17/04/2023

*Decision Capability: uitbouwen van uitlegbare beslissingsmodellen

Jan Vanthienen
jan.vanthienen@kuleuven.be
Leuven Institute for Research in Information Systems

m.m.v. **Alexandre Goossens, Vedavyas Etikala**

KU LEUVEN

Me

Em. Prof. Dr Jan Vanthienen
KU Leuven
Faculty of Economics and Business
Business Information Systems Group

Research and teaching:

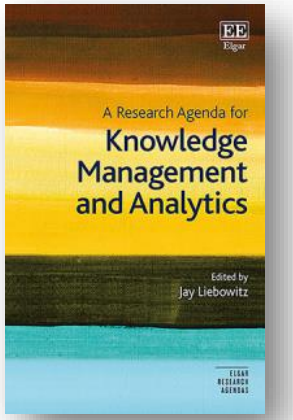
- Business rules, processes and information systems
- Decision models & tables
- AI, Business intelligence & Analytics
- Information & Knowledge Management

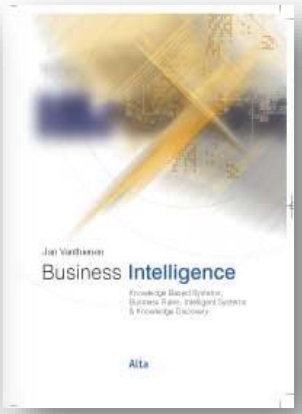
OMG WE SET THE STANDARD™ DMN Decision Model & Notation™

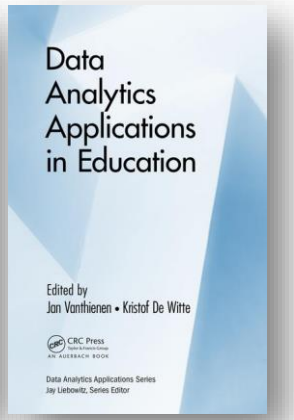
IBM Faculty Award
Belgian Francqui Chair 2009 at FUNDP

- ING Research Chair on Metadata Analytics
- Bpost bank Research Chair Actionable Analytics
- Colruyt-Symeta Research Chair Smart Data and Decisions in Marketing
- IBM Fund Intelligent Business Decision Making
- Microsoft Research Chair on Intelligent Environments
- PricewaterhouseCoopers Chair on E-Business

Email: jan.vanthienen@kuleuven.be
http://janvanthienen.net/







Contents

1. **Decisions and decision modeling**
   - DMN
   - Decision capability

2. **Executing decision models**
   - Simple or optimized

3. **Explainable decision models**
   - Flexible reasoning
   - Chatbot

4. **Recent developments**
   - Decision modeling & mining

Decisions and DMN

Decisions as a business concern

Because

- More and more **business operations are being automated**,
- and business has to react immediately, e.g. in online applications,
- and a human is often not involved anymore at the moment of transaction

It is important that

- Systems, processes and **decisions** are crafted correctly and remain correct after updates

A decision is defined as ‘the act of determining an output value from a number of input values (data), using decision logic (models), defining how the output is determined from the inputs.

Decision automation

- **Strategic and tactical decisions**
  - ➢ Often high-level, long-term, big impact, infrequent, unstructured

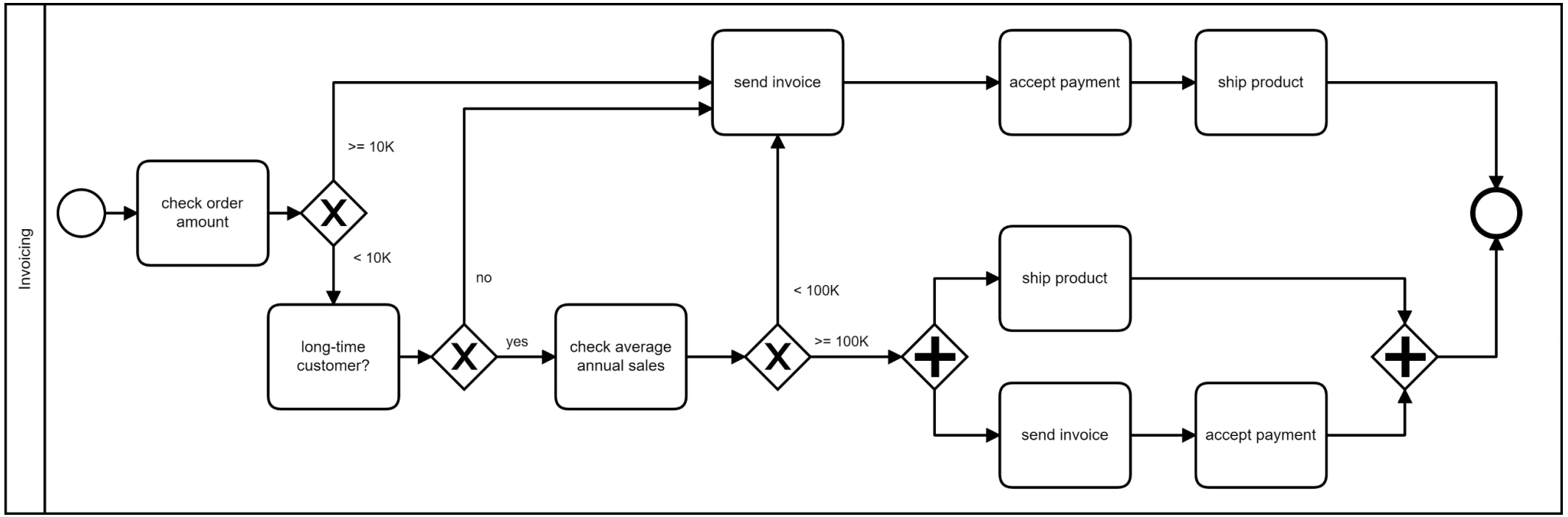
- **Operational decisions**
  - ➢ Daily, high-volume, often structured, standard procedures
    - Mathematical (and analytics) models

    or

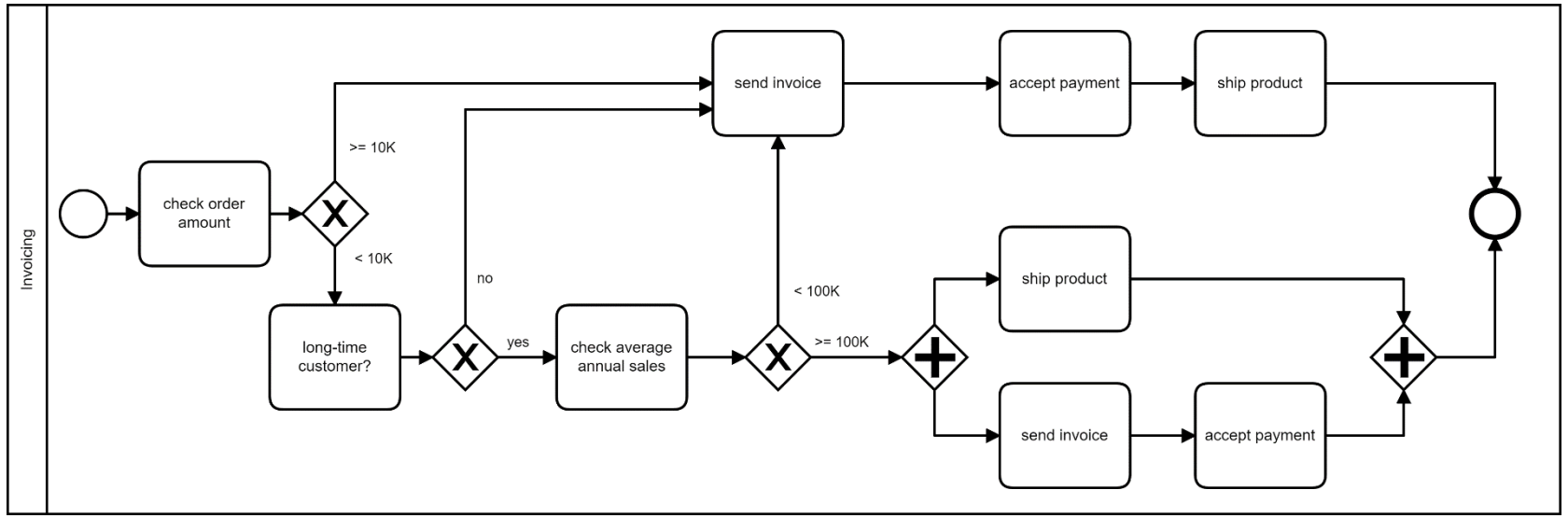
    - **Standard operating procedures based on company rules, policies and regulations present in operational business processes**

      - E.g.: eligibility, price, insurance, discount, theft rating, customer offer, retention, supplier selection, hire, credit, …

Decision and process: an example



What is really the issue here?



Sometimes we expect payment before we ship (to be on the safe side)
and sometimes we ship and send the invoice in parallel (only for good customers with a small order amount)

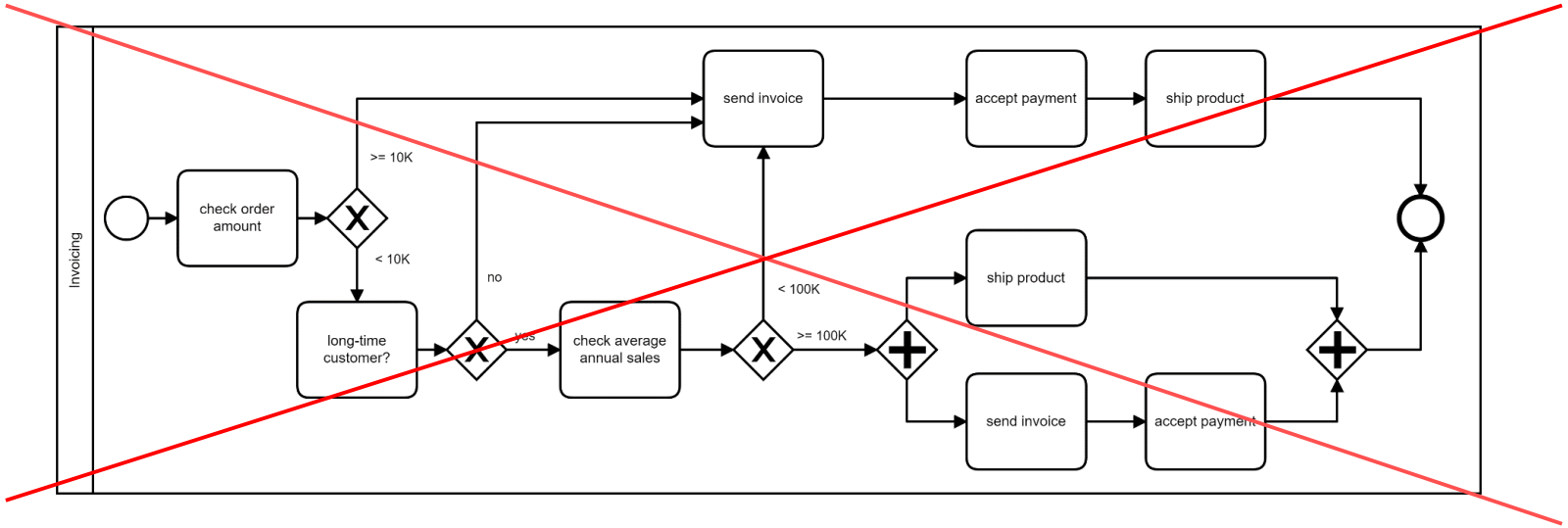
The real decision therefore is:

