QuILT 2.0, or the new area of the asynchronous distance education

László Pitlik, László Pitlik (jun) (MY-X team)

The HTML5-based streaming version can be used through this URL: <https://miau.my-x.hu/miau/quilt/2020/quilt2/launching2020III25/part0.html>, where part Nr.1 is a kind of warming up scene about the involved artificial actors/actresses based on the online services of ttsdemo.com

Abstract: The LLL (life-long-learning) is a mostly asynchronous process where the person who has in general ad hoc problems/questions tries to cover the lack of information/knowledge in an incalculable/iterative way through searching, interpreting, discussing, etc. The QuILT 1.0 demonstrated support materials for this complex phenomenon starting from 2019 (<https://miau.my-x.hu/miau2009/index.php3?x=e0&string=quilt>). The asynchronous learning expects that the Students want to increase the level of their sovereignty concerning problem solving. Now, in 2020 – after the prompt declaration of the necessity of distance education because of COVID-regulations, the situation is given to start with QuILT 2.0 where the role of synchronous lectures - having probably only a 5%-share in the knowledge management/cumulation (c.f. learning) processes (<https://miau.my-x.hu/miau/quilt/2020/teaching_is_learning.png>) - can be reduced (quasi to the zero-level) and each other alternatives/treatments may be involved in order to realize more and more learning success. QuILT 2.0 is the unlimited complexity of supports generating sovereignty of Students quasi without verbal/contact components in the communication processes concerning shifting paradigms and increasing the capability of qualitative publishing. The tutorials in frame of QuILT 2.0 simulates/demonstrates real/reality-oriented situations from the strategic level to the deep operationality. Each (written) question of the Students will be answered in an asynchronous way – and these question-answer pairs will build a kind of online FAQ behind each (mostly rel. short) tutorials. Spontaneous (also written) communication can also be realized in form of emails between conductors and Students. The QuILT 2.0 system works with artificial players generated by ttsdemo.com. The conductors/teachers become a kind of directors of content/speed/visualization like in case of a looper-driven one-man-show in the music industry.

Keywords: robot teacher, robot journalist, advanced Turing test, investigative journalism, automation, efficiency, big data, solver-based modelling/interpreting/hermeneutics, intuition generating process, artificial intelligence, OAM

# Prologue

Publishing as such seems to be an artistic performance where the personality of the author(s) should play a relevant role from point of view of the Readers. In the field of the scientific publications, the AWS (academic writing skills) are accepted as a frame for quality assurance. The AWS logic is a product of the magic of words. It is not a part of the KNUTH’s universe (where knowledge/science is what can be transformed/transferred/translated into source code).

This paper presents how and why it is possible to create a robot journalist who is capable of executing a kind of investigative journalism process in a way where the Turing-test should be positive. The Turing-test (<https://en.wikipedia.org/wiki/Turing_test>) is responsible for detecting robots with a low competence level. An advanced Turing-test could also be created where the human beings are compared with robot performances (e.g. on the field of the investigative journalism and/or on the field of the intuition generating processes). The advanced Turing-test means: Students may not have a lower level than robots because the existence of robots assumes that the components of the journalism (publishing as such) and connection between them could be identified and simplified for source code writing.

# Pre-history of robotized publishing

The following articles/documents are worth reading in order to see the complexity and the simplicity in a parallel way:

* <https://miau.my-x.hu/miau/258/kome_v1.docx> - about robot writing
* <https://miau.my-x.hu/miau/196/Pitlik_Robotlektor_Roll-up_85x203cm_HU_EN2.pdf> & <http://miau.my-x.hu/miau/181/etdk_2013_v4.doc> - about robotized quality assurance (robot lector)
* <http://miau.my-x.hu/miau/208/idealis_lektor_v1.doc> - about the activity patterns of a lector
* <https://miau.my-x.hu/myx-free/index.php3?x=test1> – Academic Writing Skills – from an other point of view
* <https://miau.my-x.hu/myx-free/nos/> & <https://miau.my-x.hu/myx-free/index_fifawc2010_en.php3?x=soccer_news_en>- about the possibility of a robotized sport-journal
* <http://miau.my-x.hu/miau/207/tdk_letra_v2.doc> & <http://miau.my-x.hu/miau/222/tdk_letra_v7.xls> - about how to create a really good best student paper
* <http://miau.my-x.hu/miau/224/jo_fogalma_otdk_biralat_anonimizalt_2.docx> & <http://miau.my-x.hu/miau/200/otdk_v2.doc> & <http://miau.my-x.hu/miau/200/otdk_v1.doc> - about the anomalies in the competition of best student papers, etc.

