaar

Het totaal van betalingen voor late betalingen van een **Vast te stellen naheffingsaanslag** moet berekend worden als de som van de verschillende **Belastingrente verschuldigd wegens te late betaling** die voldoen aan de volgende voorwaarden :

- de **Betaling** is belastingrente verschuldigd wegens te late betaling;

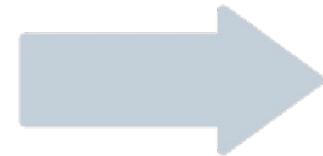
Regel Totale aan betalingen over heel grondreger in den best ligt en
geldig sinds

Het totaal van betalingen over heel grondreger is de **Betaling** van een **Vast te stellen naheffingsaanslag** moet berekend worden als de som van de verschillende **Belastingrente verschuldigd wegens te late betaling** die voldoen aan de volgende voorwaarden :

- de **Betaling** is belastingrente verschuldigd wegens te late betaling;

Indien er aan alle volgende voorwaarden wordt voldaan :

- de **Betaling** is niet tijdig betaald;
- de **Betaling** is verschuldigd na de eerste half van het kalenderjaar waarin de **aangifte** is ingevuld of voor de beschikking is geheeld van de **aangifte** bij **betaling** van de **Betaling**;
- de vast te stellen naheffingsaanslag van de **aangifte** bij **Betaling** is vastgesteld op of na 1 april in het kalenderjaar volgend op het kalenderjaar waarin de **aangifte** is ingevuld of voor de beschikking is geheeld van de **aangifte** bij **betaling** van de **Betaling**;



Waarom is de Belastingplichtige belastingrente verschuldigd wegens te late betaling?

Regel Belastingrente verschuldigd wegens te late betaling 01

Bronnen:

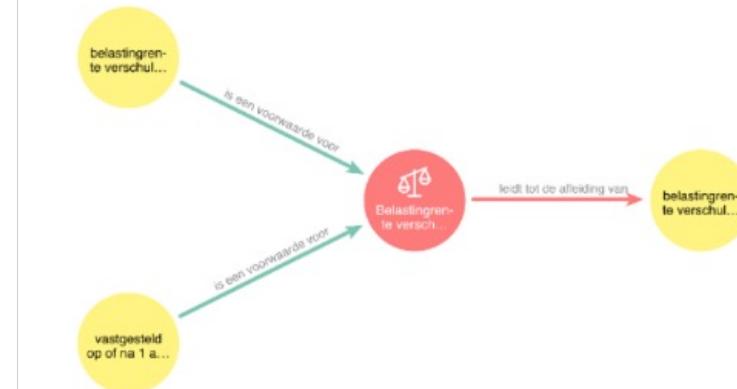
1. [http://wetten.overheid.nl/nl/c1.3:c:BWBR0002320&hoofdstuk=VA&artikel=30h&lid=4&z=2022-12-08&g=2021-01-01](https://wetten.overheid.nl/nl/c1.3:c:BWBR0002320&hoofdstuk=VA&artikel=30h&lid=4&z=2022-12-08&g=2021-01-01)
H VA art 30h lid 4 wet BWBR0002320

geldig vanaf 2021

Een **Belastingplichtige** is **belastingrente verschuldigd wegens te late betaling**

Indien er aan alle volgende voorwaarden wordt voldaan :

- de vast te stellen naheffingsaanslag van de **aangifte** van de **Belastingplichtige** is vastgesteld op of na 1 april in het kalenderjaar volgend op het kalenderjaar waarin dividend ter beschikking is gesteld;
- ten minste één van de betalingen van de **aangifte** is **belastingrente verschuldigd wegens te late betaling**.



from natural language representation of rulesets towards a visual and questionable representation in a Neo4j GraphDB

How to create explainable decision-making systems?