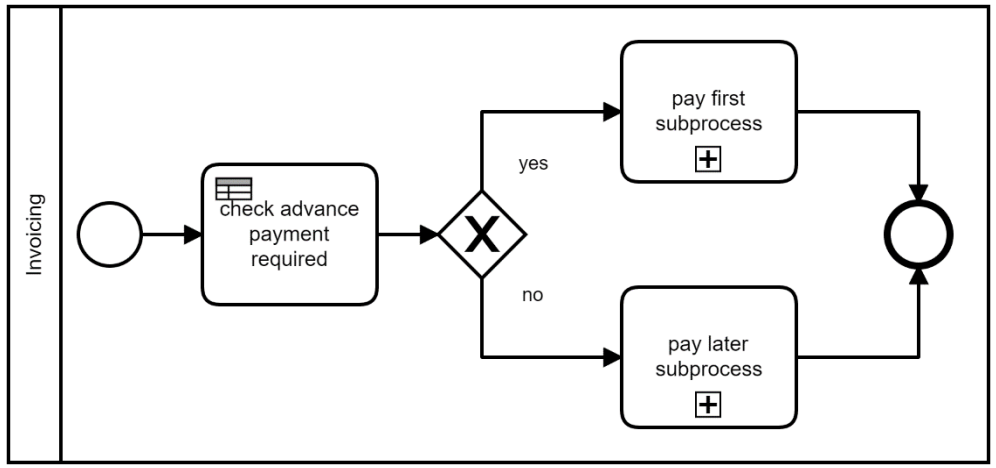
Do we require advance payment? (and what is a good customer?)

That decision seems to depend on order amount, customer history and average annual customer sales.

Where is this decision in the model?

The simpler process with a decision





Prepayment is not required for loyal/good customers when the OrderAmount is small (<10K)

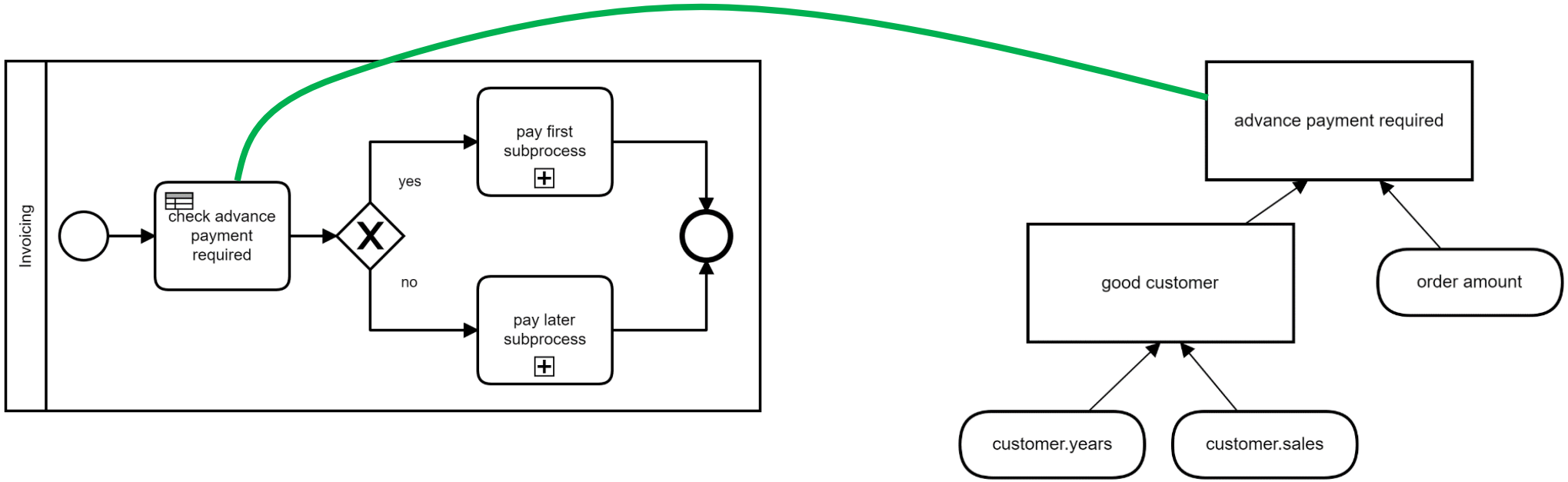
A loyal/good customer is defined as such if their AnnualSales is high (>=100K) and their Customeryears is more than 5

DMN

- **Decision Model & Notation (DMN)**
- **DMN 1.0 published by Object Management Group (OMG) in 2015**
- **Now DMN 1.4 (2022), DMN 1.5 approved (2023)**
- DMN is an executable notation for the precise specification of business decisions and rules by the business. It is designed to work alongside BPMN (or CMMN) providing a mechanism to model decision making.

- **Typical applications: decisions in finance, insurance, healthcare, legal, …**

The process and the decision requirements



Why separate decisions and processes?

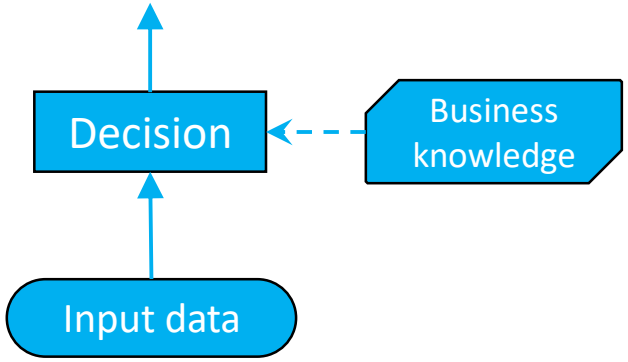
- Processes can be stable and simpler
- Decisions have improved visibility and can be reused
- Decision logic can be maintained separately
- Different stakeholders, separation of concerns

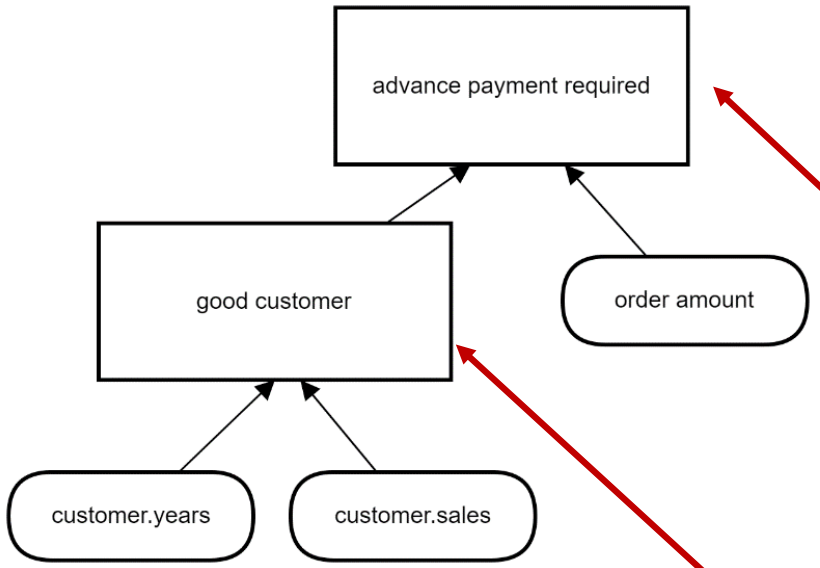
Decision

Input Data

Informaton requirement

The decision logic





advance payment required

AdvP

View DRD

U	Input +		Output +	
	good customer	order.amount	advance payment required	Annotation
	boolean	integer	boolean	
1	true	< 10K	false	
2	true	>= 10K	true	
3	false	-	true	

good customer

GoodC

View DRD

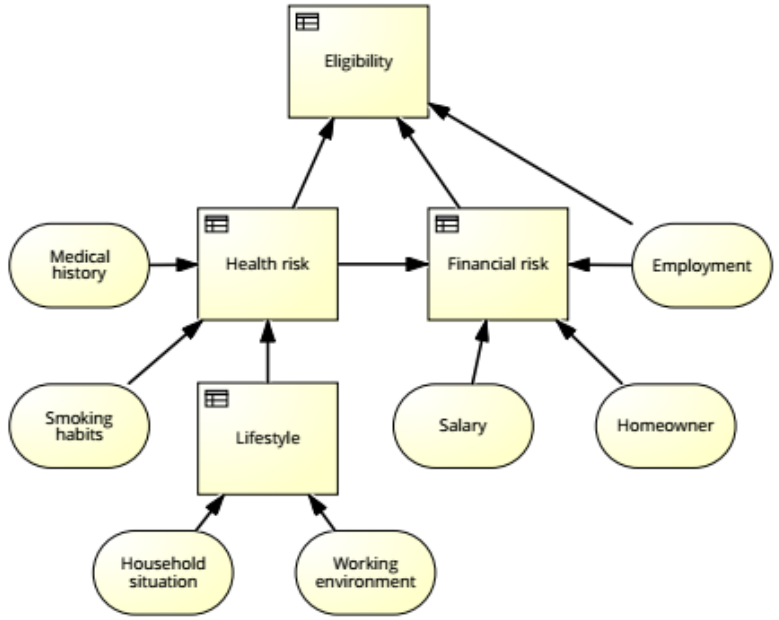
U	Input +		Output +	
	customer.years	customer.sales	good customer	Annotation
	integer	integer	boolean	
1	< 5	-	false	
2	>= 5	< 100K	false	
3	>= 5	>= 100K	true	

Prepayment is not required for loyal/good customers when the OrderAmount is small (<10K)

A loyal/good customer is defined as such if their AnnualSales is high (>=100K) and their Customeryears is more than 5

How to model decisions?