# Robot-Teachers/Conductors vs. Robot-Journalists

The Quilt-based (version 2.0) asynchronous distance education define artificial players as layers of the personality of a “schizoid” (multi-layered) conductor who should be able to demonstrate more point of views concerning teaching/learning methodologies and also involve virtual Students into the discussion processes (<https://miau.my-x.hu/miau/quilt/2020/quilt2/lauching2020III25/part0.html>) about KNUTH-relevant competences (like robotized publishing).

What the Part Nr1 of the new QuILT-services can offer is (see: <https://miau.my-x.hu/miau/quilt/2020/quilt2/lauching2020III25/part1.html>) which relevant steps/components can be identified in the process of intuition generating and publishing.

The robot conductor (Con-Duck-Thor) supported by virtual Students (Stew&Dent) helps for real Students to understand (step by step) the elementary components of publishing starting from one single keyword to the finalized article. Each step is that clear and clarified that the development of real source code is quasi trivial. What means: a robot journalist could be really existing based on a big-data-concept and solver-oriented analytical steps for interpretation of the pictures derived from the raw data assets.

What is not acceptable as human performance are the following publishing strategies:

* creating a stream of subjective opinions, (unproved conspiration) theories about WHY-syndromes (e.g. why is a corona virus exist?) by the author
* answering trivial questions (like how values of a phenomenon are changed where the answers can be derived e.g. in a spreadsheet-environment within seconds)
* creating questionnaires with attitude-question (like how import is for you to know about IT-solutions for the distance education? 1<10) and/or organising interviews in order to write arbitrary sentences,
* using questions like “how many **percent** can be realized in a knowledge acquiring process based only on lectures” – without declaring what the real unit and/or maximum of the expected “**percent**” is?
* etc.

What is complex enough to create an article about it are following approaches (c.f. previous list above):

* deriving/proving theories/hypotheses based on big-data (instead of conspiration theories)
* creating simulation systems to see connections between inputs and outputs (instead of why-oriented ideas)
* development of log-systems instead of self-“evaluation” processes
* creating artificial index-values (e.g. safety of the medical system of a country) in order to substitute entirely subjective abstraction (like subjective satisfaction concerning the medical system of a country), etc.

# Can complexity be demotivating?

Complexity can be demotivating if the complexity can not be reduced and described in a way where the KNUTH’s universe is already existing. It means an article needs a lot of thinking experiments where the arch-stone and each other stone-part of the arch can be set to the right position on a totally empty/blank sheet of paper. A thinking experiment should always be as details as possible – using each components of the system of publishing. The following QuILT-performance tries to support the understanding of the logic of a particular thinking experiment with all their components, details and connections.

# How to use the QuILT 2.0 system

If you see the first three parts (Part Nr.0 and Part Nr.1a and Part Nr.1b)

* <https://miau.my-x.hu/miau/quilt/2020/quilt2/lauching2020III25/part0.html>
* <https://miau.my-x.hu/miau/quilt/2020/quilt2/lauching2020III25/part1a.html>
* <https://miau.my-x.hu/miau/quilt/2020/quilt2/lauching2020III25/part1b.html>

then you will have a lot of critical impressions like

* why do we need artificial players at all?
* why not only texts?
* why it is that slow?
* why it is not cut?
* etc.

There are answers to these potential questions like

* it is important to impersonalize the knowledge, to make it independent from real persons because the further LLL activities will also bring contacts to strange sources and it is important to be capable of evaluating information independent from the person having it / parallel: it is also important to have the chance to demonstrate different thinking patterns (see both conductor-layers and Student’s point of views)
* texts are in general good learning materials – but the experiences show semester by semester that the motivation level for reading is relatively low – on the other hand: the voice-and-video-support makes possible to change exposures – it means: reading can become as a monotonous challenge and hearing can maybe solve this distress situation
* the QuILT 2.0 system works without any technical and financial background, there this version seems to have a good price-performance ratio / on the other hand, this technology can integrate everybody (both Students and Teachers) where it is more important to be capable of thinking experiments than to be able to produce in Hollywood-quality
* the relatively slow process (because of avoiding the improving force fields of cutting-effects) makes possible to have own opinion/notes and this effect generate the needed sovereignty
* …

The mp4-files can be converted to mp3-files and they can be used e.g. during travelling. The rel. small mp4 files can be reached in a direct way by using the shortened URL: <https://miau.my-x.hu/quilt2> where each relevant fragments can be identified – among them the part\*.mp4-series. These mp4 files can be also used in smart phones.

It is also important to know: the online services of ttsdemo.com are limited: e.g. 600 characters can be prepared for one speech-activity. The QuILT 2.0 based on the presentation possibilities of ttsdemo.com

The detailed information will be presented in muted form in order to have the possibility to teach others through involving the short streams with specific mouse movements (highlighting important details) and also in order to enforce memory exercises based on the presented interpretations before by the artificial players.