Two complementary approaches

Using simpler decision processes for inherent interpretability

Creating explanations to clarify the behaviour and decisions of systems

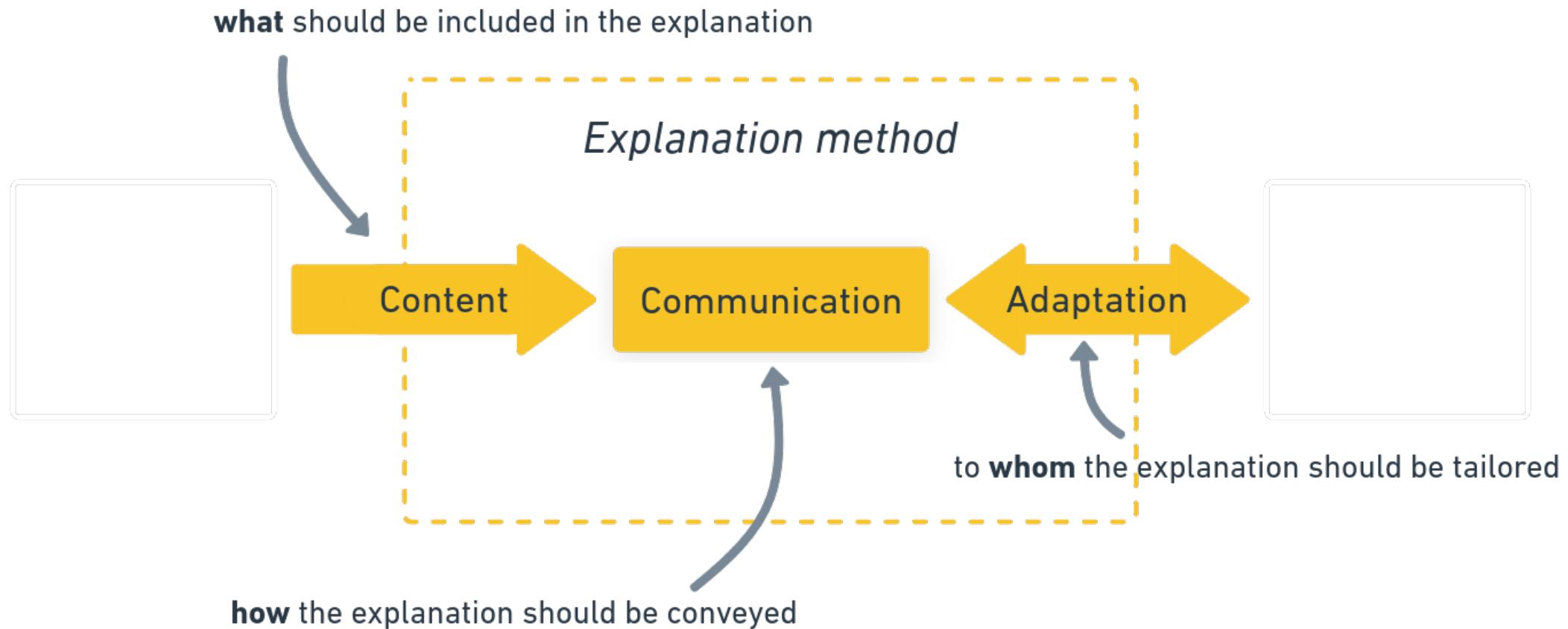
Interpretability

Explanation

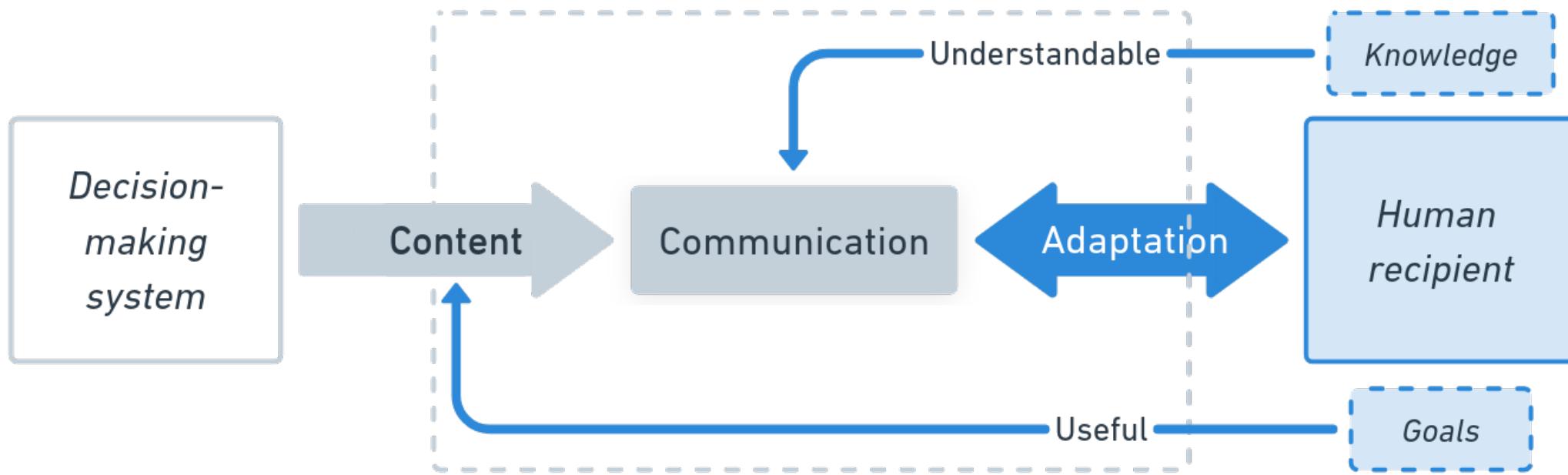
Language Engineering

Objective of Thesis

What are the components of explanation?



How to provide human-centred explanation?



How to provide human-centred explanation?

Insights from social sciences and humanities

- Explanation is "an answer to a (why-)question" (Miller, 2019)

How to provide human-centred explanation?

Insights from social sciences and humanities

- Explanation is "an answer to a (why-)question" (Miller, 2019)

So, we should be able to **answer all the recipient's questions.**

Actually, not *all* questions!
only data and rationale related



How to provide human-centred explanation?

Insights from social sciences and humanities

- An explanation is "an answer to a (why-)question" (Miller, 2019)

Methods from (explainable) artificial intelligence

- Question bank and taxonomy (Liao et al., 2020)
- Explanation generation algorithms (Lim & Dey, 2009)
 - for each explanation type
 - aided by visualisations of a conceptual model

How to provide human-centred explanation?

Insights from social sciences and humanities

- An explanation is "an answer to a (why-)question" (Miller, 2019)

Methods from (explainable) artificial intelligence

- Question bank and taxonomy (Liao et al., 2020)
- Explanation generation algorithms (Lim & Dey, 2009)
 - for each explanation type
 - aided by visualisations of a conceptual model

How to provide human-centred explanation?

Insights from social sciences and humanities

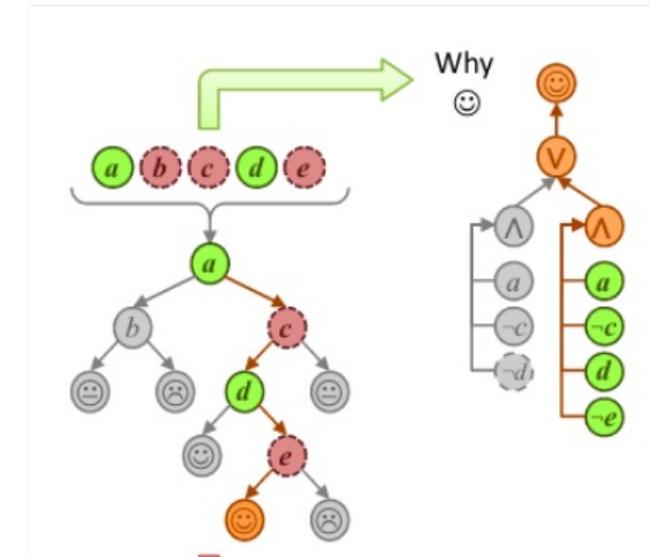
- An explanation is "an answer to a (why-)question" (Miller, 2019)

Methods from (explainable) artificial intelligence

- Question bank and taxonomy (Liao et al., 2020)
- Explanation generation algorithms (Lim & Dey, 2009)
 - for each explanation type
 - aided by visualisations of a conceptual model



How? By using a graph database!



How to implement the explanation method?