1. **Decisions requirements**

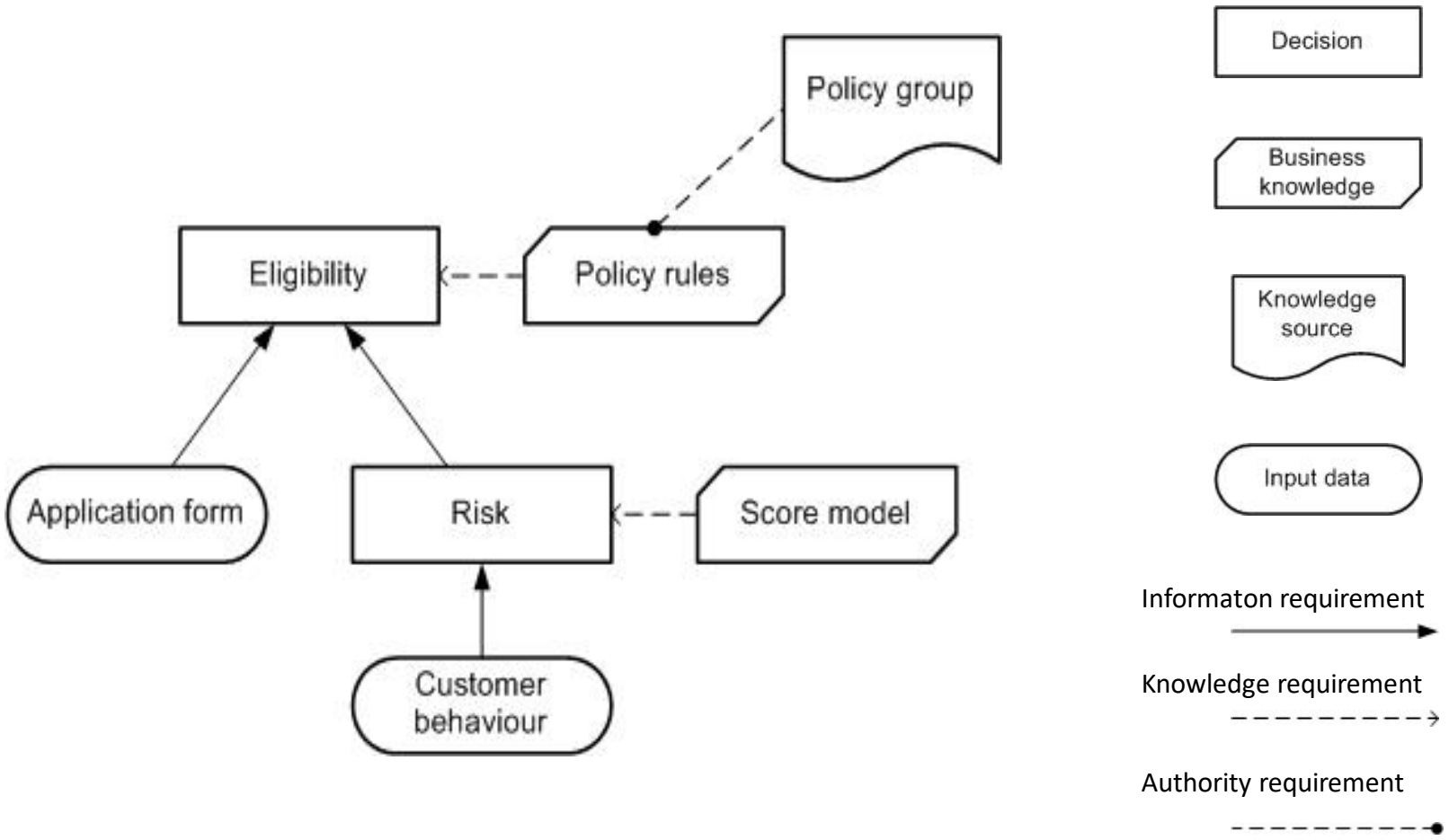


Decision requirement diagram for the credit eligibility decision

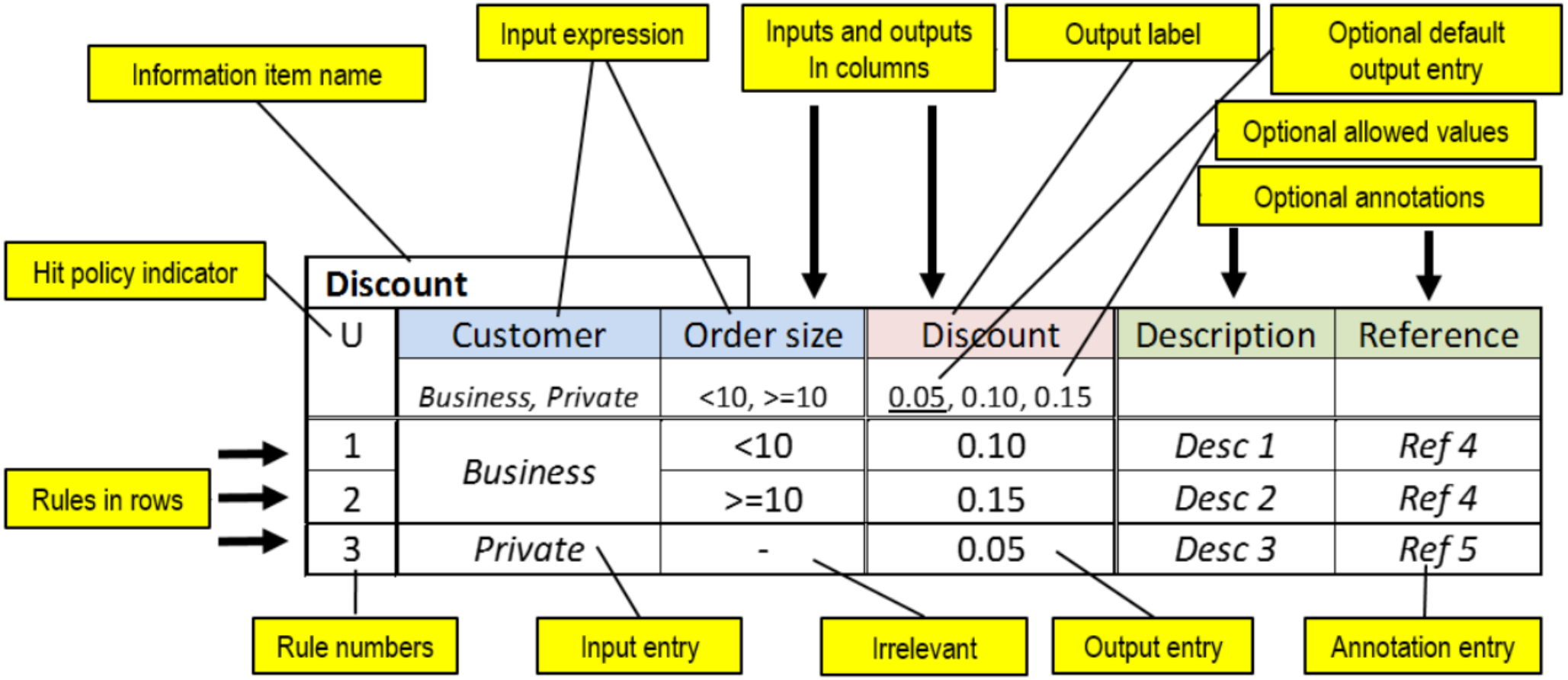
2. **Decision logic**

Applicant Risk Rating			
U	Applicant Age	Medical History	Applicant Risk Rating
1	> 60	good	Medium
2		bad	High
3	[25..60]	-	Medium
4	< 25	good	Low
5		bad	Medium

Decision Requirements Graph



DMN decision table



Discount

U	Customer	Order size	Discount	Description	Reference
	Business, Private	<10, >=10	0.05, 0.10, 0.15		
1	*Business*	<10	0.10	*Desc 1*	*Ref 4*
2		>=10	0.15	*Desc 2*	*Ref 4*
3	*Private*	-	0.05	*Desc 3*	*Ref 5*

Figure 8.2: Decision table example (horizontal orientation: rules as rows)

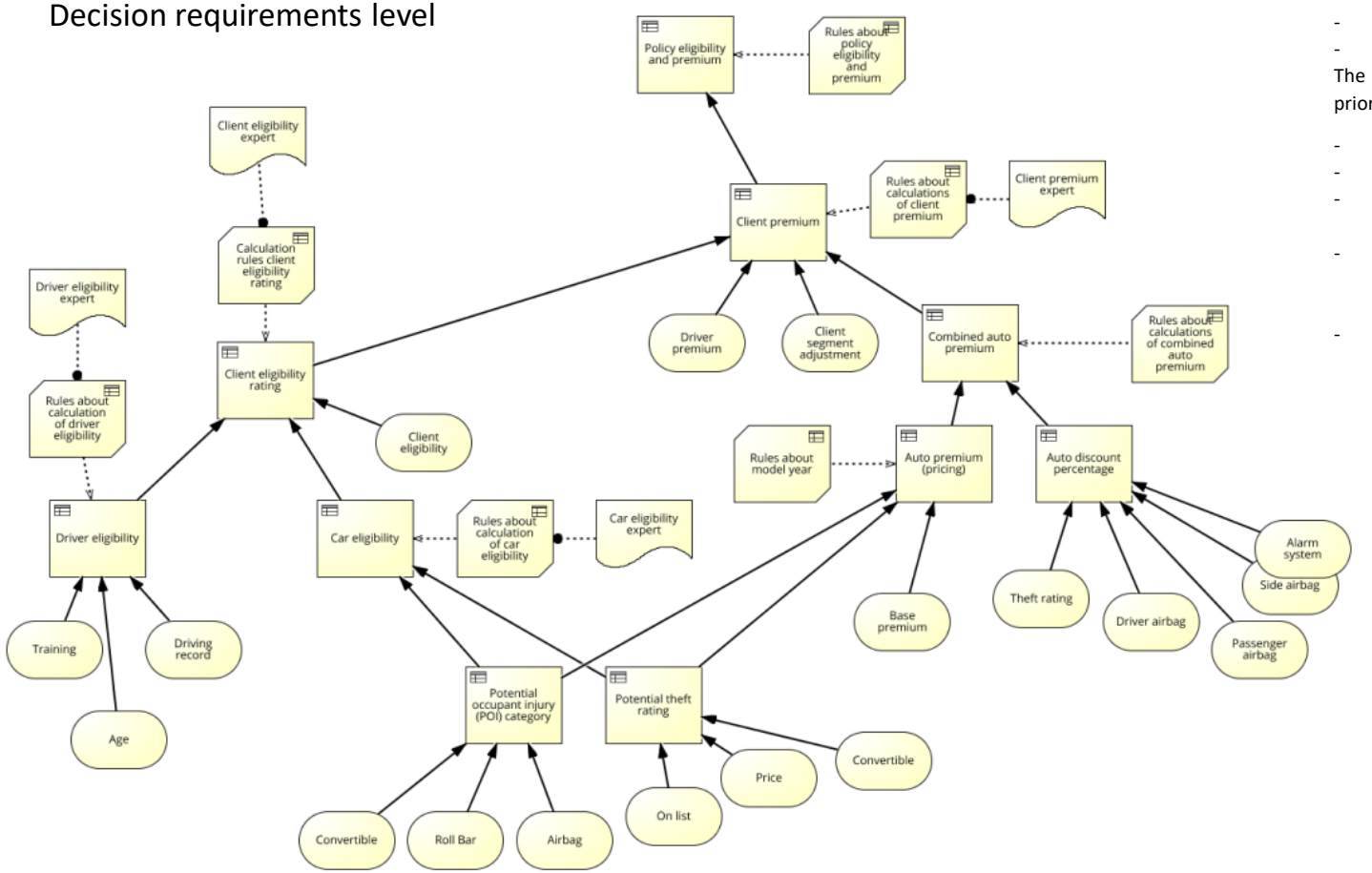
Decision capability

Why would we only care about the (big) data or the processes?
Decisions are important for business

- **What is the decision?**
  - eligibility, price, insurance, theft rating, customer offer, retention, supplier selection, hire, credit, …
- **Who owns the decision?**
- **Who makes the decision every day?**
- **What triggers the decision?**
- **How can we improve the decision?**
- **What is required to make this decision?**
  - Information requirements
  - Knowledge sources (regulations, analytics, expertise)
  - Other decisions
  - Decision logic