The parts make possible to discuss about their content. This discussion will also be in an asynchronous form – it means each question and answer will be moderated before and the FAQ-like results will be readable in the MIAU WIKI system:

* <https://miau.my-x.hu/mediawiki/index.php/QuILT2_part0>
* <https://miau.my-x.hu/mediawiki/index.php/QuILT2_part1a>
* <https://miau.my-x.hu/mediawiki/index.php/QuILT2_part1b>, etc.

Each impulse should be communicated based on the always open email-channels.

Each part can be extended, and a lot of new details can be supported through new mp4-solutions. The FAQ-oriented support is faster but specific question will surely need captured streams too in future.

Potential offers for final publications:

* the QuILT 2.0 system can be criticized and/or
* improved e.g. in form of thinking experiments
* (c.f. <https://miau.my-x.hu/miau/quilt/2020/sovereignty-pla.docx> with recommendation for education-oriented topics and their data assets)

# Story board: The case of H1N1!

From now on each text part written for THOR will have a turquoise background colour, and the same logic is valid for each other player (virtual actor/actress): DENT will have the colour-code of magenta, STEW will be grey highlighted. CON’s colour will be the red and DUCK “can not have an other” colour as yellow. The colour-scheme is quasi a randomized one.

The whole story with generated voices can be streamed here: <https://miau.my-x.hu/miau/quilt/2020/quilt2/lauching2020III25/part1.html>. It is worth reading this article before switching to the animated version:

|  |  |
| --- | --- |
| Persons | Messages of PART1a - … how to produce an OAM behind an article… |
| THOR | In medias res: The corona virus is not the first virus with relevant impacts in the modern history of the human beings. On the other hand, good structured data assets about the recent corona virus pandemic process are hardly to identify. The quasi closed case of H1N1 virus concerning the years 2009-2010 can be used to model what should we do for the data asset management in case of the corona virus in order to be capable of deriving conclusions based on big-data-sources and based on solver-oriented analytical techniques? |
| DENT | Stop! Please! Stop-stop-stop! This was a really rough start, Professor. Is my interpretation correct, if I say: the keyword we have is the abbreviation of H1N1? If somebody read the previous pages as recommended, it seems to be trivial, that the first step should be an easy one. It means the robot journalist should also be capable to have a clear starting parameter. And this can only be one single keyword because the next step should be a searching process… |
| THOR | Excellent! You are right in each aspect. The robot journalist needs at least one single keyword to start with searching. For searching, the robot journalist will use for example the Google-Search-Services as almost everybody would do. It is not necessary to have any pre-knowledge about the background layers of the keyword. It is also not necessary to have a precise goal before starting with search activities. It means the results of the searching needs no human-like interpretation processes because we do not really know what our brain really makes with these results. |
| STEW | Let us summarize briefly the situation. Do we need just one single keyword without any pre-defined purpose? And if we do not have more than a keyword how will we go on if we will have the searching results? I have already read the recommended documentation and there were a lot of other keywords like big-data. Is my assumption proper if I think the robot journalist will prefer searching results where data are available? But data can be present in text stream and also in more structured form – for example in matrices. |
| THOR | Yes, your interpretations are correct! Robots are relatively simple creatures. If a robot journalist can identify structured data where the legends or to be more precise: the row header positions and or the column header positions consist the searched keyword of H1N1 then these results will be used for writing an article. Namely, the header positions and other words about affected phenomena have mostly a clear relationship to the starting keyword of H1N1. So simple is the way therefore to the OAM (you know: to the object attribute matrix). |

|  |  |
| --- | --- |
| Persons | Messages of PART1b - ...efficiency problems with PDF-based communication... |
| CON | It is nice to follow your conversation. But, please, do never forget the Knuth’s principle: Knowledge is what can be transformed into source code. Concerning to your previous sentences, I think, we need a lot of short presentations about the potential results of the different searching processes. We should see, what are the results, if we really have one single keyword? We should also see how we could improve the ratio of more structured data among the results. What do you think? |
| DENT | I agree - definitely. It is trivial that we will have PDF-files among the results if we do not use specific extensions to the one single keyword we have. These PDF documents can contain Tables. But – you know – PDF-files can be converted to DOC-files quite simple – however it is never guaranteed that the numbers of the tables keep their structures. Nobody likes retyping too much therefore it is important for increasing the efficiency of the data collecting processes that we use sources where structures will be kept. |
| CON | At first, we need a test process with the only keyword of H1N1. As you can see, the first results belong to Wikipedia in different languages. Wikipedia is not a really good source for structured data. Wikipedia delivers mostly texts. But the HTML-sites make possible to create tables. These tables can be transferred into Excel. But the footnotes and the format of the numbers make the data processing not really easy. Parallel to tables, there can also be charts, figures, maps. Unfortunately, they can not be processed with Excel in a direct way. Therefore, Wikipedia is not a preferable source. |
| XXXX | Please, watch the next video too, where the described phenomena can be detected… |
| DENT | Thank you for the presentation. It is now clear what kind of sources, formats we can not involve into the data collecting efficient enough. I think there are further traps not supporting our goals. For example, it would be nice to see a demonstration about the problems with PDF-converting. I am afraid not only table but also parts of text are integrated into the raw PDF-file in a way, that we do not have any chance to use it – just in form of retyping the relevant information units. Could you show us a presentation? |
| XXXX | Please, watch the next video too, where the described phenomena can be detected… |
| STEW | Thank you for this presentation too. It is not even motivating to see, that a lot of anomalies can really be detected behind the common formats like PDF. We could see the possibility of the converting, but the converted content can not be used in an efficient way. Parts of texts are pictures and vice versa – parts of tables could be converted. Obviously, nobody cares about efficient data processing – although the PDF-s as such are voluminous. I think, I do understand here and now what the difference is between data-driven thinking and the arbitrary opinions. |
| DUCK | It is a great pleasure for me to be a witness of the development! Now, it seems to be it is ultimately time to speak about the basics of the similarity analyses. You could see before, for example in Wikipedia that the statistical values belonged to countries. Therefore, the robot journalist can declare, that the OAM will contain as row header countries. Countries could also be identified in the corrupt converted PDF. The Wikipedia tables and PDF-tables contains potential attributes too: for example, data about vaccines. Attribute names will be set as column headers into the OAM. |