1

Knowledge Graph

decision models in a structured form

Abstract Syntax Graph

2

Global Graph

conceptual decision model containing key semantic information

Dependency Graph

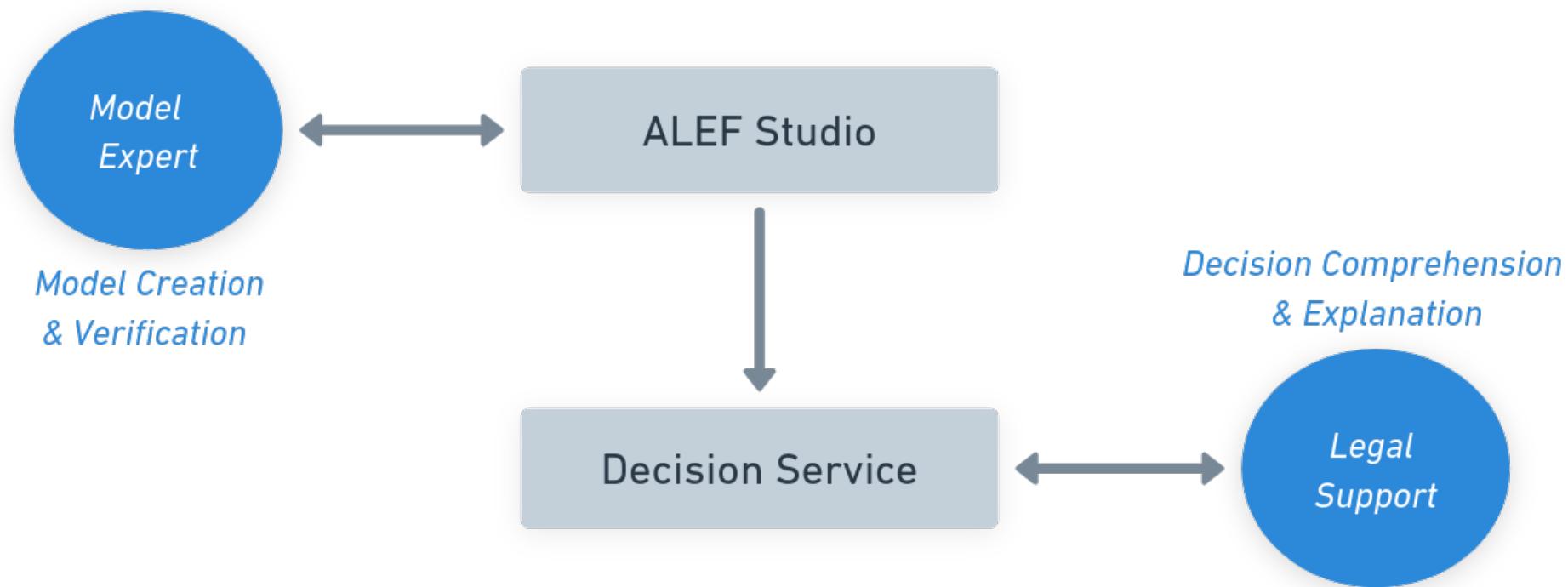
3

Local Graph

instantiated model containing the values of one specific decision

"Filled" Global Graph

How to decide what is key information?



How to decide what is key information?

Recipient's knowledge

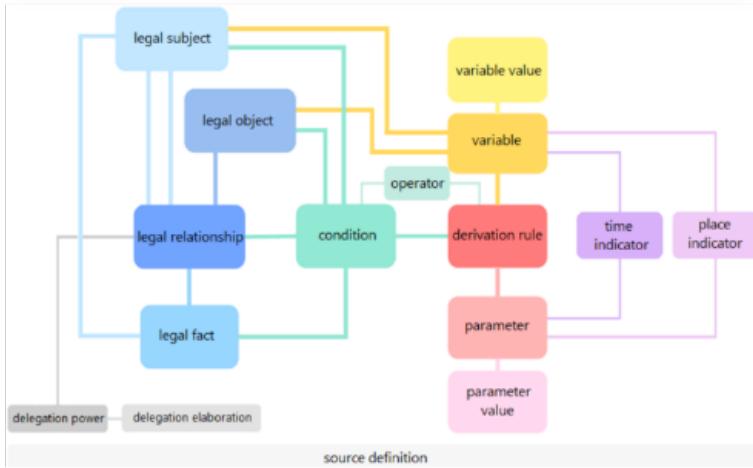


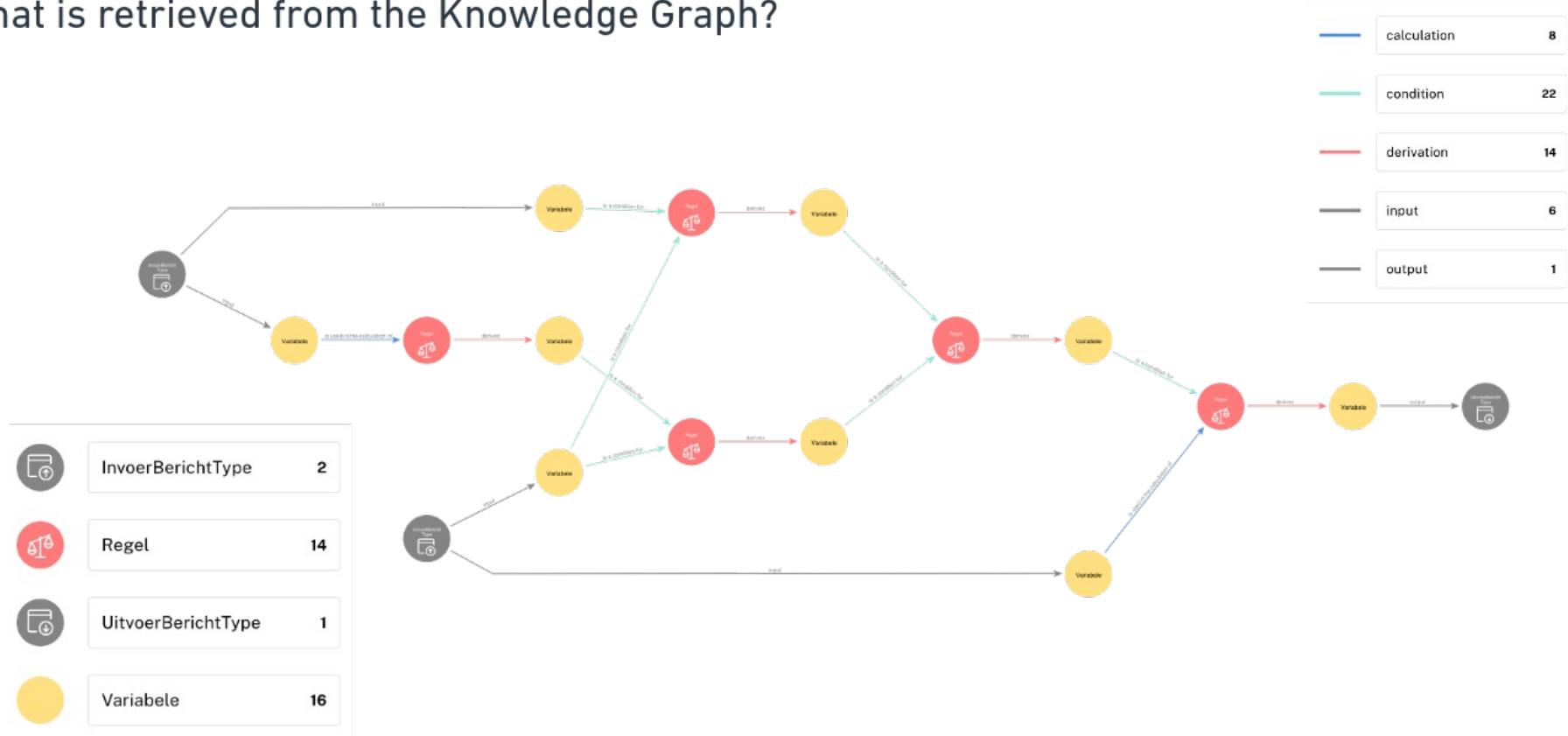
Diagram representing a framework for legal analysis at the DTA, translated from (Lokin et al., 2021)

