Modelling challenges on the field of IT-security (based on artificial intelligence)

László Pitlik (sen), László Pitlik (jun), Marcell Pitlik, Kodolanyi Janos University / MY-X-research-team

**Abstract**

IT-security is a complex phenomenon. Its complexity level leads the authors (quasi automatically) to the strategy in case of a planning process that they suffer under all the negative force fields of the problem called “magic-of-words”. The magic-of-words as such means a real problem can never be descripted deep enough based only on texts. ChatGPT is a characteristic example for this trap: chatGPT is gossiping and gossiping (let alone: using virtual facts and irrational rules – at least partially) till the human actor do not ask for a source code. The magic-of-words as one of the oldest knowledge representation forms is therefore not capable arbitrary enough to approximate the principle of Knuth (c.f. knowledge/science is what can be transformed into source code – each other human activity is a kind of art). The magic-of-words-phenomenon is a kind of artistic performance – like this abstract.

This article is a case study about two real modelling processes in frame of educational activities where Students (preferring IT-security as a potential specialization of the BPROF-education) had the first task to declare both a title for the own thesis and an abstract describing the own thesis (based on the following structure: tasks, goals, targeted groups, utilities, solutions, future, etc.). Based on the title and the abstract, the next task was: identifying at least one single OAM-oriented (decision-making) problem as a focus of an application (because a BPROF-thesis is a publication about the own application/source-code - being complex enough for a real service). An OAM is an object-attribute-matrix containing log-based raw data about comparable objects. Attributes should be such phenomena where the derivation/measurement of the real value for a given object can (at least theoretically) be automated.

If an OAM is imaginable (it means it can contain random, but realistic values), the next task is: deriving detailed demonstrated data processing possibilities as test-cases for source-codes in future). A data processing possibility is a model (c.f. knowledge representation) where each step of the more or less complex data processing can be realized for the random data in the OAM and finally, the results can be (in form of text) interpreted as a kind of operative decision-making support. The hermeneutics (interpretation text) is again an action in frame of the magic-of-words. This arch (starting from the magic-of-words – then declaring operative programmable data processing steps – and landing again in the world of the magic-of-words) is the general pattern of the human innovation processes. Therefore, the magic-of-words seems to an essential part of the operativity. Without the massive subjectivities generated by the magic-of-words, the clarifications of the source-codes can not be achieved and evaluated as the highest value in the Knuth-universe of operativity/operability.

If a Student is capable of constructing an OAM-based solution (mostly involving a lot of data processing techniques like naïve aggregation of single impacts through averages and/or regression and/or chained similarity analysis as a specific part of AI – in co-operation with the conductors/teachers) then the last task is: adapting the own experiences for a new challenge where the same thesis should be supported through one or more new OAM-variation(s) and adequate data processing and interpretation steps as a proof of the needed sovereignty concerning creation challenges in future enterprises/institutions/projects.

Demo for further details: [System-modeling – Miau Wiki (my-x.hu)](https://miau.my-x.hu/mediawiki/index.php/System-modeling) - <https://miau.my-x.hu/mediawiki/index.php/System-modeling>

# Introduction

The BPROF education on the field of computer science means bachelor of profession (Computer Science Operational Engineering). The education is massive practice oriented. Therefore, it is relevant that each course should be capable of supporting the final thesis and the needed computer program (application) needed behind it. The system modelling course can be interpreted in diverse ways, but a trivial interpretation is the modelling of the system affected by the above-mentioned application.

The BPROF educations have different specialisations (like IT-security) where each specialisation as such is a quasi unlimited domain. Therefore, most of the keywords concerning a specialisation are still too vague in order to define the necessary levels of the operationalism (c.f. Annexes: Demonstration materials: Magic-of-words).

This publication summarizes the experiences of two projects:

* <https://miau.my-x.hu/mediawiki/index.php/System-modeling#Expert_system_for_fine-tuning_IT-security_systems_.28ES-FITS.29>
* <https://miau.my-x.hu/mediawiki/index.php/System-modeling#Automation_of_Incident_Response_Planning_in_IT_Security_.28AIR-P-ITS.29>

Further details can be seen in the MIAU-WIKI article about a system modelling course: <https://miau.my-x.hu/mediawiki/index.php/System-modeling> (where MIAUWIKI is a collaborative frame for education and Student’s publication – c.f. <https://miau.my-x.hu/> with own online modelling tools like COCO (component-based object-comparison for objectivity - (MY-X FREE: <https://miau.my-x.hu/myx-free/>).