The collected data and their sources can be studied in the background XLS too: <https://miau.my-x.hu/miau/quilt/2020/projekt_h1n1/OAM1_h1n1.xlsx>

|  |  |
| --- | --- |
| Persons | Messages of PART1c - …the value of knowledge or how to use already downloaded data… |
| DENT | Okay, we know by now a lot about the negative cases – it means about sources where we will have nothing else just troubles. Fortunately, the robot journalist can also know that the basic search activity with one single keyword can deliver information about institution/organization like Wikipedia, WHO, etc. These information units can be identified in the domain or server names. Parallel, robot journalist can know that specific additional keywords can increase the chance to find tables, matrices. These extensions are for example: statistics, figure, table, etc. |
| DUCK | That is right! We can check immediately this hypothesis. You can see the impact of additional keywords in the searching phrase at once. It is important to know it is worth searching not only with English keywords! Let alone, it is very important to define the objects – in our case the countries and maybe a continent like the EU in order to increase to probability to have good results in the first page of Google. As it can be seen, the enterprise Statista is a provider where you can find appropriate information. Even the highlighted images belong to Statista.  A képen képernyőkép látható  Automatikusan generált leírás A képen képernyőkép, monitor, számítógép, laptop látható  Automatikusan generált leírás |
| DUCK | It is also relevant, that the results of the Google Search can also be translated by Google as in case of each other Internet site. However, the identified charts, figures can consist description elements for example country names in our case in German where the Google Translator does not have any access rights because these information units behave themselves like images behind the PDF file before. In ideal case, the robot journalist will have timestamp-like data too. It means besides country names and statistics it is also important to know about the time aspects.  A képen képernyőkép, számítógép, monitor, laptop látható  Automatikusan generált leírás A képen képernyőkép, monitor látható  Automatikusan generált leírás |
| DENT | But how can we copy and or download the identified data in text format? Figures, charts, animations are images, and these can only be processed through retyping the needed values and descriptors like country names, timestamps. Stew! You are our expert for ethical hacker tricks and tips. Do you know what have we to do in such a case? I think, the robot journalist and therefore we all need some idea to ensure efficiency in the data processing. Stew! It is now – your - turn! |
| STEW | Aye-aye, Sir! Sir, I can immediately present you not only one but two parallel possibilities. There are two spells for such kind of HTML-oriented problems. The one spell is: CTRL+U, and the other one is F12. Of course, spells should be trained like in Hogwarts even for Harry Potter. In these cases, training means you should know about one specific HTML-tag-pair. These tags are the opening and closing tags for HTML-tables. By the way: these tags can also be used in searching phrases. This is a technique-oriented support parallel to the keywords before about statistics, tables, etc.  A képen képernyőkép látható  Automatikusan generált leírás |
| DUCK | Excellent! And one further spell - if we want to use such kind of analogies – is the hotkey CTRL+F. This button combination makes possible to search within the source codes. Namely, the spells of Stew lead us directly into the deepness of the world of HTML-tags and Java-Scripts. The first views demonstrate the codes after the command of CTRL+U. With CTRL+F it is possible to identify the opening tag and the needed data can be read at once. For transferring the data into Excel, it is necessary to select each character including the closing HTML-tag for tables.  A képen képernyőkép, képernyő látható  Automatikusan generált leírás A képen képernyőkép látható  Automatikusan generált leírás |
| DENT | I have tried the other spell – it means the F12 for opening the development tool behind the browser of Google Chrome. As everybody can immediately see, the opening tag can be identified and after selecting of the whole – very-very-long – row, and we can check the value for Germany. The number is 17265. This is the same value as before in image-view and also in case of CTRL+U. Fortunately, the Excel can transform the HTML-codes, and we and the robot journalist will have the needed data in a useful format.  A képen képernyőkép látható  Automatikusan generált leírás A képen képernyőkép látható  Automatikusan generált leírás |
| CON | Finally, let us declare something relevant one: All the spells are not important if somebody can pay for information brokering services for experts and or enterprises. Therefore, knowledge is a kind of power or even a kind of business potential. Information brokers are experts who are able to collect data in an ethical way as soon as possible. Common citizens have the possibility to choose experts and or online services – but they should mostly pay for the information. The business is simple: experts are so fast that the salary can be paid by the common citizens at any rate. |