Recipient's goals

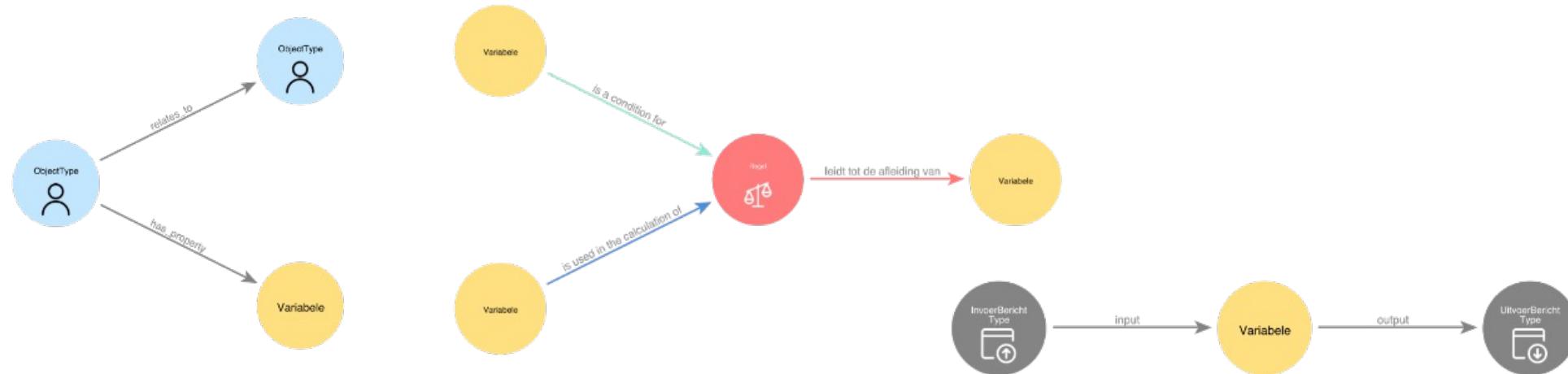
	Type	Description
Decision	What	Describes what decision(s) it made and what input it used.
	What if	Describes the system's decision when an input value changes.
	Why	Clarifies the reasoning behind the system's decision, given the input values.
	Why not	Clarifies why an alternative decision was not generated, given the input values.
Model	How to	Clarifies how to achieve a desired decision, given some (but not all necessary) input values.
	Input	Describes which input the system could use.
	Output	Describes which decisions the system could make.
	How	Clarifies how the system generally makes a certain decision.
Visualisation	Visualisation	Describes a simplified view of the system's conceptual model.
	Whether	Clarifies whether the system meets a certain requirement.

General user-type question of automated decision systems, based on (Lim and D

What is retrieved from the Knowledge Graph?



What is retrieved from each model?