Example: Insurance premiums

Decision requirements level



The **Client Premium** is based on

- The client segment adjustment (based on client data)
- The combined auto premium (based on car data)
- The driver premium (based on driver information)

The **Potential Theft (PT) Category** is subject to the following rules (in decreasing priority)

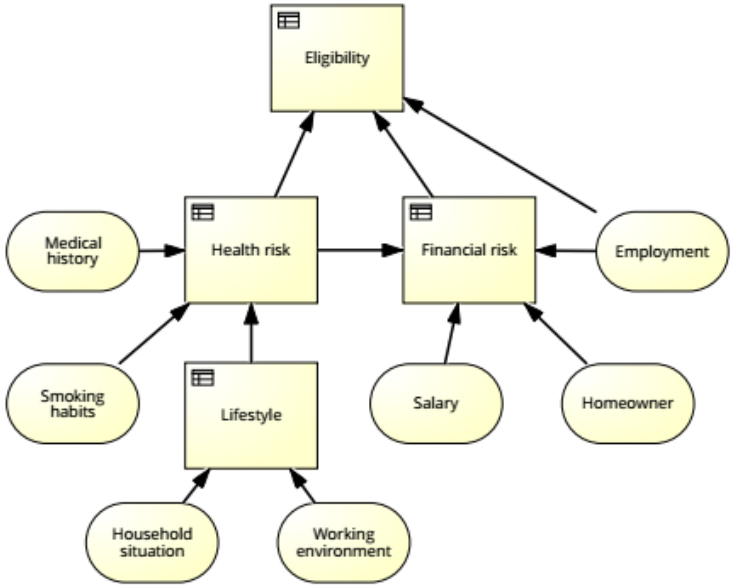
- If the car is a convertible, then its potential theft rating is high.
- If the car's price is greater than $40.000, then its potential theft rating is high.
- If the car model is on the list of 'High Theft Probability Auto', then its potential theft rating is high.
- If all of the following are true, , then its potential theft rating is moderate:
  - The car's price is between $22.000 and $40.000.
  - The car model is not on the list of 'High theft potential Probability Auto'.
- If all of the following are true, the car's potential theft rating is low:
  - …

Decision logic level

	Inputs					Outputs
U	On list		Price		Convertible	Theft Rating
	{YES,NO}		{<22 000,22 000-40 000,>...		{YES,NO}	{High,Medium,Low}
1	= YES		-		-	High
2	= NO	=	<22 000	=	YES	High
3	= NO	=	<22 000	=	NO	Low
4	= NO	=	22 000-40 000	=	YES	High
5	= NO	=	22 000-40 000	=	NO	Medium
6	= NO	=	>40 000		-	High

Issues DMN solves

- ***Separating decisions and processes***
  - Using a standard modeling notation.
- ***Decision table types***
  - Recognize, and unambiguously exchange.
- ***Decision modeling methodology***
  - Keep the insights of the past and avoid confusion.
- ***Separating decision structure and decision logic***
  - *A*llows to model decision relations, even if not all logic is in tables.
- ***Standard notation for exchange and implementation***
  - Strict notation and simple expression language (FEEL).



Decision requirement diagram for the credit eligibility decision

Executing decision models

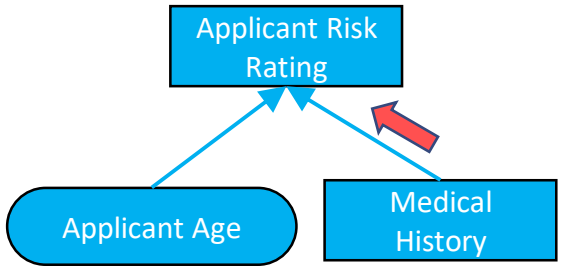
Decision execution

Use this executable DMN specification to **make a decision**

- Straightforward processing of input cases, with all inputs made available (subdecisions first)

Applicant Risk Rating

Applicant Age	Medical History	Applicant Risk Rating
> 60	good	Medium
	bad	High
[25..60]	-	Medium
< 25	good	Low
	bad	Medium



If the Applicant Age is >60 and the Medical History is bad, then the Applicant Risk Rating is High.

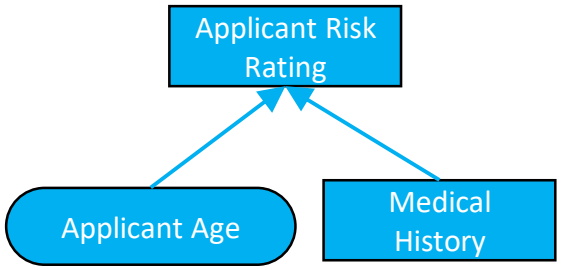
Decision execution

Use this executable DMN specification to **make a decision**

- Straightforward processing of input cases, with all inputs made available (subdecisions first)
- **Smart** processing of input cases, with only relevant inputs retrieved

Applicant Risk Rating

Applicant Age	Medical History	Applicant Risk Rating
> 60	good	Medium
	bad	High
[25..60]	-	Medium
< 25	good	Low
	bad	Medium



For an Applicant Age [25..60] we do not need to look up the Medical History.

Irrelevant information items, or irrelevant dependencies (subdecisions) can be skipped depending on the specific case.

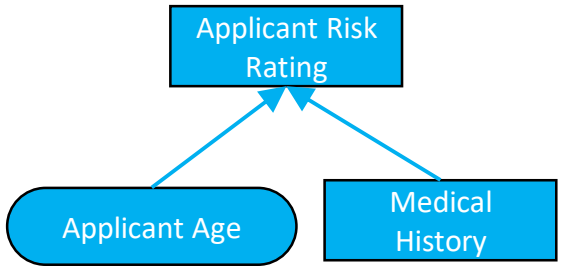
Decision execution (even with unknowns)

Use this executable DMN specification to **make a decision**

- Straightforward processing of input cases, with all inputs made available (subdecisions first)
- **Smart** processing of input cases, with only relevant inputs retrieved
- Reasoning with some information items, and others **unknown:** what is the output range?

Applicant Risk Rating

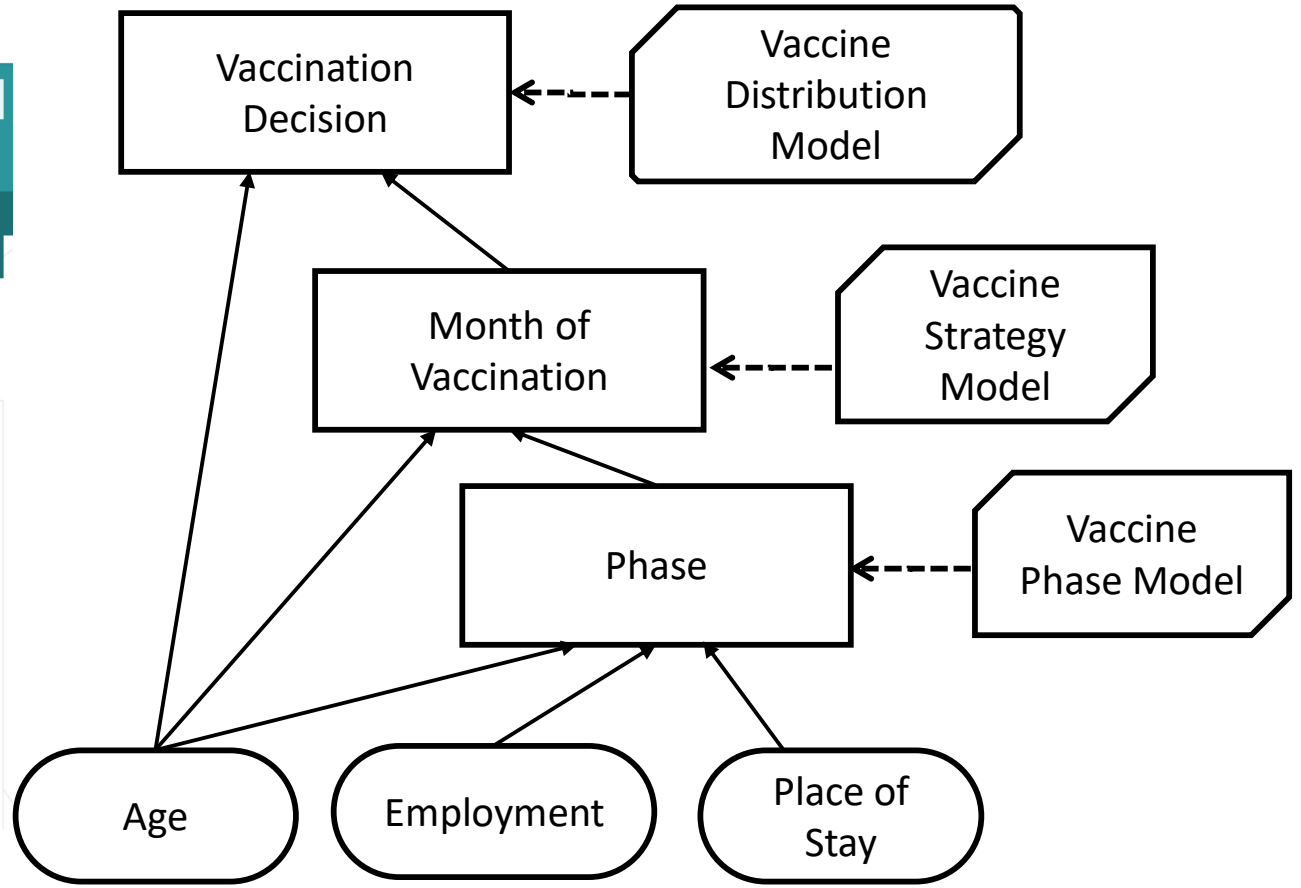
Applicant Age	Medical History	Applicant Risk Rating
> 60	good	Medium
	bad	High
[25..60]	-	Medium
< 25	good	Low
	bad	Medium



For an Applicant Age<25, the only possible outcomes are Low or Medium.