Important declaration: descriptive thesis can not be accepted! If chatGPT is capable of writing a thesis (and chatGPT is capable of doing this), then the expected level of a thesis should be set higher. On the other hand: an application should have real input/test data/cases, real customers/users, real utilities.

# Identification of decision situations

The first step of the modelling is: identifying decision necessities because an application (concerning the Knuth’s principle[[1]](#footnote-1)) is a kind of automated activity pool. The title and the abstract of a thesis (=application description) contains a lot of keywords being complex enough to expect a deeper interpretation. The above-mentioned two project-URLs present a set of docx-files demonstrating the keywords and their high-levelled complexity – being incapable of characterising a problem for modelling in a realistic way. The keyword-oriented descriptions are examples of the magic-of-words-situations. Footnotes and/or remarks can be inserted into the word-of-magic-streams if these remarks make possible to involve more technocratic terms (like expert systems: or more precise inductive expert system or manual-driven expert systems).

In order to be able to continue to the next step in the modelling phase, it is necessary to choose single keywords like

* password-attack and/or
* email-phishing, etc. (see MIAUWIKI: docx-files in the system modelling article).

This identification is a kind of arbitrary activity concerning the affected Students and conductors. The more is the number of arbitrary focused potential problems (micro-projects), the higher is the potential to select a subset of them having real connections to each other and to the keywords (to the magic-of-words-force-fields).

# Operative micro-projects

In this chapter, two arbitrary micro-projects will be presented. Both of them are basing on at least theoretically existing (in the reality randomized generated) log-data (structured in form of OAMs):

## Email-phishing

Email-phishing can be defined in different ways. A potential dramaturgy of a decision situation can be seen in the following construct (see Figure#1):

* a user can receive emails (objects being identified through the email-address of the sender – see row-headers in Fugure#1)
* the naïve user can expect an IT-security service (c.f. firewall, spam-filter, etc.) where potential dangerous emails can be filtered in an automated way
* this automated way can be a kind of risk estimation robot:
	+ this robot identify/use keywords (see column-headers – see Figure#1)
	+ in frame of text mining services, the number of these keyword in an email can be derived (see cells with borders in Figure#1)
	+ the number of critical keywords (and/or phrases and/or distances of words, etc.) can be interpreted based on trivial rules (c.f. direction codes in Fugure#1 where 0 = the-more-the-more, and 1=the-less-the-more based on the last RANK()-function-parameter in MS Excel): the more is the number of a critical word, the higher is the risk/danger concerning the particular email (object)
	+ the robot has to be capable of aggregating the sub-risk-potentials of the different words (attributes)



Figure#1: Raw data for risk generation on the field of email-phishing (source: <https://miau.my-x.hu/mediawiki/index.php/System-modeling> - <https://miau.my-x.hu/bprof/2023/data_for_email_phishing_risk_estimation_3.xlsx> – Sheet: “raw OAM”)



Figure#2: Naïve risk-estimation (based on average ranks) – (source: <https://miau.my-x.hu/mediawiki/index.php/System-modeling> - <https://miau.my-x.hu/bprof/2023/data_for_email_phishing_risk_estimation_3.xlsx> – Sheet: “ranked OAM with naïve solution”)

A naïve aggregation can be realized based on the well-known (yet mostly dangerous) averages where the ranks of the row data in a column should firstly be derived like in the school. The average of the ranking numbers in case of an object (email) can be interpreted in a trivial way: the less is the average ranking value, the higher is the risk of the particular email. This simple methodology can be used independent from the number of objects (emails). On the other hand: a kind of benchmark (threshold) should be defined for deleting an email. This level (rule) is mostly a subjective value: e.g., if the risk potential of an email is higher than the maximum potential value in a database, then this email may be deleted or an alert message should be generated, etc.

Based on this micro-project, it can be understood, what is an OAM-based decision making process. Open question is, how to derive the raw (log) data of Figure#1 in an automated way. This is a new project where new simplifications (models) should be worked out.

Figure#3 presents a relevant problem: How can we “convince” that our naïve aggregation is robust enough? If the objects (rows, emails) are sorted according to the naïve risk potential (average ranks), then pattern (barcodes) could be seen in ideal cases. Here and now, based on randomly chosen OAM-values, there is no sign of a pattern. Observers can eventually see a kind of reddish feeling towards the top and parallel, a greenish feeling towards the bottom of the OAM. Therefore, this aggregation (this kind of risk-derivation/definition) as such is a source of further risks where emails will be stopped although they are correct emails from realm, but new partners and really risky emails can be processed without any alerts.