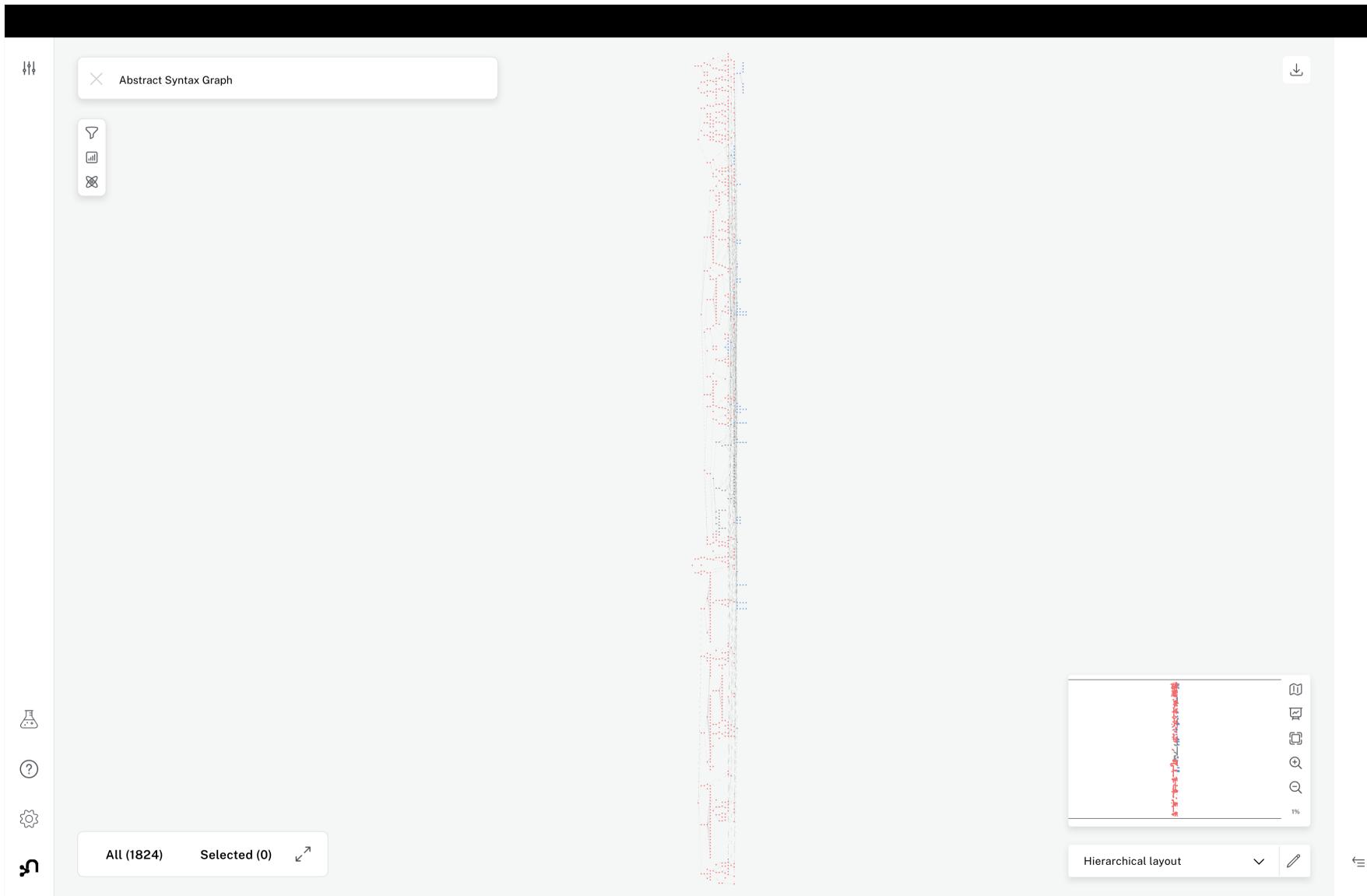
data in gegevens.mps

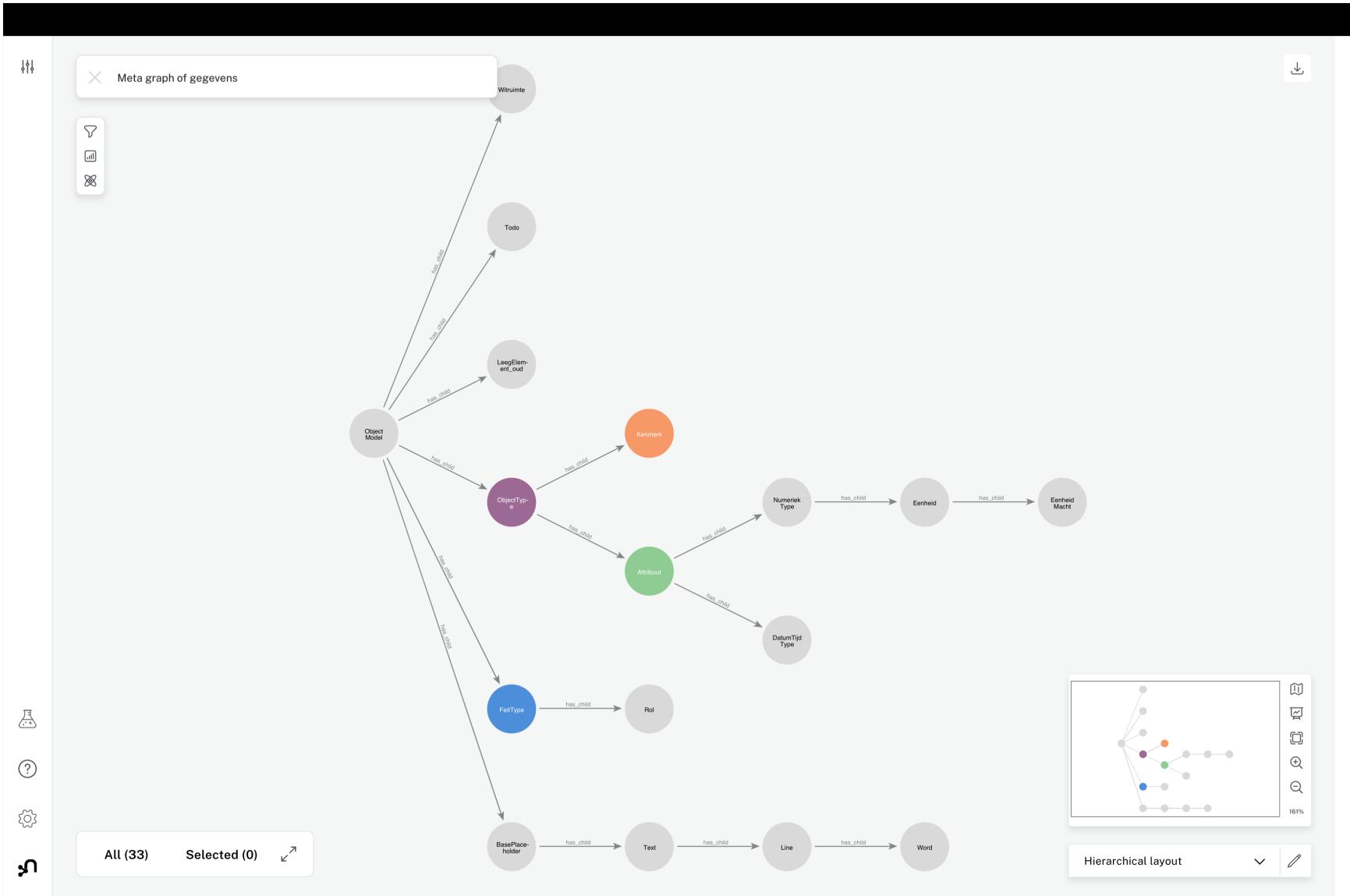
rules in regels.mps

services in services.mps

Demo Knowledge Base





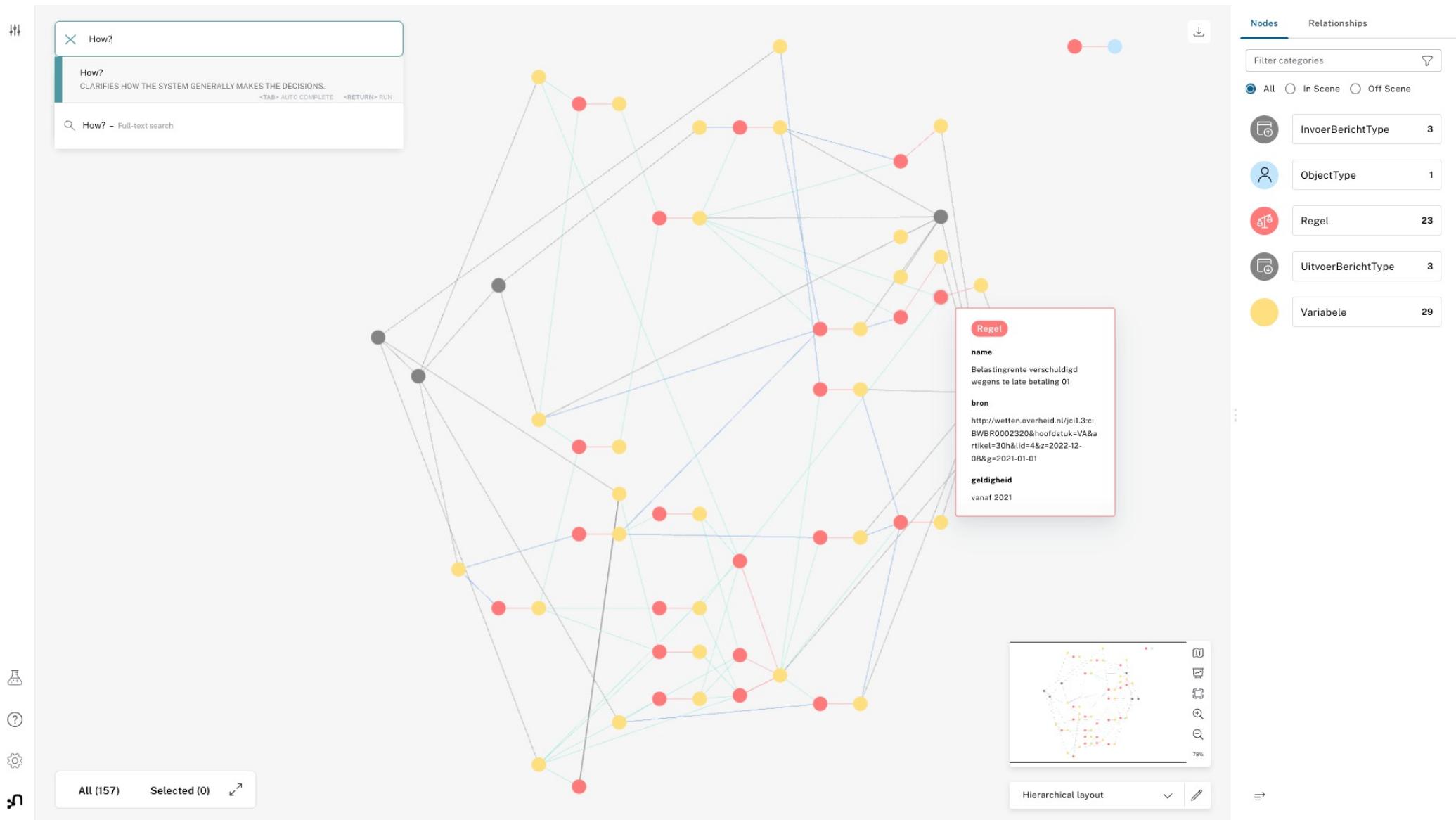


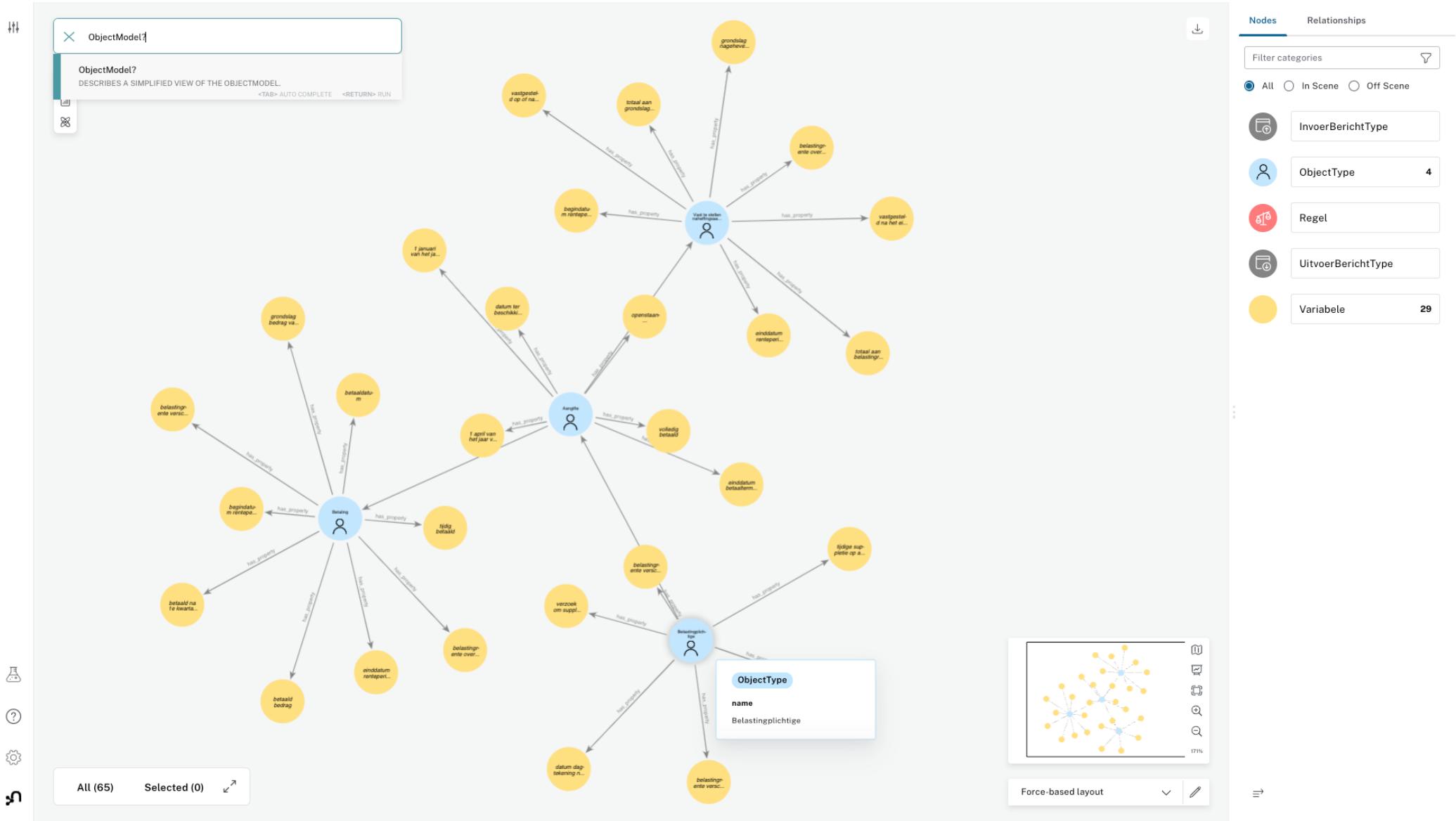
Demo Global Explanation

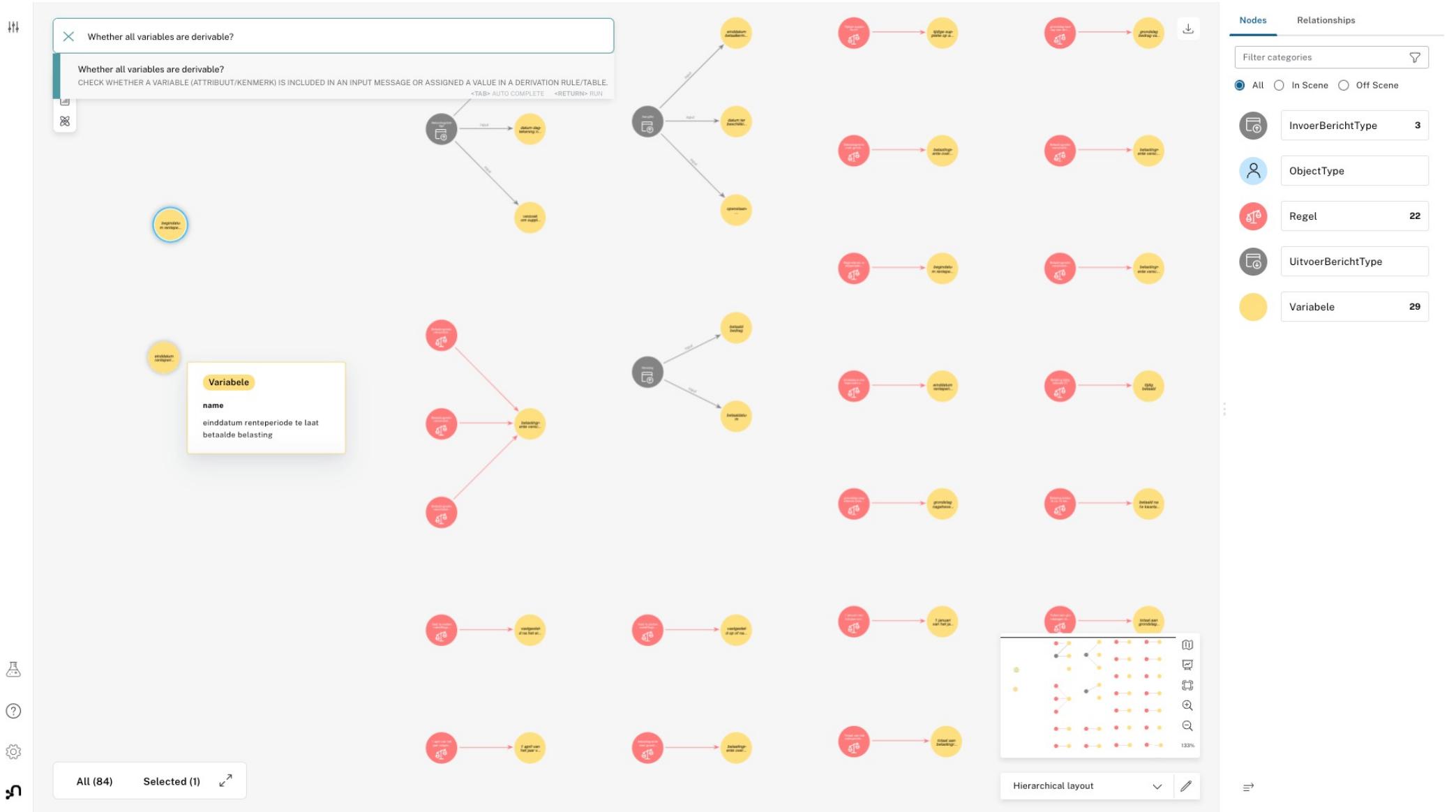
Model expert – Model creation and verification

Model Expert - Creation and verification

	Type	Description
Model	Input	Describes which input the system could use.
	Output	Describes which decisions the system could make.
	How	Clarifies how the system generally makes a certain decision.
	Visualisation	Describes a simplified view of the system's conceptual model.
	Whether	Clarifies whether the system meets a certain requirement.