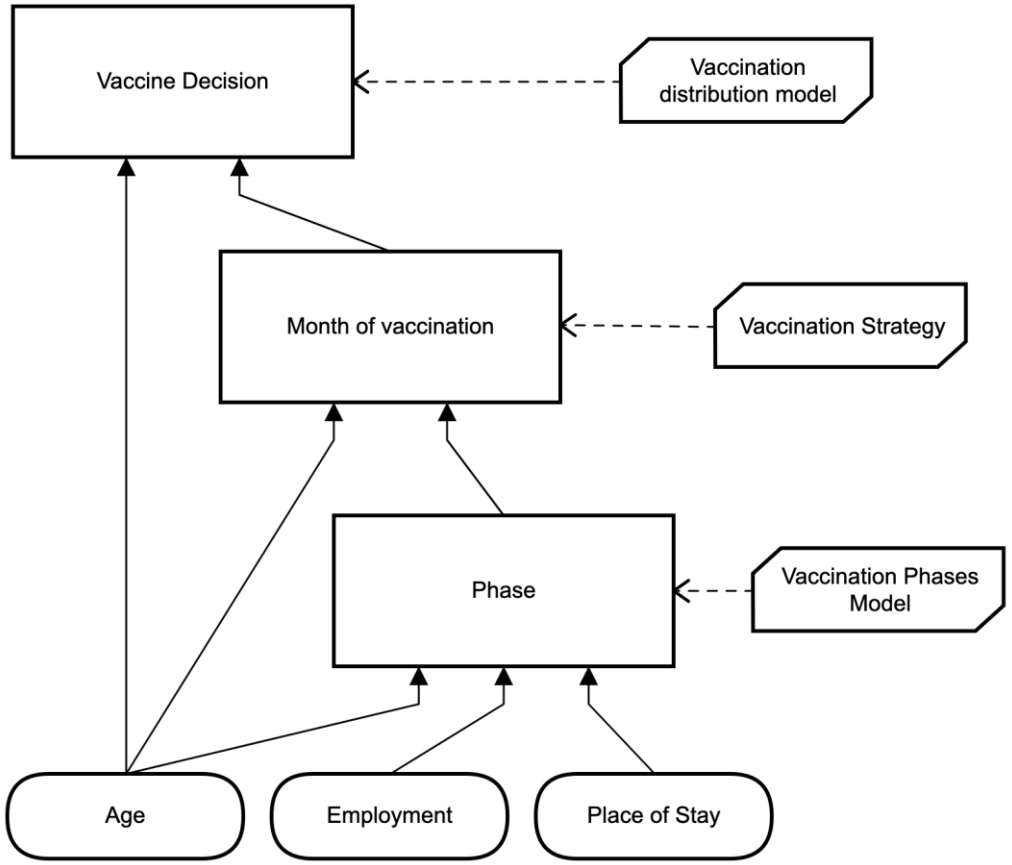
This already gives a range of possible decision outcomes.

Example: Covid Vaccination Decisions in Belgium



https://www.info-coronavirus.be/en/vaccination/

Decision model for Determining Vaccination Decision



Vaccine Decision

U	Age	Month of Vaccination	Vaccine	Number of doses
1	[0..16]	-	No vaccination	0
2	]16..18[	-	Pfizer	2
3	[18..100]	January	Pfizer	2
4		February or March	Moderna	2
5		April	AstraZeneca	2
6		May	Johnson & Johnson	1
7		June	AstraZeneca	2
8		July	Moderna	2

Phase

U	Place of Stay	Employment	Age	Phase
1	"Hospital" or "Residential care center"	-	-	Phase 1
2	"At home"	Healthcare	-	Phase 1
3		Police	-	Phase 2
4		Retired or Other	[0..65[	Phase 3
5			[65..100]	Phase 2

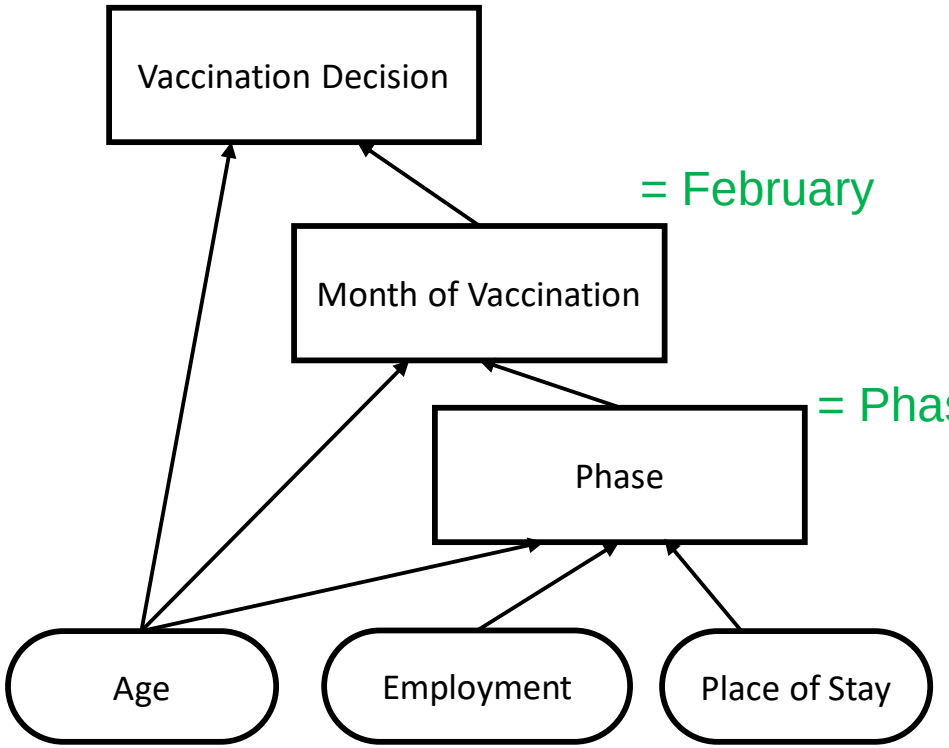
Month of Vaccination

U	Phase	Age	Month of Vaccination
1	Phase 1	[49..100]	January
2		[0..49[	February
3	Phase 2	[49..100]	March
4		[0..49[	April
5	Phase 3	[49..100]	May
6		[28..49[	June
7		[0..28[	July

Decide Vaccination

User: I am a 27-year-old healthcare worker living at home, which vaccine do I get?

= Moderna, 2 doses



= February

= Phase 1

Vaccine Decision

U	Age	Month of Vaccination	Vaccine	Number of doses
1	[0..16]	-	No vaccination	0
2	]16..18[	-	Pfizer	2
3	[18..100]	January	Pfizer	2
4		February or March	Moderna	2
5		April	AstraZeneca	2
6		May	Johnson & Johnson	1
7		June	AstraZeneca	2
8		July	Moderna	2

Phase

U	Place of Stay	Employment	Age	Phase
1	"Hospital" or "Residential care center"	-	-	Phase 1
2	"At home"	Healthcare	-	Phase 1
3		Police	-	Phase 2
4		Retired or Other	[0..65[	Phase 3
5			[65..100]	Phase 2

Month of Vaccination

U	Phase	Age	Month of Vaccination
1	Phase 1	[49..100]	January
2		[0..49[	February
3	Phase 2	[49..100]	March
4		[0..49[	April
5	Phase 3	[49..100]	May
6		[28..49[	June
7		[0..28[	July

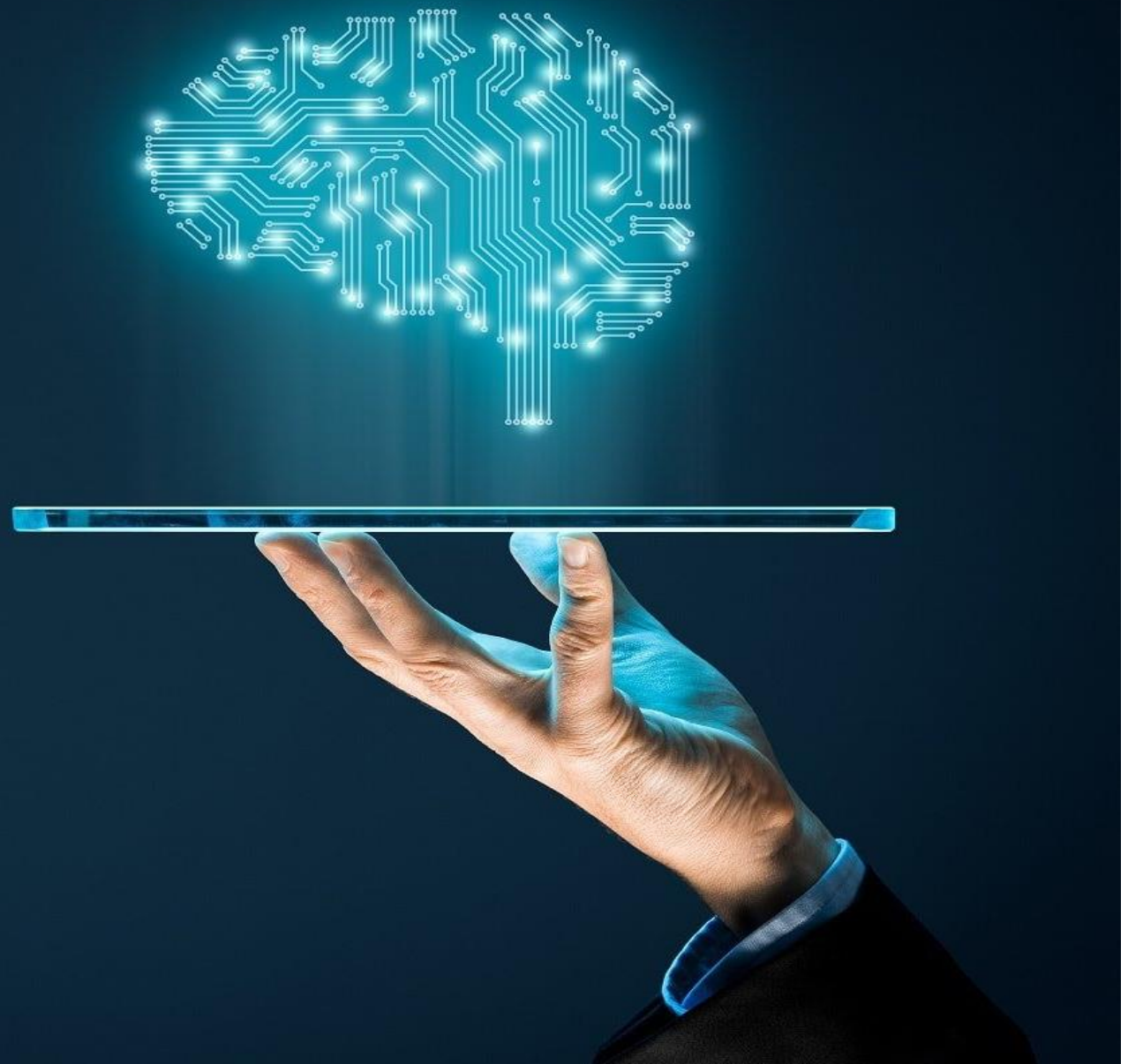
= 27-year-old

= Healthcare

= At home

Decision outcome: 2 doses of Moderna in February

Explainable Decision Models



For every automation, there is a hidden service cost

- Every automation/digitalization removes physical contact

→ Removes opportunity to ask questions

  - Customers have many questions, and it is difficult to ask questions when service is done online: Why, Why not, How, What to do?
  - How to give advice, support, explanation?

→Call centers, FAQs, online documentations, chatbots, …

  - Do not always meet customer demands: waiting times, incomplete, impersonal
  - High labor cost

You have built a model/application/process to decide

But after (or even before) the decision, the customer has some questions !!!

What do you do?