Conclusion: The above-presented risk-estimation had no positive and/or negative examples for a kind of machine learning (c.f. supervised and unsupervised learning). On the other hand, this knowledge representation is not a totally random approach: the rules/direction about the general characteristics of risks (see: the more is the number of a critical word in an email, the more is the risk) let create a simply (naïve) approach in frame of a fast software development process. The information added value (c.f. cost of a software vs utilities through the software) can be low in case of a naïve solution, but it can also be good enough. It depends on the competitors (other knowledge management solutions).



Figure#3: Sorting of objects – (source: <https://miau.my-x.hu/mediawiki/index.php/System-modeling> - <https://miau.my-x.hu/bprof/2023/data_for_email_phishing_risk_estimation_3.xlsx> – Sheet: “visualization”)

More details about the constant values of 1000 (called Y0) in Figure#2 and about the sheets “online solution” + “comparison” can be read in chapter “Advanced Levels”.

## Password-attacks

A parallel challenge (compared to the email-phishing-problem above) is the password-attack-problem where the dramaturgy of the decision situation can be seen as follows:

* a server log can contain items about login-processes
* a given login-process is an object
* attributes can be phenomena having direction rules[[2]](#footnote-2) (like the shorter is a password, the higher is the risk to be attacked)
* the challenge is trivial: which login seems to be risky enough[[3]](#footnote-3) to be stopped (c.f. fraud detection concerning bank transactions)
* parallel challenges are:
	+ which attribute (level) can be seen as responsible for how many damages? (and/or)
	+ which attribute (level) can be seen as responsible for the binary success (1;0) of an attack?

The two last (so called parallel) challenges need a production function as such (see Figure#4 and more details in the subchapter about the production function - below):



Figure#4: Possible production functions – (source: [https://miau.my-x.hu/bprof/2023/data\_for\_password\_attack\_models%20(5).xlsx](https://miau.my-x.hu/bprof/2023/data_for_password_attack_models%20%285%29.xlsx) – sheet “password-attack-log-OAM”)

In case of an antidiscrimination model, the human expert has to define the direction rules (the-more-the-more or the-more-the-less). In case of a production function, a direction rule is a very complication phenomenon (c.f. ceteris paribus, law of the diminishing returns: <https://www.google.com/search?q=law+of+diminishing+returns>, Liebig-principle: <https://en.wikipedia.org/wiki/Liebig%27s_law_of_the_minimum>).

The signs of the correlation values between X(i) (e.g., length of a password) and Y (e.g., damage) are not totally independent from the direction rules. On the other hand direction rules (e.g., the-more-the-more) can also be more sophisticated e.g., optimum-like. The doubled-attribute-sets are capable of interpreting optimum-like relationships as ceteris paribus views in case of additive models. Regressions need polynomials for the same purpose. Neural networks are capable of deriving ceteris paribus shapes for arbitrary constellations.

# Adaptation projects

After closing the first micro-projects, each Student have to solve seemingly new problems based on the previous experiences with the micro-projects. Each Student has access to the documentations of each micro-projects. Therefore, the motto (learning by doing) is valid not only for the own practical experiences but also concerning the activities of other Students (c.f. learning from the mistakes of others and/or from best practices).

## Ideal Student (prima primissima awards)

This challenge is a classic challenge for supporting the interpretation processes of the antidiscrimination models. Public procurement should also be a good frame, but the mafia-like thinking patterns are (fortunately) totally strange for the most Students.

The dramaturgy concerning the school life should be trivial for Students:

* there are a lot of marks/grades/scores about different exams/courses
* who is the best Student? (c.f. profiling)

More details: <https://miau.my-x.hu/bprof/2023/analogy_v2_v3.xlsx>

## Ideal Company

The challenge about the ideal company (enterprise, partner) has following dramaturgy:

* an IT-expert receives the same salary-value (as offer) from different enterprises
* the IT-expert want to identify the enterprise with the less risk to be attacked

More detail: <https://miau.my-x.hu/bprof/2023/company_ideality_index_formula>

## Ideal weighting

The naïve approaches can identify that the hidden weights are always 1:1:1:…

The question is: Is there a better weighting (without conscious break of the rationality of the competition – like in case of public procurements in general)?