|  |  |
| --- | --- |
| Persons | Messages of PART1d - …finalizing an OAM… |
| DUCK | The OAM is a key for automated intuition generating. It means to interpret the world around us in an automated way. The robot journalist and so, all of you have each necessary component for building an OAM. We have the searching possibilities and tools for their fine tuning. We are capable identifying objects and attributes. We also have the ethical tricks to identify and copy HTML-tables letting execute an efficient data processing. I should repeat my first sentence: The OAM is a key for automated intuition generating. It means to interpret the world around us in an automated way. |
| THOR | In case H1N1 we could download spreadsheet-compatible data about infection and death rates of countries in the EU. We have the population-density, and the level of urbanization. It means the ratio of people living in cities compared to the total population number. We could download a PDF with data about vaccination of different groups of the society like clinical risk groups, pregnant women, health care workers, staff of long-term care facilities, residents of long-term care facilities, and older population groups. Finally, we have the time as a strong structure to say what is X and what is Y. |
| DENT | This is a very informative, strong-structured figure. But I have a lot of question: What you mean if you speak about X and Y? Why is time that important? How we can derive the so-called directions? Of course, I will immediately try to approximate the parameters of ideal Students who has not only questions but parallel the potential answers concerning them are also given. A good question defines at once its potential answers. The most trivial case is a question with two answer options: yes and no. Each question can always be transformed into a lot of yes-no-questions. For example: |
| DENT | I could have even asked: Does X stand for the independent variables? Answers: yes and no. Because, I think attributes with the X-sign should be the so-called independent variables or causes in a causal interpretation frame. The attribute marked with a Y-sign is therefore the dependent variable or with other word the consequence-variable. Time should be important to be capable of interpreting causes and consequences. Causes, reasons should happen earlier then consequences. |
| DENT | And finally, directions should be the potential types of the relationships between an X and the Y. The most simple types are proportion-oriented types: the-more-the-more and the-more-the-less. More precisely: the more is an X-variable the more is the Y variable. In case of H1N1: the more is the rate of infection (where infections happen before deaths) the more should be the rate of deaths caused by the viruses. However, relations between an X and the Y can be more complex. We can speak about optima where neither the lowest nor the highest value of X make possible to increase the values of Y. |
| THOR | Exactly! So, one single row of an OAM is a sentence or a rule or a declaration - assumed that each statistical value is correct enough. If the values of the X-variables in case of Belgium are that low and or that high as they even are in the database, then the Y-variable should exactly have the collected value. If we have a lot of objects – in case of the H1N1-project a lot of countries, then we have our database for simulations. Unfortunately, we do not know in general which relations are given between the variables. Fortunately, the robot journalist has an automated tool for this challenge! |

|  |  |
| --- | --- |
| DENT | But before we try to be introduced into the world of the robot analysts, it seems to be important how we can build an OAM based on the huge number of data-puzzle-pieces? I think it can not be the only way creating references in Excel - position for position - between the raw data and the OAM-cells. I expect if we have structured raw data then we should have the chance to integrate them to a standard structure for report-generation. Stew! Are you still awake? Do you have some idea, some new spells or even wizard-services? |
| STEW | Oh-yes, I am awake –of course- and yes, I do have the super tool! And yes, this is a wizarding surface. It is namely the pivot table generation. In order to use the Excel-pivot-tables, it is necessary, that the needed data can be integrated into a standard object-oriented structure. This structure makes possible to derive an OAM through the pivot wizard without doing anything with the data positions in a manual way. Of course, it is also possible, that the references to the row data positions will be integrated into the OAM-structure in a direct way. But the raw data may never be retyped! |