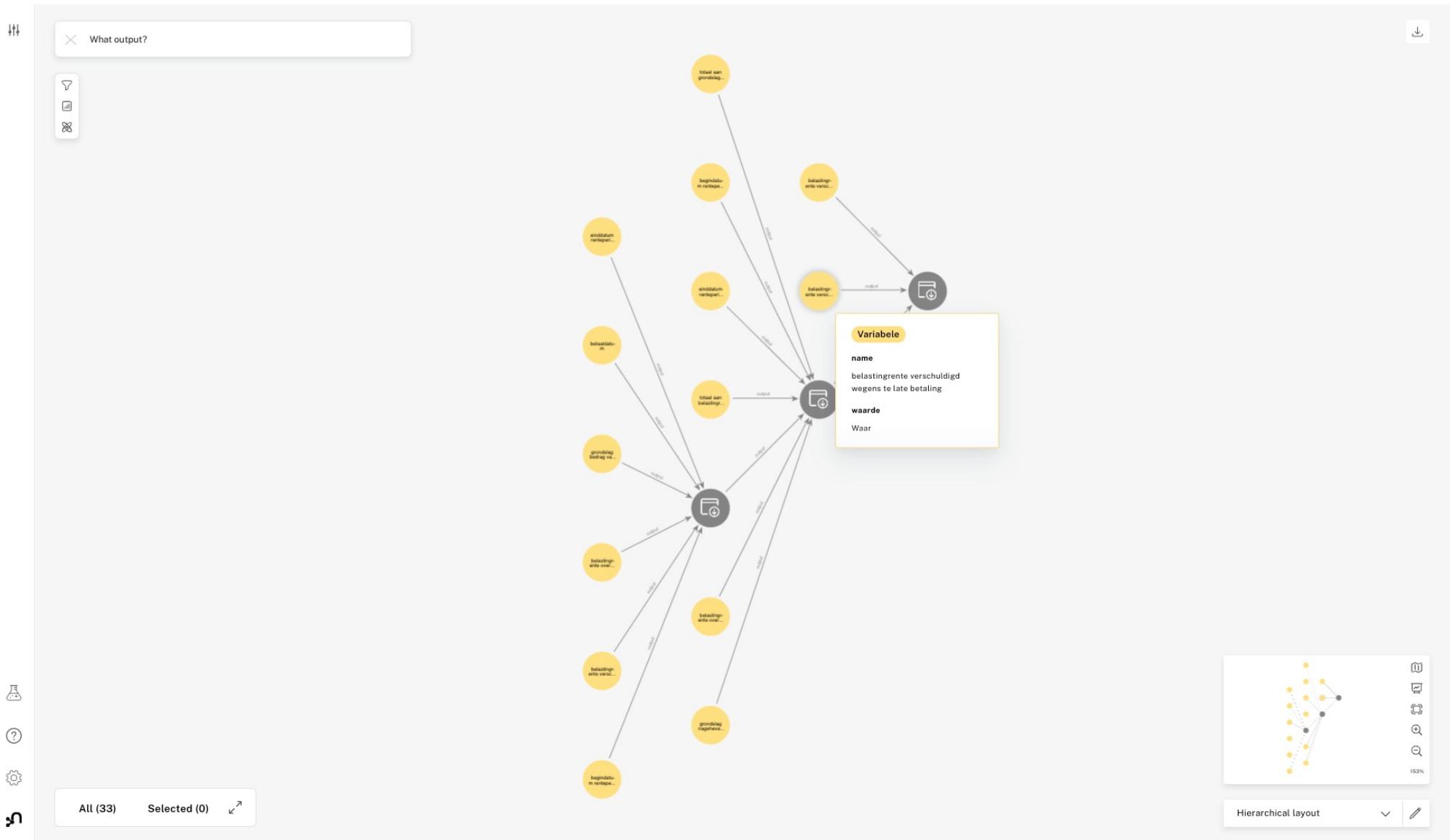


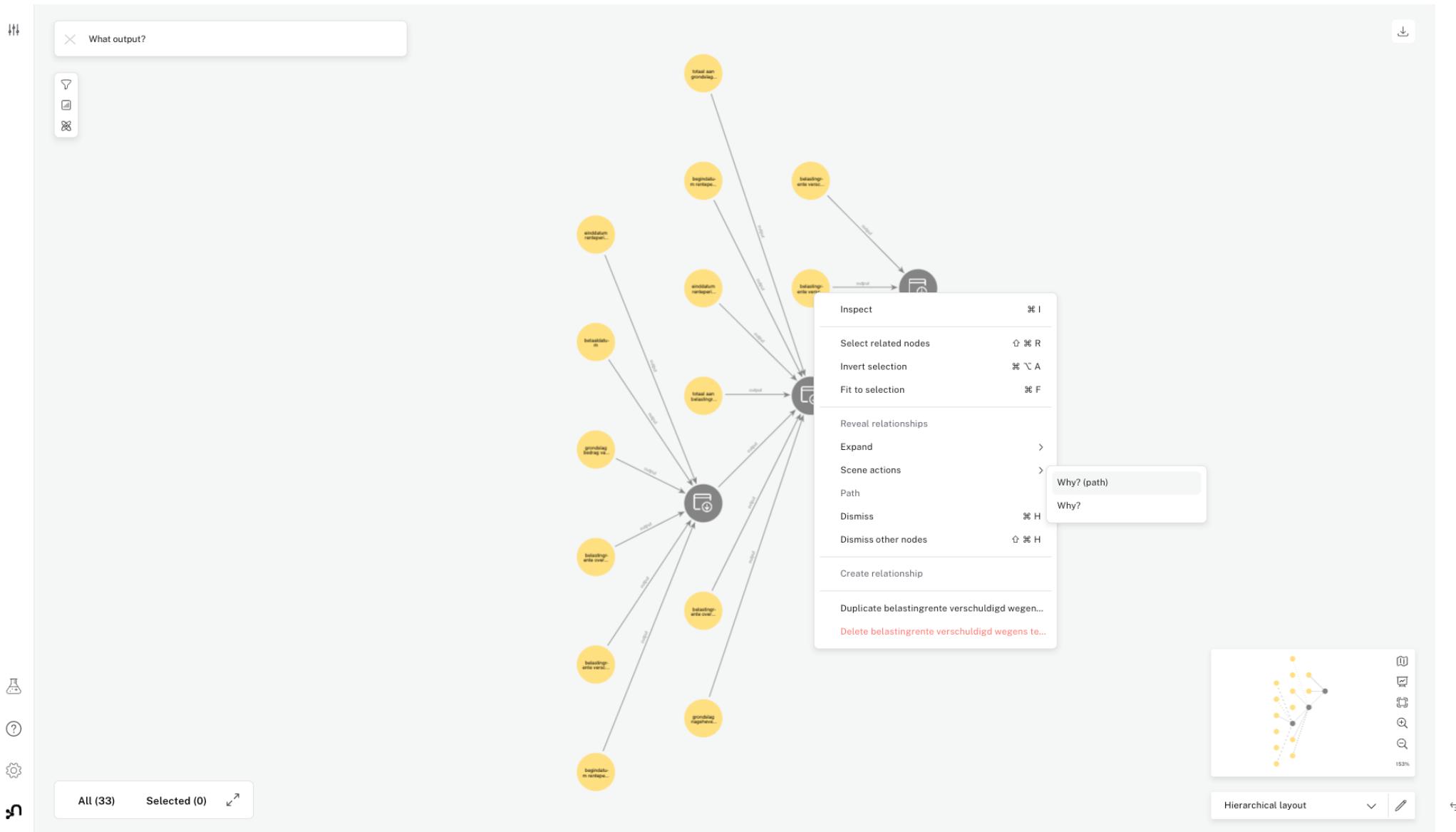
Demo Local Explanation

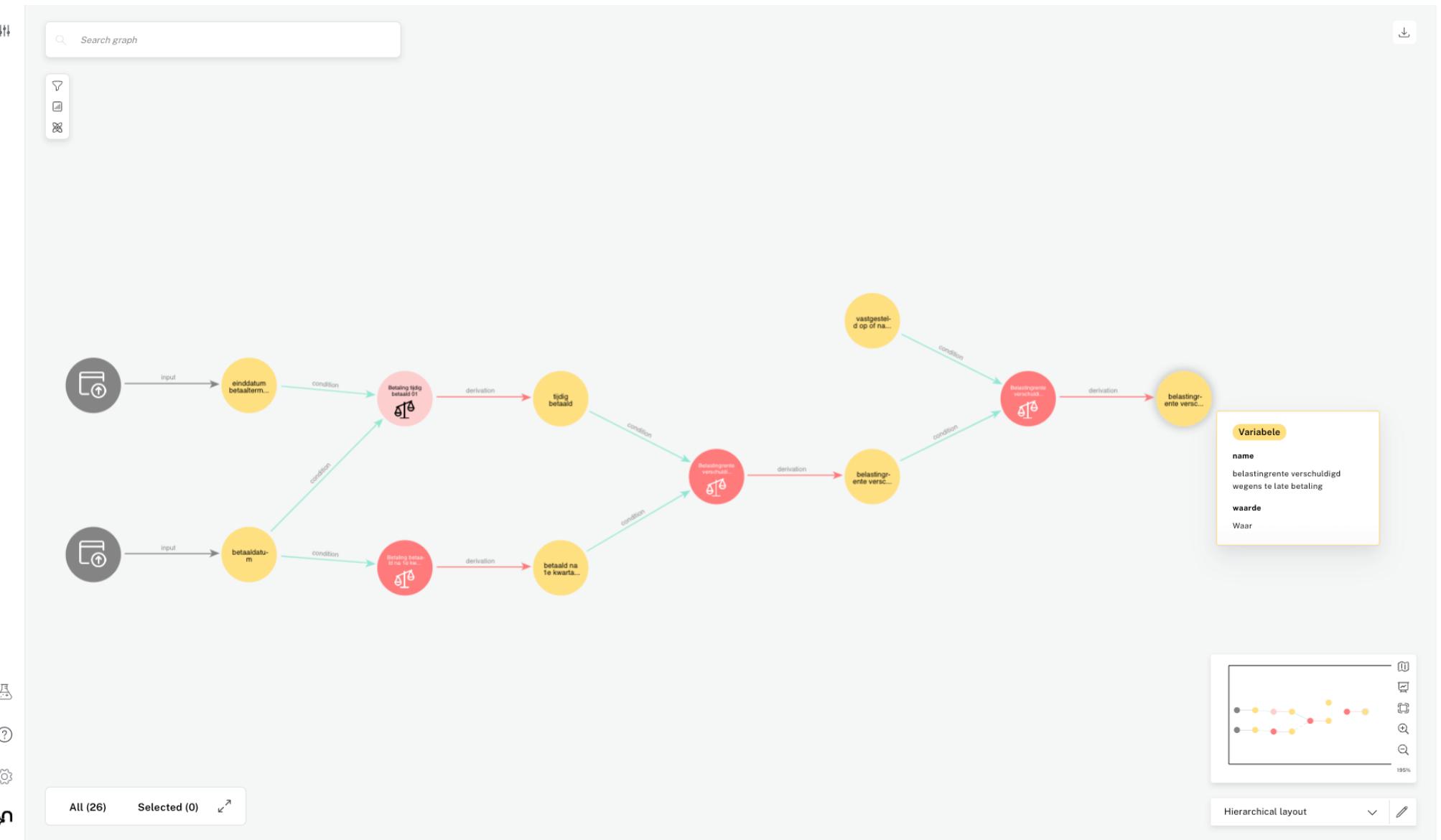
Legal support – Decision comprehension and explanation

Legal support - Explanation

	Type	Description
<i>Decision</i>	What	Describes what decision(s) it made and what input it used.
	What if	Describes the system's decision when an input value changes.
	Why	Clarifies the reasoning behind the system's decision, given the input values.
	Why not	Clarifies why an alternative decision was not generated, given the input values.
	How to	Clarifies how to achieve a desired decision.







Future research

- Reasoning properties
- Dynamic modeling
- Participant study
- Insights from application domain

Liao, Q. V., Gruen, D., & Miller, S. (2020). Questioning the AI: Informing design practices for explainable AI user experiences. Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, 1–15. <https://doi.org/10.1145/3313831.3376590>

Lim BY, Dey AK. Assessing demand for intelligibility in context-aware applications. In: Proceedings of the 11th International Conference on Ubiquitous computing. Association for Computing Machinery; 2009. p. 195-204.

Lokin M. Wendbaar wetgeven: de wetgever als systeembeheerder. Vrije Universiteit Amsterdam; 2018.

Miller, T. (2019). Explanation in artificial intelligence: Insights from the social sciences. Artificial Intelligence, 267, 1–38. <https://doi.org/10.1016/j.artint.2018.07.007>

Short paper of presented ideas

Zuurmond, S., Borg, A., van Kempen, M., & Wieten, R. (2023). Human-centred explanation of rule-based decision-making systems in the legal domain. arXiv preprint arXiv:2310.16704.