## Involving ChatGPT

About AI may not be spoken from now on (2023) without involving chatGPT into the problem solving processes (c.f. OAMs (objects and/or attributes) should better be identified, and/or OAM-based calculation schemes should better be identified, etc.).

ChatGPT is a robust tool, but it is a kind of risk generator, if somebody does not want to be better than chatGPT. Therefore, a chatGPT interview is such a basis where the Students have to be capable of criticizing each aspects of chatGPT’s reaction…

Demo: <https://miau.my-x.hu/miau/297/Chat-GeePeeTee_vs_Pratchett-G-Pete_nr1.docx>

# Advanced levels

Basic level is the level where the first experiences can be collected without deep mathematical background (and/or in an intuitive/instinctive way). The basic level is mostly not the level being useful enough, but the basic level should be presented in order to demonstrate a realistic benchmark for the advanced approaches:

## Antidiscrimination models

The following documentation offers a kind of entry into the new world of the similarity analyses (c.f. <https://miau.my-x.hu/miau/196/My-X%20Team_A5%20fuzet_EN_jav.pdf>). The antidiscrimination variant (Y0) of the COCO (component-based object comparison for objectivity) approach makes possible to derive an optimized answer for the question: Whether each object (email) can have the same evaluation values (risk-index)? The stair-case function are more complex (sophisticated) than the hidden weighting behind the naïve aggregations (averages) and/or the regression models (as an other well-known but too inflexible) knowledge representation form (e.g. - <https://miau.my-x.hu/bprof/2023/damages_regression.xlsx>). The similarity analyses have an inner quality assurance layer: direct and inverse inputs can be defined and both alternatives have to lead to mirrored results in order to accept a model and/or an estimation for a particular object.

COCO-Y0-models can be seen e.g., here:

* basic level:

<https://miau.my-x.hu/bprof/2023/data_for_email_phishing_risk_estimation_3.xlsx> - sheet “online solution”

* advanced level:

<https://miau.my-x.hu/miau/297/AREA-BASED-ANTI-DISCRIMINATION-MODELS.xlsx> - in several sheets with the constant value of 1000000

Results (estimated risk-index-values and or their ranks) of optimized and naïve models can be different compared to each other. The volumes of the differences make visible how naïve is a simple solution. If the difference is low or there is no difference, then the decision situation is robust. The existence of robust comparisons is the reason why naïve solutions are ever used in the history (see: <https://miau.my-x.hu/bprof/2023/data_for_email_phishing_risk_estimation_3.xlsx> - sheet “comparison”).

## Production functions

Production functions are models where the consequences of different input-constellations should be interpreted. Production functions are patterns. The naïve forms of production functions are manual-driven rules (e.g., if the number of a keyword is higher than a subjective set threshold then an alert should be sent to the user).

More sophisticated production functions generate a simulator where consequences (outputs) for arbitrary input-constellations can be derived (e.g., regression models, decision trees, staircase functions, neural networks, etc.).

Further details about non-causal alternatives and/or about detecting overfittings in case of too flexible knowledge representations: <https://miau.my-x.hu/miau2009/index.php3?x=e0&string=l.v.radi>

The increased flexibility of models leads mostly to a decreasing interpretability. To estimate the rational level of flexibility is the one of the most relevant challenges on the field of AI (c.f. similarity-based models with doubled-attribute-sets are rel. already flexible but still fully interpretable:

COCO-STD-models (incl. examples with doubled-attribute-sets) can be seen e.g., here:

* [https://miau.my-x.hu/bprof/2023/data\_for\_password\_attack\_models%20(5).xlsx](https://miau.my-x.hu/bprof/2023/data_for_password_attack_models%20%285%29.xlsx) – sheet “password-attack-log-OAM(2)”
* <https://miau.my-x.hu/miau/296/joker2_for_biological_systems.docx> + <https://miau.my-x.hu/miau/296/spel_joker2b.xlsx> + <https://miau.my-x.hu/miau/296/spel_joker2.xlsx> + <https://miau.my-x.hu/miau/296/spel_joker2_negative.xlsx>

## Critical aspects of the potential-star-methods

The naïve approaches can lead to infographics (data-visualisation effects) – like the potential-star-method. In this case, parallel existing attribute values are used for generation of a radar-diagram. The area of a polygon should represent the value (risk-level) of the object.