|  |  |
| --- | --- |
| DENT | As I always do, I am trying to reproduce your idea, your instructions. It seems to be quite simple, mate! Even objects like Croatia because of lacking data can be excluded with a mouse-click from the OAM. The attributes can be placed in the expected position if you have a X-signs numbered. On the other hand, a kind of final checking is always necessary. For example, the maximum and minimum values for each attribute should be derived and interpreted. For an interpretation are definitions of the variables needed. In an ideal case, the ratio of the maximum and minimum values is limited. |
| XXXX | Please, watch the next video too, where the described phenomena can be detected… |
| CON | The previous video highlighted a lot of important steps of the quality assurance. For example: if we do not need the digits after the integer values, then these number should not be visualized. If we need a report, it is always to check, how many data unit are behind a cell. Here and now we may have only one raw data in every cell. If total and or subtotals are not necessary, they may not be visible. Maximum and minimum values should be checked. In our case, the density is realistic. The infection rate is critical. In general, for comparing we need relativized data instead of absolute numbers. |
| DENT | All right. Yet, I have new questions: What do you mean with the evaluation of critical in case of the infection rate? Besides, what kind of attributes got made invisible through the group of “no”? I could translate the German header-positions in this group, and I think, the no-group is the group of the attributes with the risky absolute values. The absolute values should however be collected to be capable of calculating the relative values – for example the number of infected persons divided by the number of the total population. Is this interpretation correct? |
| CON | Your interpretation about the relative and absolute values is correct. Good job! The data positions about the infection rate are critical because the maximum and the minimum values demonstrate a difference which can not be explained at once. It is important to know, that the number of infected persons and the number of the deaths are independent from each other in the logic of the statistics – unfortunately. Now, we come to a point what seems to be that important that we need a little bit more to talk about. Mr. Thor, what may be declared to this critical issue? |
| THOR | Well, the number of the infected person is never a good number. Only few people are tested in general. So, the real volume of infections can never be measured. The number of deaths is also a fuzzy number. The real causes for death can probably be derived through autopsy. But a person executing a bungie jump can have a heart attack although the rope will be cut by a killer so that the jumper will never know about this action. In this morbid case, the cause of the death will be the crash and not the heart problem. |
| DUCK | Now, short about a relevant part of our history: Winston Churchill once reportedly said that “the only statistics you can trust are the ones you have falsified yourself”. With other words: “I only believe in statistics that I doctored myself”. They mean - for good interpretations we always need exact definition about our attributes. Parallel, in an artistic performance, Vonnegut designed a table about the most important persons of the second world war – about Churchill, Hitler, Roosvelt, Il Duce, Stalin and Tojo. The next video presents a core message: |

|  |  |
| --- | --- |
| Persons | Messages of PART2a - …working with an OAM… |
| THOR | Well, we have now an OAM. This matrix is the key for the publication. The robot journalist does not have real competences in the field of statistics, epidemiology, etc. Yet, it was possible to identify relevant phenomena – it means attributes and objects – and data sources letting use them in an efficient way. The data could be structured at least in a manual way – but in optimal case based on the pivot wizard in Excel. The robot journalist is interested to derive questions and answers never existing before and being extreme for news. The OAM is the answer for each lack of human competences. |
| THOR | The OAM is a system – in general the OAM is a causal system. The independent variables are the causes for each change – it means for the consequences - concerning the dependent variable. Therefore, the potential questions were existing before the OAM with H1N1-relevant content could be filled. The OAM is a kind of closed universe supporting all human purposes concerning the attributes and objects of this universe. This universe is a system and therefore system-theoretically it means context free knowledge can be used to interpret it. |
| THOR | And before somebody ask me to stop with this kind of high-levelled magic of words and to present at last examples, I change my strategy on my own and I will demonstrate some general questions. These general questions are: Which correlation level can be reached based on flexible online robot analyst or with other words: based on online analytical tools? Correlation is an index value being capable of describing the similarity between the consequences and their estimations based on a model using the independent variables. |
| THOR | And you want surely and immediately to ask me that I have to use from now on just a H1N1-relevant terminology. Well, the question before can be translated into a context relevant form as follows: How close can a flexible model estimate the values of the death rate of countries based on all collected variable like vaccination rate of different groups, population density, grade of urbanization, infection rate? Your next remark should be at once: Why it is important to know whether the collected data are capable of a close enough estimating of the facts or not? |
| THOR | Well, the human and animal brains and probably the plants and one-cell-systems too, they do the same interpretation task since ever. And they do this based on the strong force fields of the evolution and selection step by step better and better. It seems to be at least so. Back to the focused H1N1 terminology and parallel to the recent situation determined through the corona virus pandemic processes, we need for rational, sustainable actions interpretations, models, simulation possibilities for the whole society and parallel for our own decisions. |
| THOR | Therefore, a model being robust enough based on the available data, leads us to further questions. Before they will be listed, it is necessary to highlight that this project about H1N1 can not use a real big-data background, but the modelling process is quite the same independent from the number of data units. The next very-very relevant aquestion should be: when may be seen model as good enough or when is a model better than an other one? This learning material is not defined for answering this probably most interesting question. But the following URL is maybe a good starting point: <https://miau.my-x.hu/miau/196/My-X%20Team_A5%20fuzet_EN_jav.pdf> |

