The future of pharmacovigilance with the use of artificial intelligence sounds good

Marwen Bouaziz, Data Scientist
Isabelle Laugel, Principal Consultant

Life Sciences Expertise

imaginalsoce for your data

www.LifeSciencesExpertise.com

Introduction

The future of pharmacovigilance with the use of artificial intelligence sounds good! Indeed, the use of artificial intelligence (or AI) to facilitate the treatment of pharmacovigilance cases is much potential. How can we merge artificial intelligence and pharmacovigilance? In this document, we will discuss the opportunities that AI represents for PV data, and also make sure we stay aware of the challenges involved in implementing this technology.

This document will first introduce the basics of AI, then show how it can be used in case processing, and finally we will see how various Drug Safety activities can benefit from AI.

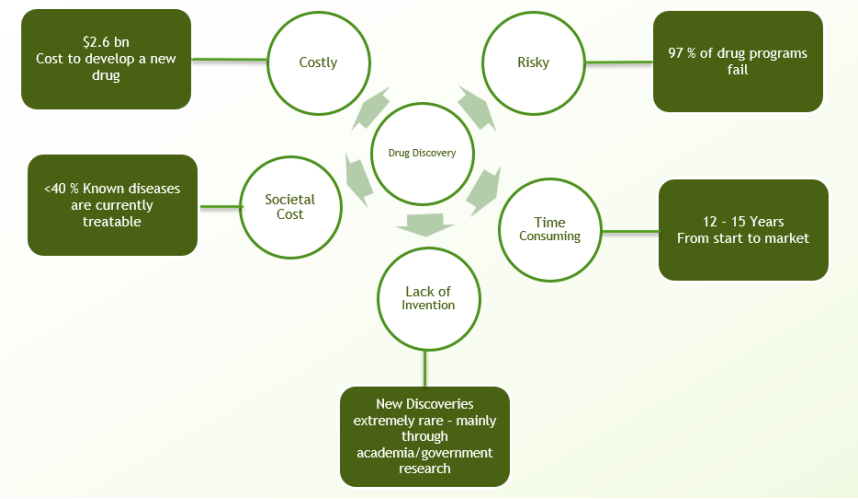
However, whenever we try to introduce new technologies in regulated activities, we always worry about the same major questions:

- Are we still compliant with regulations?
- Is the work required worth the investment?

We will try to answer the following key question: “Is it really safe to transfer Pharmacovigilance responsibilities to software that claim being able to think, even though we all suffer from software bugs and from cyber-attacks?”

Drug development

Thanks to increase of data and size of medical database, Artificial Intelligence is used a lot at various stages of the Drug Development process, including clinical research and patient management, in order to optimize drug discovery and reduce the risk taken by patients.



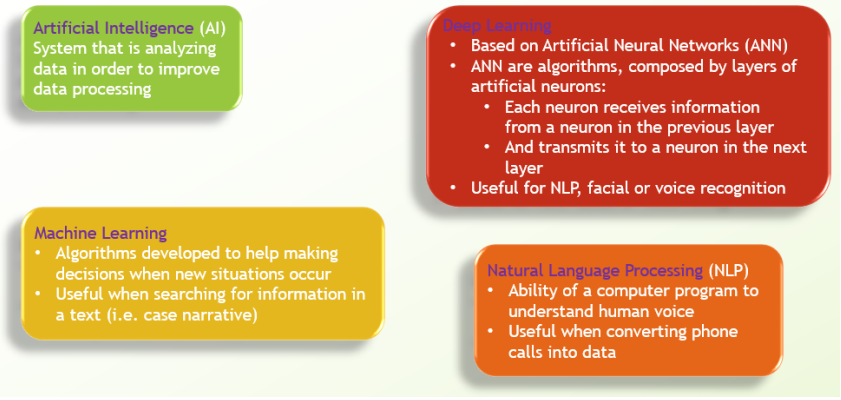
Drug discovery is risky (97% of drug programs fail) and costly (the average cost to develop a new drug is $2.8 billion [1]). There is also a societal cost, due to the fact that less than 40% of known diseases are currently treatable. On top of that, new discoveries are extremely rare, mainly through academia and government research. A lack of invention despite the fact that the time consumed in drug discovery is important (between 12 and 15 years from start to market).