This approach is known since decades, but there is a massive critical aspect: the sorting/ranking the attributes has a robust impact to the calculation of the aera of the polygons!

If a ranking/sorting of the attributes is accepted, then the calculation-formula for the area can also be used in antidiscrimination models instead of the trivial addition (or even multiplication – like in case of biological systems: c.f. <https://miau.my-x.hu/miau/296/spel_joker2_negative.xlsx>).

The AI as such can be seen as a set of different knowledge representation forms. A new formula (especially a geometry-driven formula) is a totally new approach even if this formula is included into a similarity analysis (which is itself a new point of view compared to neural networks, fuzzy-logics, decision trees – because a similarity analysis is partially each of them).

Demo materials: <https://miau.my-x.hu/miau/297/AREA-BASED-ANTI-DISCRIMINATION-MODELS.xlsx>

# Finalising a thesis

The above-mentioned possibilities to create (here and now basically decision-support) models can be seen as a set of LEGO and/or puzzle pieces. This set is available for each problem (for each thesis, application). The LEGO/puzzle pieces can be combined, they are basically context-free.

A thesis can be finalized based on one single model (if it is sophisticated enough e.g., through automation of log-handling and data visualization and hermeneutical modules, etc.) or based on a lot of (in)dependent models.

Writing source codes without a set of models (as test cases) is not impossible but highly not recommended!

# Annexes

# Demonstration material: Magic-of-words



(Translation: automatically in the browser by Google Translate – original language = Hungarian)

*Original text: Fronthatás ugyan nem terheli szervezetünket, azonban a légkörnek melegfronti jellege lesz, mely különösen megviselheti az időseket és a krónikus betegeket, de az egészséges szervezetre is hatással lehet. Az arra érzékenyeknél fejfájás, alvászavar, fáradékonyság jelentkezhet, továbbá fizikai és szellemi teljesítő képességünk is romolhat. Ennek köszönhetően sokaknál figyelmetlenség, koncentrációs zavar jelentkezhet, ami növeli a balesetek előfordulásának kockázatát. A szív-és érrendszeri betegeknek különösen megterhelő ez az időjárási helyzet. Náluk jellemzően vérnyomás-ingadozás, szédülékenység alakulhat ki. Készítette dr. … . Orvosmeteorológus*

The text does not have any aspects where a check mechanism could be activated in order to derive what was correct and what was incorrect in the “declarations” of the “medical meteorology” report. It is even not to clarify: who should be observed at all?

Critical parts in detail:

* “can take part”
* Who does count as “elderly”?
* Who does count as “chronical ill”?
* but “can also”
* Who does count as “sensitive”?
* it “may”
* “headache”🡨when should be said, it is a headache if we do not know, who is affected at all?
* “sleep disturbance” 🡨when should be said, it is a …
* …
* …
* …
* …
* 🡨quasi **each word** is critical concerning a potential log-data collection… (like in case of astrological texts[[4]](#footnote-4) since ever in the history)…
1. [https://miau.my-x.hu/miau2009/index\_tki.php3?\_filterText0=\*knuth](https://miau.my-x.hu/miau2009/index_tki.php3?_filterText0=*knuth) [↑](#footnote-ref-1)
2. The most critical aspect of an OAM-driven AI-approach is the definition of the attributes. Theoretically, attributes can be each potential existing phenomenon because each phenomenon should have direct or indirect connection to each other attribute. The artistic performance of an AI-service can be seen in the definitions of the attributes based on human intuitions. The human-oriented attributes can be interpreted easier (for human beings – c.f. direction-rules in this article). The attributes defined by humans can be therefore an advantage but also a disadvantage (see GO-robots without or hardly existing human interpretation capacities)… [↑](#footnote-ref-2)
3. This interpretation of a challenge is the same like in case of the email-phishing (c.f. antidiscrimination modelling). [↑](#footnote-ref-3)
4. On the other hand: Astrology seems to be the oldest/first activity in the history of data science! It means: based on the geometry of starts and/or planets (= solid data asset), it is possible to involve the same AI as in case of meteorological measurements. The only restriction is: the declarations (like life expectancy and/or existence of hidden cancer syndromes in case of one/more people) should be checkable at once and/or in future in an exact way like in case of meteorological forecasts… What can never be checked, that is a kind of magic-of-words! [↑](#footnote-ref-4)