|  |  |
| --- | --- |
| THOR | The next interpretation level in case of a good model is, what kind of objects – here and now countries – can be estimated close enough and which produce extreme errors or differences between the fact about the death rate and the estimations about the same phenomenon? As you could see and hear, we have now the first extremity what our robot journalist should handle with. This is therefore a kind of thinking experiment which can be played, executed without real data. |
| DUCK | I think the next analogy could be helpful: If you tend to make a survey and design a questionnaire later then you should be capable of interpreting the expected data without seeing them. It means you must be capable of generating randomized answers concerning your questions and also capable of analysing them. Why? Because the analysing of the real data will inspire you to read interpretation into them. A consequent, consistent, existing rule system for analysing arbitrary randomized or real answers is protected from this risk. Ein system created before can not be corrupted through fresh data. |
| DENT | I think, you know by now that we need again examples. It is clear that the robot journalist got programmed based on the system-theoretical backgrounds. It is also clear that we human beings needs a kind of quite universal frame for our thinking experiments. But here and now we can see a classic antagonism. We are still not matured enough to make thinking experiments in an unlimited complex level, and parallel, in order to become more and more matured, we need the theoretical complexity. This is a power field to form our personality – if we “fight” with the challenges. |

|  |  |
| --- | --- |
| Persons | Messages of PART2b - …the first models… |
| THOR | All right, it is time to create our first model and parallel, we will also see the online analytical tool. You could believe me – this tool is “toll” if somebody learned German! Worth knowing: The Excel Solver is a useful add-on and the Solver is being capable of substituting the online tool - in a limited volume. On the other hand, Google spreadsheets can also involve online solver engines (like the NEOS server). Summa summarum: if you have a raw OAM and you also have a connection to a solver-based service – for micro problems it is irrelevant whether online or offline – then you can go on. |
| THOR | Fortunately, the usage of online services does not need any mathematical pre-knowledge. Online analytical engines could even be installed in smart phones because the only input they use is a ranked OAM. The raw OAM contains values with their identified or calculated units like population density in capita per square kilometre. A ranked OAM contains however integer values without any units where in case of each attribute only ranking values are stored in the OAM. Ranking values mean: the object having the highest or lowest value can have the ranking value Nr.1. |
| THOR | If you still remember, Dent spoke already about the types of relationships between independent and dependent variables. The-more-the-more approach can be re-defined for the ranking of the values of one single variable as follows: the highest value of a variable will have the ranking value of Nr.1. And vice versa: A the-less-the-more approach means – the lowest value of a variable will have the ranking value Nr.1. Ranking values have seemingly a reduced information value because the differences between the ranking levels will not be processed. |
| CON | May I interrupt you? On the other hand, ranked values are kind of standardized values with a lot of advantages. These advantages are for example: it is not necessary to use delimiters like dots or commas and digits after the integer values. Therefore, the X-part of the OAM can always be checked in a trivial way: the minimum value is always 1 – it means Nr.1 in case of each X-column. The maximum value of each X-column is the number of the objects. This logic makes the transferring of OAMs relatively easy and the volume of this data communication is as compact as possible. |
| DENT | If I understood you correct, then the ranking values and all advantages will only concern the so-called X or independent variables? Assumed an answer of YES, let us make a short thinking experiment: This standardization ensures therefore a kind of context free situation in case of each OAM. The same ranked OAM can namely be calculated or derived based on a lot of raw X-values. It seems to be for me as a new impulse why the robot journalist and the robot analyst do not need the context-interpretations human beings execute. Is this derivation or this logical way correct? |
| THOR | Excellent! You identified here and now one of the core characteristics concerning the robotizing. I should highlight again and again: the way to an OAM did not need any contextual interpretations as human beings do. With other words: the variables in an OAM will validate each other because these attributes are available online for the same object-set. And if the building of an OAM can be automated without human-like context interpretations then the using of an OAM should also be automated, and also without human-like context interpretations. |
| DUCK | The well-known analogue situation is the automated generation of the annual final reports of enterprises. These reports contain an OAM with 2 objects: previous year and current year. The number of the attributes is balance-dependent and quasi unlimited. The robots produce texts based on the values of the OAM. The logic of interpretations is pre-defined. For example: the profit is higher in the current year compared to the previous year. The increasing will also have percentual values. Such kind of reports can be downloaded in Hungary for each Internet user without any restrictions. |
| DENT | I think, this is clear enough. But we should also clarify what we have to do with the Y-values of the OAM? It seems to be trivial based on the previous declarations that the Y-values will have their single or common unit or with other words: they are quasi raw values even if they can be derived through calculations. In case of the H1N1-project: instead of the raw X-values of the calculated relative infection rates, we will have ranking values in the OAM, but the parallel calculated Y-values of the death rates will have the calculated unit – for example capita vs. 1000000 population. |