We are all conscious that AI can reduce the cost of developing a new drug that Health Authorities are encouraging the introduction of AI in clinical trials in order to reduce the risks taken by patients. However, what is happening in Pharmacovigilance? We will focus here on concrete examples where AI can be useful in Drug Safety activities, and see how it can facilitate your daily work.

What is Artificial Intelligence?

First, let's see rapidly what is exactly meant by "Artificial Intelligence" and what are the key components of this technology that can be useful in PV.

AI aims to simulate human intelligence by using informatics and mathematics methodologies. It's a system that is analysing data in order to improve data processing. The final goal is to treat tasks that are currently realized by humans, in a more timely and efficient manner. Today, all the largest companies in the digital market and the information technology industry, which are processing massive quantities of data, try to apply AI solutions for handling specific issues.



We will focus on Machine Learning (ML) which represents algorithms developed to help making decisions when new situations occur and that help automating data processing. ML is training and improving itself by "learning" from big databases.

ML is useful when you are searching for information and case narrative in a text.

We will also focus on Natural Language Processing (NLP). It corresponds to the ability of a computer program to analyse and understand a human voice. To exemplify, it can be useful when you want to convert a phone call into data.

Finally, we will talk about Deep Learning. Deep Learning is a family of Machine Learning methods, based on Artificial Neural Networks (ANN). ANN are algorithms (whose operations are inspired by biological neural networks), composed by layers of artificial neurons. Each neuron receives information from another one from the previous layer and transmit it to another one from the following layer. Deep learning algorithms are led to significant advances in NLP, and also in facial recognition and voice recognition. We will see concrete examples where Deep Learning was implemented in Drug Safety: we will see the difference between human intelligence and artificial intelligence for processing individual case safety reports, and how we can use deep learning for labelling adverse drug reactions on social networks.

It is important to keep in mind that the terminology used in AI is clearly confusing.

In reality this technology is really powerful where thinking and intelligence are not required – the machine is not thinking or learning, it is simply doing a fantastic job when completing repetitive tasks: today our PV systems turn Drug Safety Experts into “Super Data Entry Operators”, with AI, the “non-intelligent” part of your job is done by the machine, and you interfere where your knowledge is required.

Why should AI be used in PV?

Application of Artificial Intelligence in pharmacovigilance means new methods to improve case management and monitoring the risk-benefit balance.

AI is interesting whenever you want to automate repetitive tasks involved in Data Processing. AI also help you to identify situations similar to your current case that may suggest a potential signal.

So maybe you wonder what the difference between traditional programming and AI is. Your actual PV system is already able to identify duplicate cases, maybe even to help you in Signal Detection. So why would you need AI?

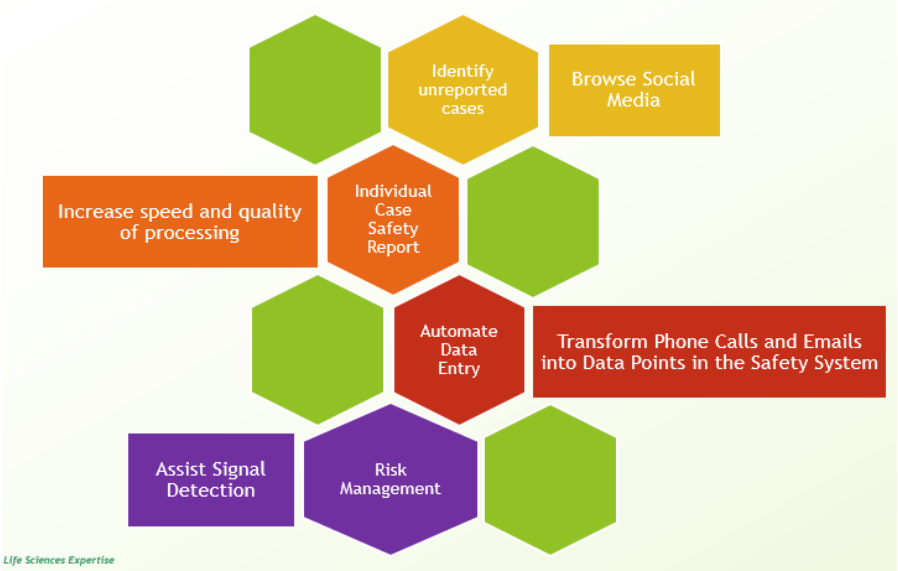
There is a big difference between traditional programming and AI: traditional programming can only work on planned scenarios, AI can also provide you directions for new situations and support complex decision making. And this specificity is very valuable in Pharmacovigilance. The PV specialist can focus on reviewing data and concentrate on medical evaluation of AEs.