- Ignore the questions. It's all on the website
- Produce some superficial, glossy brochures
- Organize a call center and let them wait
- Build a support application (double work)
- Build some FAQ pages
- Create some user self help groups
- Hope that machine learning will solve your problem
- Refer users to social media
- Use WhatApp to refer them to email that sends them to phone …
- Use some kind of dedicated chatbot solution
- Hope GPT-5 will solve your problem

Problems with chatbots, web pages, FAQs

- They are labour intensive
- They duplicate the decision knowledge
- They are problem-specific
- Hard to maintain
- Every problem domain requires a new chatbot
- They only contain the basic questions and answers
- They hardly refer to customer data
- They only deal with simplified cases

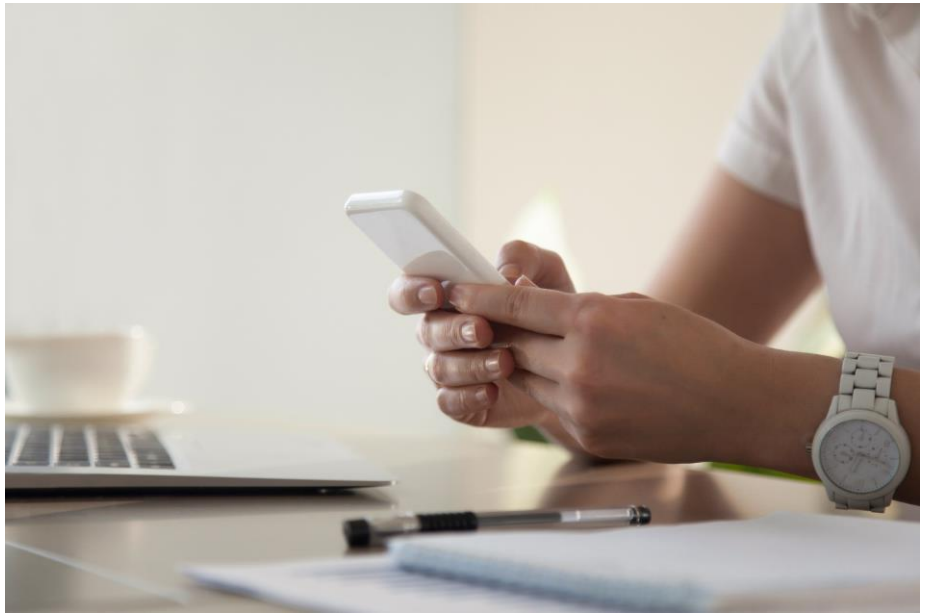
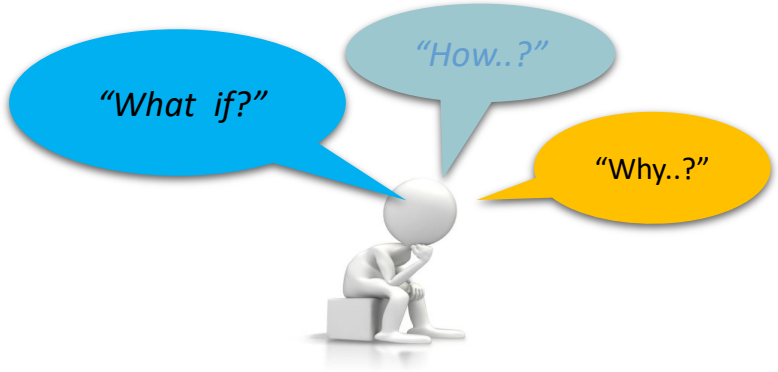


Image by yanalya on Freepik

You built a decision model!

Now what?

- **Use this executable specification to allow flexible reasoning (What-If, How, How-To) (DecisionCAMP 2021)**



What questions can you ask a decision model?

***What** is my decision?*

*"**How to** reach a decision outcome?"*

"**Why** is this the outcome?"

"**What if** ..?"

"What vaccination do I receive, if I am a 27 years old health care worker?"

"I have heard Moderna is better, How can I get Moderna as my vaccination?"

"Why is vaccination Pfizer suggested to me?"

"If I am retiring next month, what would be my vaccination?"

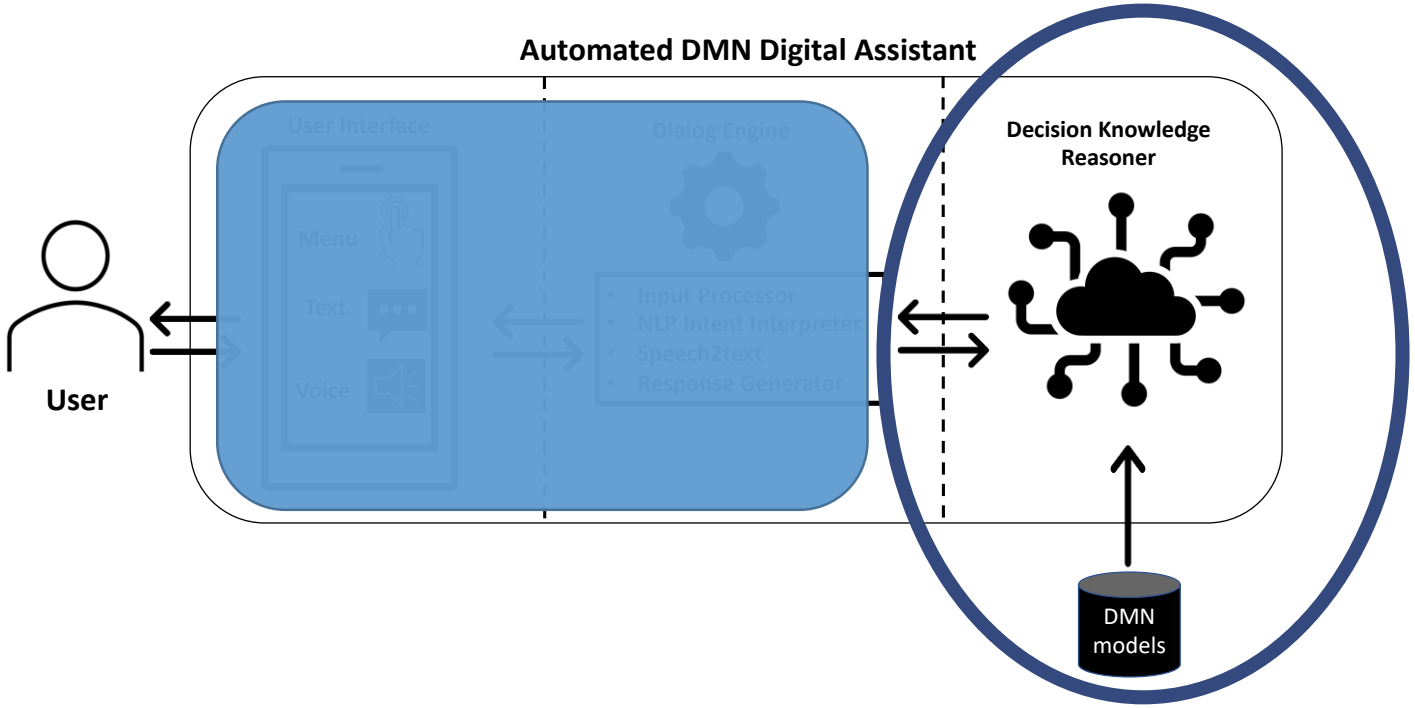
Decision and explanation scenarios

- **Make decision**
  - with complete information. [one outcome]
  - with only relevant information. [one outcome]
  - with incomplete information. [list or range of possible outcomes]
- **How to:** Seek goal
  - reasoning towards a given outcome. [list of satisfactory rules and conditions]
- **Explain Why:**
  - why did I receive this outcome? [list of rules leading to the outcome]
- **What if:**
  - one of the inputs changed? [rerun the decision]
- **Sensitivity**
  - Which values can I change and still get the same output? [relevant input ranges for the given outcome]

You built a decision model!

Now what?

- **Use this executable specification to allow flexible reasoning (What-If, How, How-To) (DecisionCAMP 2021)**
  - Given some input information items, or an obtained outcome, use the reasoning mechanism for other forms of inferencing. Or simulate with other input data.

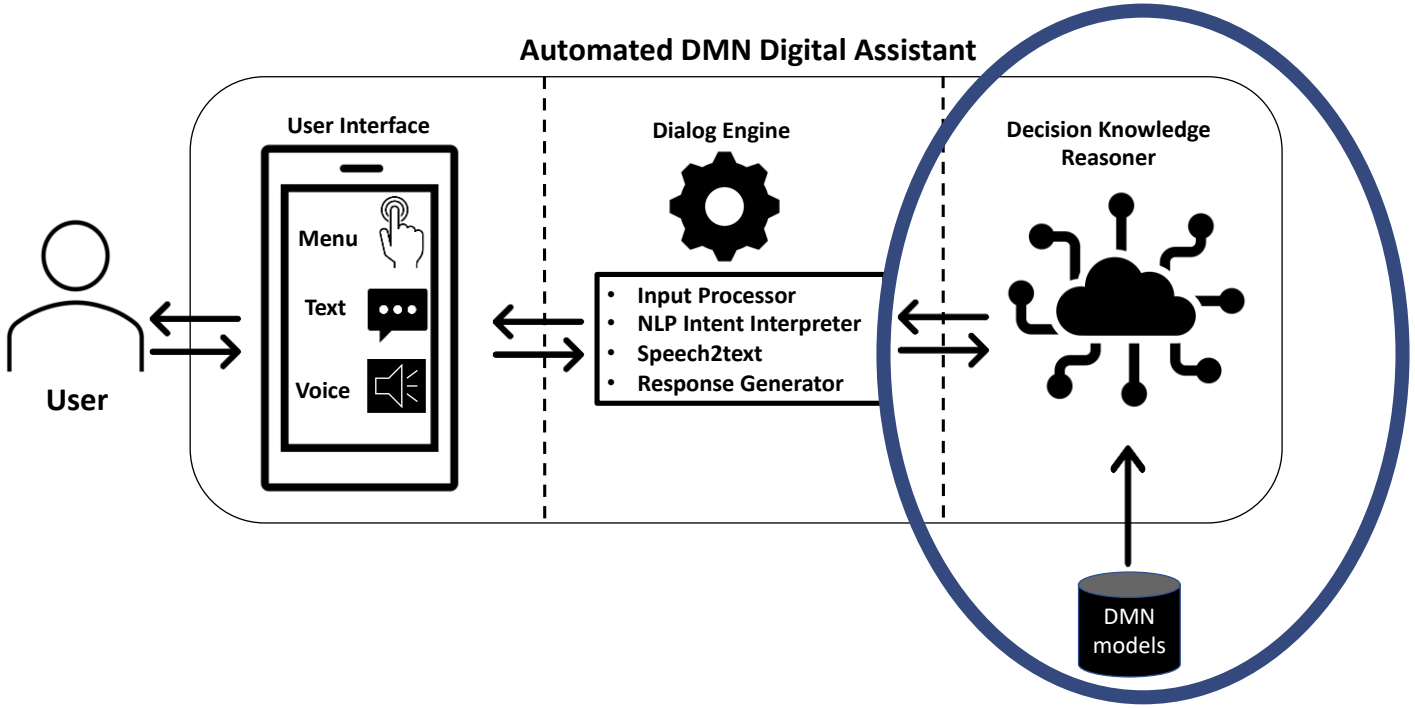


You built a decision model!

Now what?