| THOR | I can confirm your interpretation about the raw or calculated Y-values with their own single unit. And parallel, this is the answer to the limitation of the context free characteristics. The ranked X-part of an OAM can belong to quasi unlimited real situations or contexts, but the Y-values are quasi the guarantee for the needed uniqueness. The Y-value can be seen as a kind of fingerprint. And now, the next presentation will demonstrate the first modelling steps in frame of the online analytical tool. This tool can be accessed with the following shortened URL:  <https://miau.my-x.hu/cocostd> |
| THOR | You know, in this phase of the co-operation, we will only use muted presentations in order to catalyse your personal interpretations. In advance, I have to say, that we will see the raw OAM with units and also the ranked OAM without units for the X-part and with a unit for the Y-column. The X-values are enforced integer values, and the Y-values are integer too but not ranked. We will see 27 objects – it means 27 countries of the EU. And we will also see the 9 attributes: 6 attributes about vaccinations and the population density, the grade of urbanisation, and the infection rate. |
| XXXX | Please, watch the next video too, where the described phenomena can be detected… |
| STEW | All right! All the theoretical aspects and the previous and recent presentations by you before about the X and Y signs and direction and by me about pivot tables can be matched. Yet, I have the question: Why did we need that long theoretical information if you said: we do not need any mathematical pre-knowledge for the usage of the robot engines? As a common citizen, I think, an application is then trivial, if we need hardly anything to use it. And we could even see: if we have an input matrix, then we can select it, copy it, and the results are immediately present. |
| THOR | Well, this is a new antagonism what you are focusing on here and now. Dent have already spoken about one antagonism before - concerning a kind of circular reference problem with becoming more matured and being more matured. Your remark let highlight one of the most important problem of the didactic: if I did not speak about something before, then somebody will me ask later to the topic - as a sign: you should have spoken about it. On the other hand, if I speak about something - let me say in a direct way - in advance, then your will be right to ask, why should we handle something that early?! |
| STEW | Okay! Then even back to the last presentation. We speak almost continuously about automation and robotizing. And the presentations are made with manual controls. I think it could be important for all of us to highlight, that even pivot table can be derived with macros. And macros are the source codes for Excel. Parallel, the usage of an online service can also be initiated based on a C-URL call where we can have parameters for the input matrix and of course for the needed online service. A C-URL-call can deliver as result the same view as before or even only that part what we really need. |
| THOR | Thank you for the relevant extension of the common horizon. The next presentation demonstrates even a C-URL call in order to deliver only estimation values for the starting OAM. We will see how the correlation will be calculated and how simple it is to visualize country-specific colouring effects based on the differences between facts and estimations. The realized correlation value based on our manual-driven directions for each attribute led to a relatively low level - although the directions can be seen as correct based on the literature. |
| XXXX | Please, watch the next video too, where the described phenomena can be detected… |
| DENT | I think we do not have any other chances concerning the definition of our directions. It seems to be relevant to say loud out all directions and the arguments behind them: The more is the rate of the vaccination in case of each risk group, the less should be the death rate - because the vaccination makes possible: to increase the robustness of our immune system. And parallel, the more is the population density, and the more is the grade of urbanization, the more should be the death rate because of increased infection risks. And finally: |
| DENT | And finally: the more is the infection rate, to more will be the death rate compared to the total population. It is therefore not really clear, why we do not have a higher correlation level – it means a better model? With other words: do we have any chance to visualize what kind of relationships or even directions between each X and the Y will lead us to a better model? I know that we have already declared that our collected data are probably not that correct. From an other point of view: when may we say: we have wrong data? |
| CON | Excellent interpretations and questions! It is always a great pleasure for me to see and hear that the quality and or risk management can also be important for the new generations. All these questions will be answered in the next presentations where you will see a simple and a general trick how we can enforce new relationships between each X and the Y based on the same online analytical engine. The simple trick is to use doubled directions. The general approach is exploring through the online engine itself, what kind of relationships could be existing at all. |
| XXXX | Please, watch the next video too, where the described phenomena can be detected… |
| XXXX | Please, watch the next video too, where the described phenomena can be detected… |
| THOR | Let us summarize the core messages of the presentations: you can double the number of the attributes with your logically derived – it means direct – directions and with the so-called inverse direction. This kind of doubled setting can be seen as a kind of question: What is the stronger direction in case of an attribute? The general trick works without directions… |

# Potential FAQ-elements

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Part | Player | Content | Question | Answer |
| 1d | STEW | Highlighted images | What is a correct number of the rows in case of 28 counties and 6+4+1+2 attributes? | 28 multiplied by 13 because each country should have a cell, a value-position for each attribute |
|  |  |  |  |  |