At which levels can AI be useful?

For extracting specific data, artificial intelligence can be very helpful in Pharmacovigilance.

There are various Pharmacovigilance activities where AI can be useful:

- First: Identification of unreported cases by browsing Social Media.
- Secondly, Individual Case Safety Report (ICSR). From a quantitative stand point this is clearly the most important step, because AI can increase speed and quality of processing.
- However, from a valuable standpoint, AI can assist Signal Detection. Risk Management is clearly a key element where AI can improve your work



There are also other points of interest. AI can automate Data Entry by transforming phone calls and emails into Data Points in the Safety System. For example: screening social networks to find new cases, or importing phone calls into your PV database. Each one of the point above mentioned will be develop in this document.

Data processed by Artificial Intelligence

Which data may be processed by Artificial Intelligence?

As we are going to see further on, using AI for identifying and processing adverse events can be very useful: the volume of ICSR increase yearly, but it is estimated that more than 90% of AEs go unreported.

The goal of a study called "Artificial Intelligence Within Pharmacovigilance: A Means to Identify Cognitive Services and the Framework for Their Validation" published by "Springer Link" was to identify areas across the PV value chain that can be augmented by applying artificial intelligence.

According to the result of the study, there are potentially 51 candidate data points that may be involved in the application of AI in single case processing [2]. Each of these data points requires a specific AI model to be defined, which means implementing them all represent lots of work.

When implementing AI strategies, you need to prioritize the data points for which you need to develop a model.

You must first find out what are your main issues:

- Do you have non-cases sent to authorities?
- Or are serious cases processed too late?
- Do you have Incomplete MeDDRA coding?
- Depending on the products you are responsible for, some data points are more important than others.

Key steps and data points

AI can help us to process data, and that it is important to identify which data we are talking about.

Whenever you have a checklist to follow, especially when your checklist is complex, you may want to consider AI to complete the task for you.

The system won't make decisions for you, but simply propose you the most reasonable set of data, based on what you approved earlier.

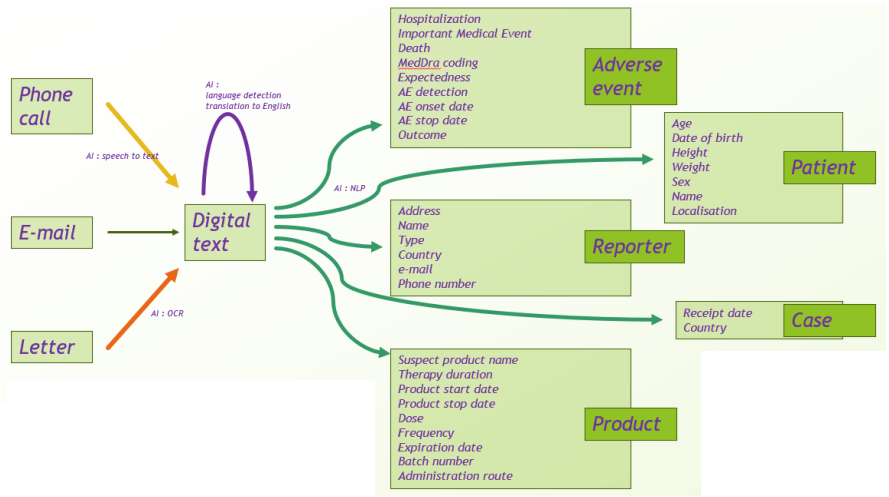
Let's see some examples of opportunities to improve Data Quality with AI in Case Processing:

- Triage: Identify better effective cases and serious AEs to ensure AEs are reported on time, and only real cases are processed as such
- Ensure consistent Medical Coding rules: facilitate Aggregate Reporting and efficient Signal Detection
- Ensure identification of suspect drugs
- Determine faster Expectedness and Causality
- Ensure consistency between Narratives and all other data
- Assist Quality Assessment

We can also mention the identification of missing information: with traditional data review it is very hard to detect missing data - you can easily identify data that is consistent, but when data is missing, how can you clearly distinguish between data not entered by error and data not entered because the information is not available?

Case workflow using AI

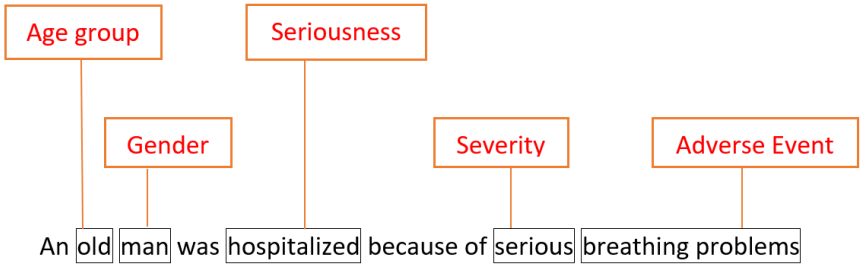
Let's have a look at a case processing workflow using Artificial Intelligence. First, you can decide to use an AI technology like speech-to-text or OCR to convert a letter, an e-mail or a phone call into digital text ready to be processed. Text processing itself starts with language detection followed by an automatic translation of the digital text into English when relevant. Then, Natural Language Processing needs to be set up for each relevant data point, including information about the patient, the reporter, the case, the product and the adverse event.



Artificial Intelligence Predictions vs. Human Processing

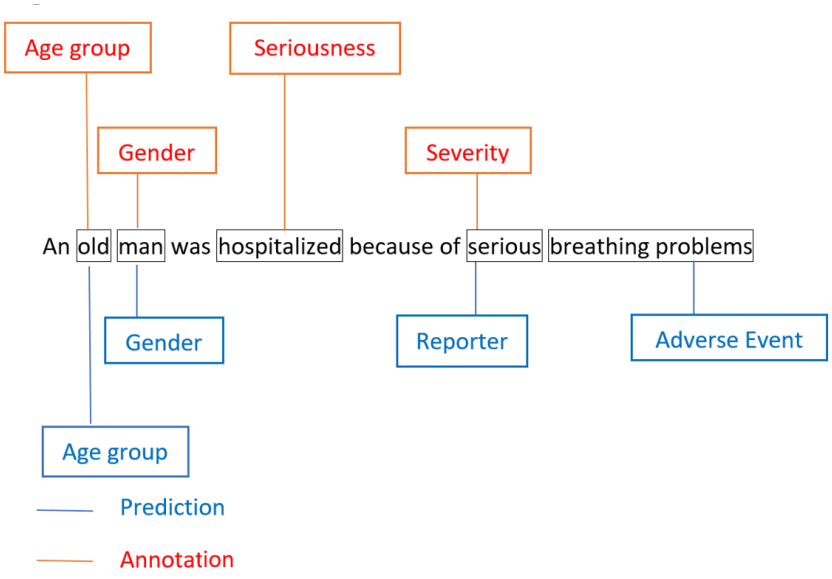
Let's take a look at how a report is processed by human data operators.