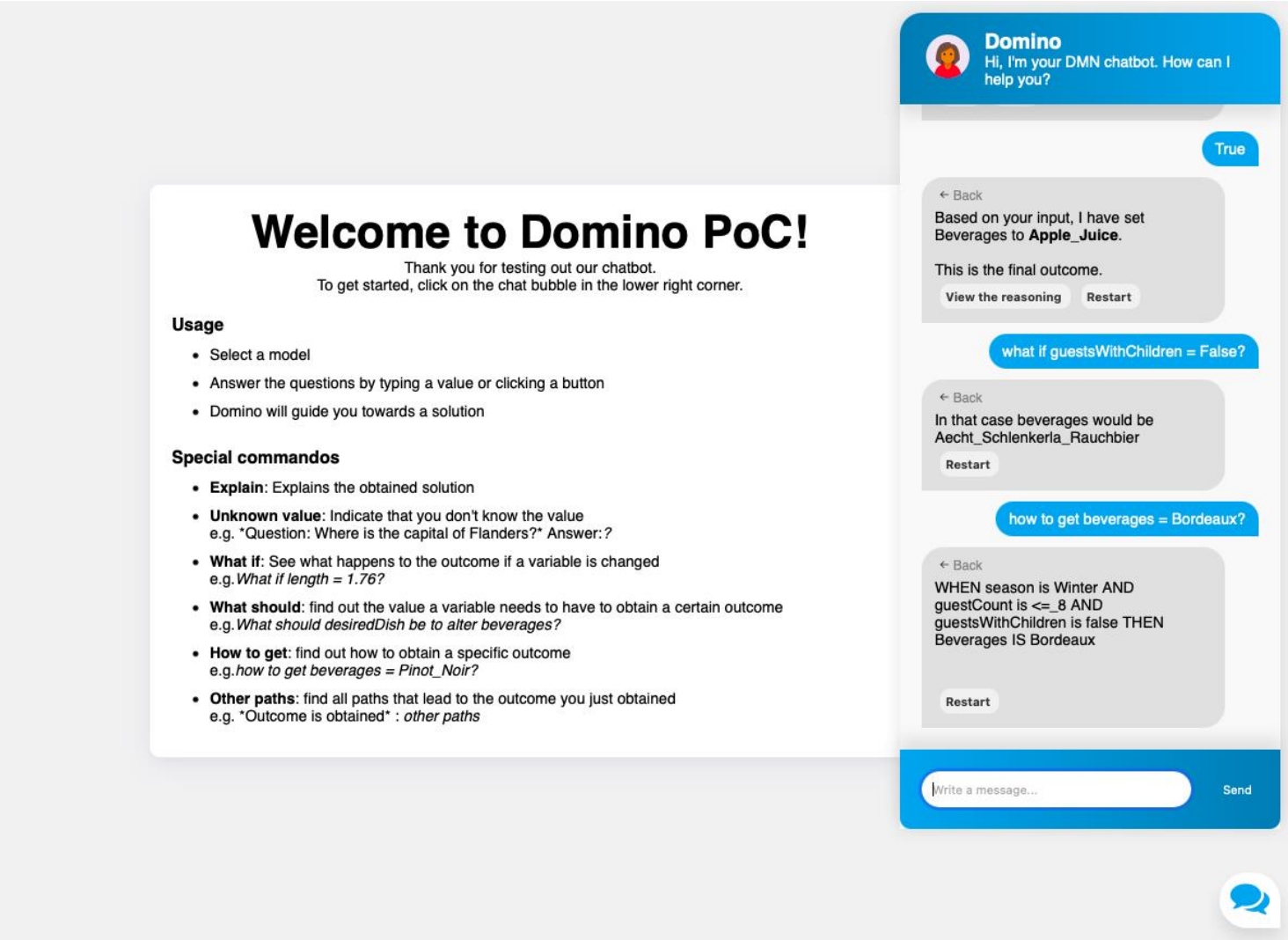
- **Use this executable specification to allow flexible reasoning (What-If, How, How-To) (DecisionCAMP 2021)**
  - Given some input information items, or an obtained outcome, use the reasoning mechanism for other forms of inferencing. Or simulate with other input data.

- **Allow the user to communicate with the model using a generic chatbot**
  - Menu, text or voice
  - NLP for intent/scenario identification
  - **No-code generic chatbot**
  - Not problem specific
  - No knowledge duplication: the DMN model IS the knowledge
  - Explainable decisions



You built a decision model!

Now what?



You built a decision model!

Now what?

- **Decide**
- **Explain**
- **Evaluate**
- **Check**
- **Analyze**
- **Monitor**
- **Predict**
- **…**

- ✓ **Use this executable specification to make a decision**
- ✓ **Use this executable specification with some inputs unknown (or not specified) to obtain a decision range (RuleML 2021)**
- ✓ **Use this executable specification to allow flexible reasoning (What-If, How, How-To) (DecisionCAMP 2021)**
- ✓ **Explain the decision model by allowing the user to communicate with the model using a chatbot (DecisionCAMP 2021)**
- ✓ **Evaluate model quality and complexity (DecisionCAMP 2019)**
- ✓ **Check if the decision model is compliant with existing rules and regulations**
- ✓ **Examine fairness of the decision knowledge**
- ✓ **Static Decision Model Analysis**
- ✓ **Policy Evaluation based on historical data**
- ✓ **Real-time monitoring**
- ✓ **Data-driven Policy evaluation and prediction**

Recent developments

Adding more AI

- **Modeling stage**
  - Mining decision models from event logs (decision mining)
  - Extracting decision models from text
    - Pattern-based extraction
    - Extraction using Deep Learning techniques (BERT)
      - Decision Dependency extraction
      - Decision Logic Extraction
    - Extraction using GPT
      - Decision Dependency extraction
      - Decision Logic Extraction

- **Execution stage**
  - Decision making using ChatGPT
  - Integrating GPT with DMN for explainability

DMN Extraction from Text

- ***TEXT :***

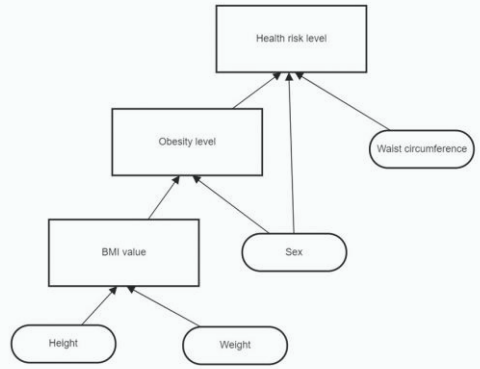
1. Decision Dependencies

   - *"The health risk level of a patient should be assessed from the obesity level, waist circumference and the sex of the patient. Furthermore, the degree of obesity should be determined from the BMI value and sex of the patient. Patient's height and weight are considered to calculate his BMI value."*

2. Decision Logic

   - *"When the patient's sex is a male and his BMI value is in between 25 and 29.9, then his obesity level is normal."*
   - *"If patient' sex is female and BMI value is above 25.0 and less than 30, then obesity level is overweight . Where as, If BMI value is 30.0 or higher, obesity level falls within the obese I range."*

- ***DRD:***



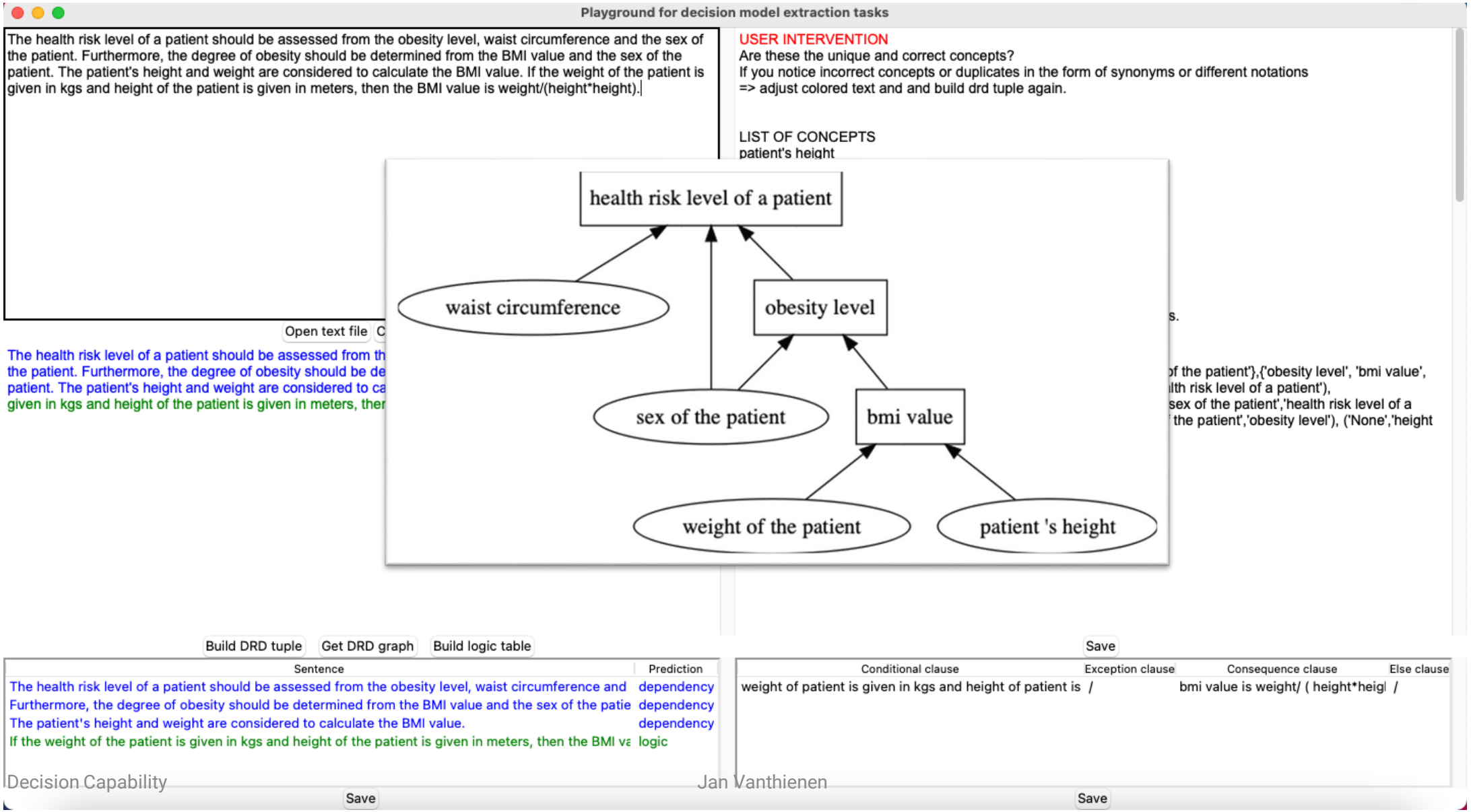
- ***Rules:***

	Inputs		Outputs
5	[25..29.9]	"Male"	"Normal"
6	[25..29.9]	"Female"	"Overweight"
7	[30..34.9]	-	"Obese I"

BMI Guidelines: https://www.nhlbi.nih.gov/files/docs/guidelines/ob_gdlns.pdf

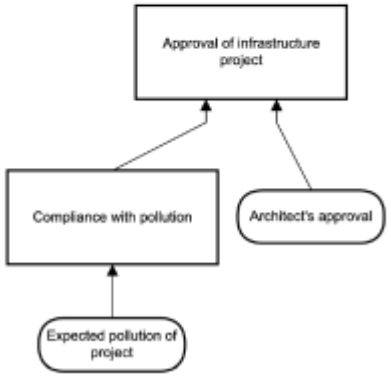
*Clinical Guidelines on the Identification, Evaluation, and Treatment of Overweight and Obesity in Adults

Decision dependency extraction from text (DL)



Decision dependency extraction from text (GPT)

Example 1. With the recent adoption of stricter environmental laws, the approval of infrastructure projects needs to take into account the amount of additional nitrogen pollution each project would cause in the environment. The maximum allowed pollution has been set at a maximum of 14 grams per hectare per year. Any project which would cause this bar to be exceeded can not be approved. Additionally, any infrastructure project also needs to be approved by an architect before a final approval can be given.



Approval of infrastructure project			
U	Compliance with pollution	Architect's approval	Approval of infrastructure project
1	True	True	True
2	True	False	False
3	False	True	False
4	False	False	False

Compliance with pollution		
U	Expected pollution of project	Compliance with pollution
1	>=14	False
2	<14	True

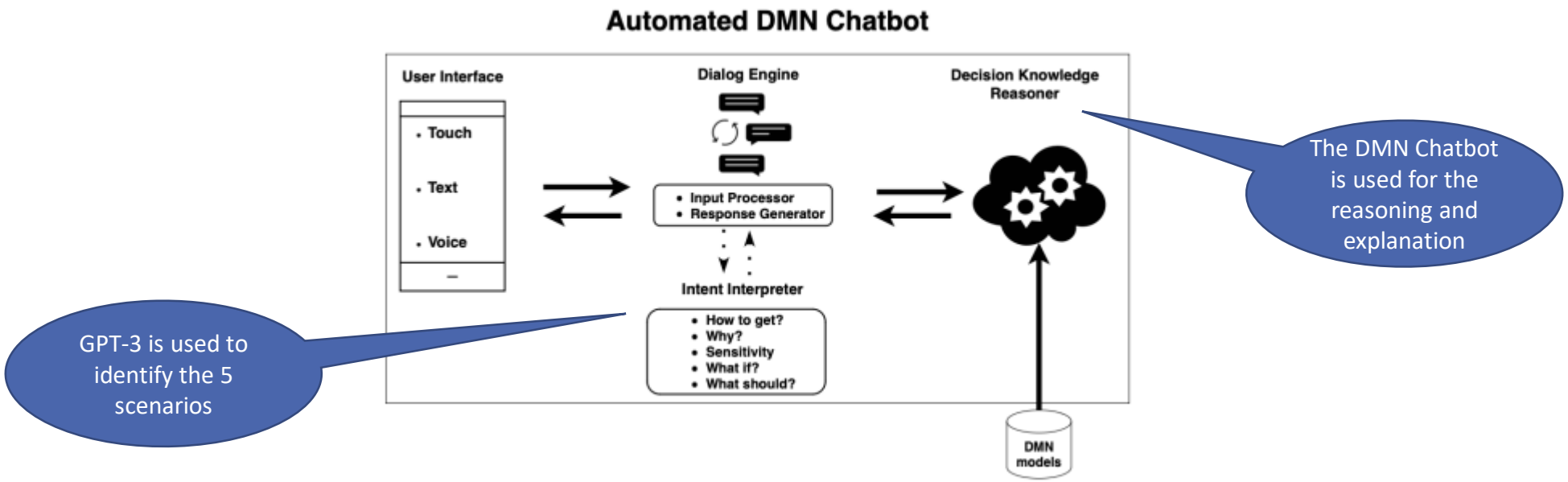
Table 3. Averages of GPT-3 and BERT models over all cases

Model	Decision and Input Data						Decision Dependencies					
	#gold	#pred	#ok	P	R	F1	#gold	#pred	#ok	P	R	F1
Total_GPT-3_0	38	40	33	**87%**	**86%**	**86%**	32	34	25	83%	74%	79%
Total_GPT-3_0.2		42	33	82%	**86%**	84%		36	25	78%	74%	76%
Total_GPT-3_0.4		40	33	87%	**86%**	**86%**		34	28	**89%**	**87%**	**88%**
Total_GPT-3_0.6		39	31	84%	80%	82%		33	23	80%	66%	72%
Total_GPT-3_0.8		39	31	85%	80%	82%		32	24	83%	70%	76%
Total_GPT-3_1		41	31	79%	80%	79%		34	24	74%	70%	72%
Average_BERT		40	37	**93%**	**97%**	**95%**		37	31	87%	**97%**	**91%**

Table 3 reports the averages for each analyzed temperature of GPT-3 over all cases as well as the average of the BERT-based approach.

When comparing the BERT approach to the best performing GPT-3 model with temperature 0 and 0.4, BERT still outperforms GPT-3 in all but one metric namely precision of decision dependencies

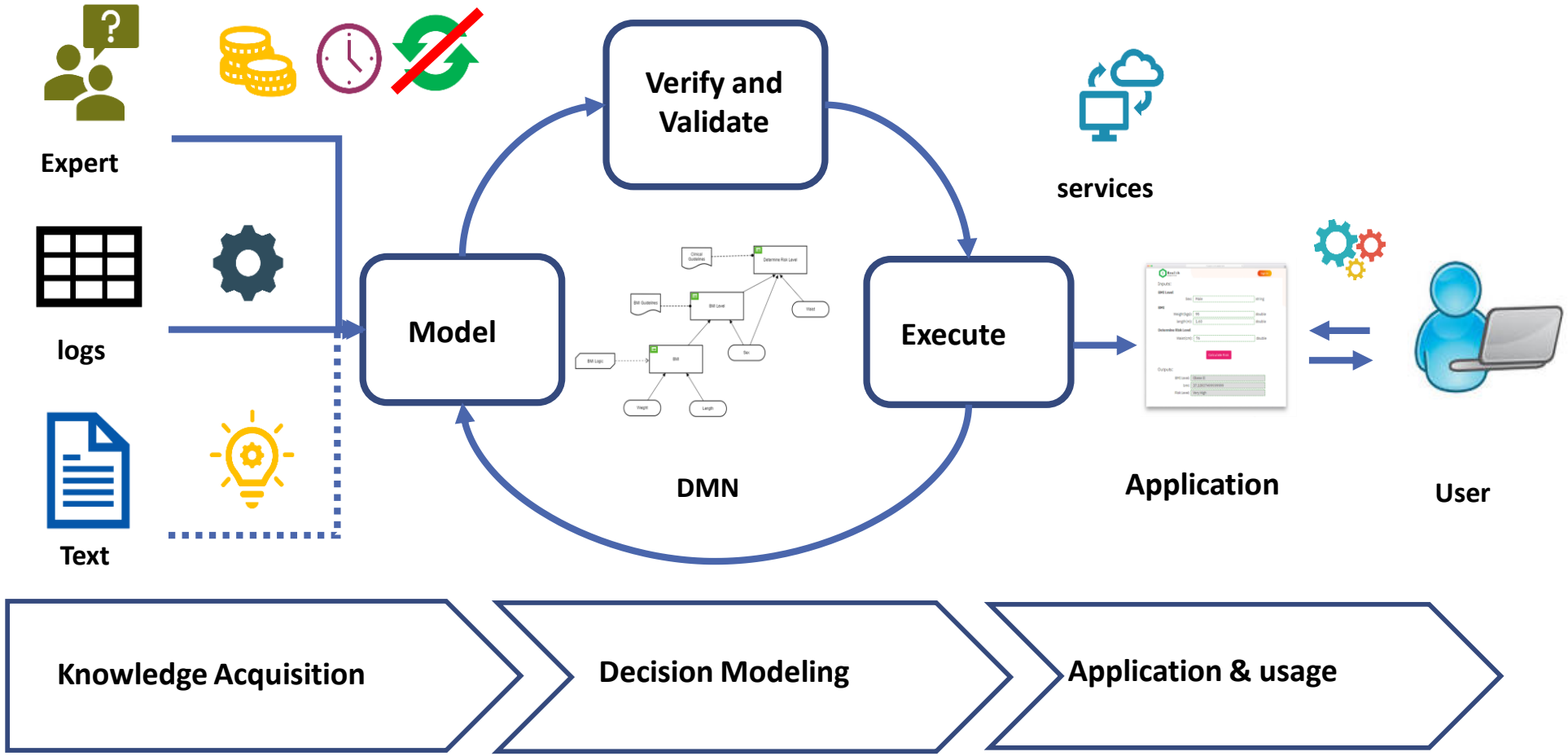
Integrating GPT with DMN for explainability



When GPT-3 has to answer the 5 questions (on two examples), only 15% of the explanations are correct and complete.

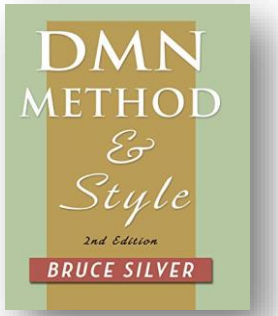
When GPT-3 is used just to identify the type of question (leaving the reasoning to the knowledge reasoner in the DMN chatbot), the results are 100% correct on these (small) examples.

Business Decision Management from A to Z

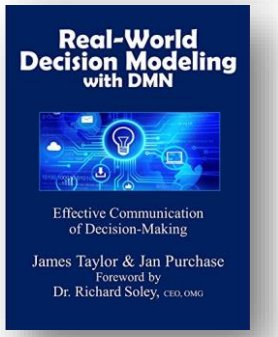


More information

- Etikala, V., Goossens, A., Veldhoven, Z.V., Vanthienen, J.: Automatic generation of intelligent chatbots from dmn decision models. In: International Joint Conference on Rules and Reasoning. pp. 142–157. Springer (2021)
- Goossens, A., Maes, U., Timmermans, Y., Vanthienen, J.: Automated intelligent assistance with explainable decision models in knowledge-intensive processes. In: Business Process Management Workshops: BPM 2022 International Workshops, pp. 25–36, Springer (2023)
- Vandevelde, S., Etikala, V., Vanthienen, J., Vennekens, J.: Leveraging the power of IDP with the flexibility of DMN: A multifunctional API. In: Rules and Reasoning - International Joint Conference, RuleML+RR. pp. 250–263. Springer (2021)
- Vanthienen, J.: Decisions, advice and explanation: an overview and research agenda. In: A Research Agenda for Knowledge Management and Analytics, pp. 149–169. Edward Elgar Publishing (2021)



DMN Method and Style, 2nd Edition: A Business Practitioner's Guide to Decision Modeling (2019)
by Bruce Silver
https://www.amazon.com/gp/product/B07P7JD28F/ref=dbs_a_def_rwt_hsch_vapi_taft_p1_i1



Real-World Decision Modeling with DMN (2016)
by James Taylor, Jan Purchase
https://www.amazon.com/gp/product/B01N1OALY4/ref=dbs_a_def_rwt_bibl_vppi_i3

Research topics: Decision & Process Modeling, Mining and Analytics
http://janvanthienen.net/research/