In this new example, we suppose that we received this information from one of our various sources of cases:



The sentence “An old man was hospitalized because of serious breathing problems” needs to be processed: the age, the gender, the seriousness, the severity, and the adverse reaction are information that must be identified. The following process is repeated for each sentence in each report by many different operators. With a manual process, it is difficult to maintain a high quality and a good performance, as some pieces of information may not be identified properly by the data processor. This is where Artificial Intelligence can interfere. Thanks to Natural Language Processing, we are able to develop models detecting the different information that must be identified in a sentence.

However, if AI can be better than a human, it may also sometimes be worse...



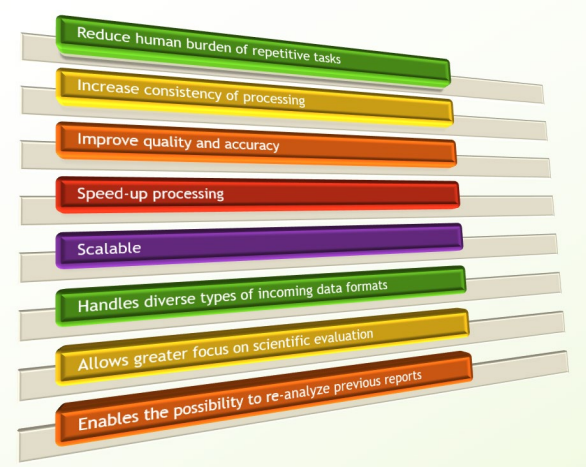
In this second example, you see in red the interpretation made by a human data operator, and in blue the processing completed via Artificial Intelligence. Our human data operator forgot to identify the Adverse Event (AE), which seems a weird error. Our AI algorithm was able to detect the AE, however, the

hospitalization information was missed, so the seriousness could not be set properly, and the severity has been associated with the wrong data point. That's unfortunately also a weird result…

AI errors can occur when facing new situations. But if this happens, we are able to improve the AI model to prevent this error from happening. With time the AI model will be more and more effective as it will face new situations. But it is important to keep in mind that AI is not a perfect solution, it also makes error. So, when implementing such techniques, you must clearly first define your priorities and objectives.

This example highlights the potential of AI in the PV field:

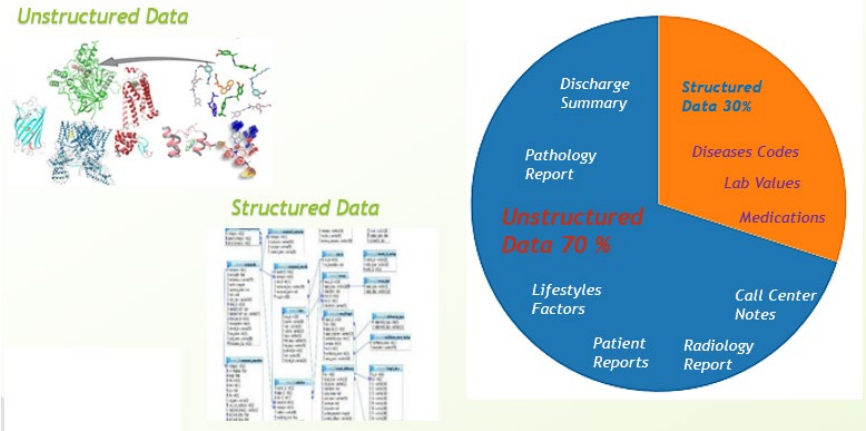
- It avoids human burden of excessive repetitive tasks, so that PV operators can focus on something more valuable than worry about the structure of their PV database
- The PV personal will still need to confirm or change the proposals made by AI, and their corrections will be used to make the program more qualitative and evolve toward better accuracy
- AI allows to increase consistency of the processing, as it does not depend on the operator, his or her mood and state of tiredness
- It is very scalable: we do not face any important issue when the volume of data to be processed increases, which is very valuable particularly for companies dealing with seasonal products: if the human operator is trained in May, you do not know what will happen in November: is this person still available? Does this person still remember the training received? However, your AI algorithm will still be reliable
- Another important point is that you can always re-analyse previous reports. This can be useful to get a better consistency of older data.



From unstructured data to structured data

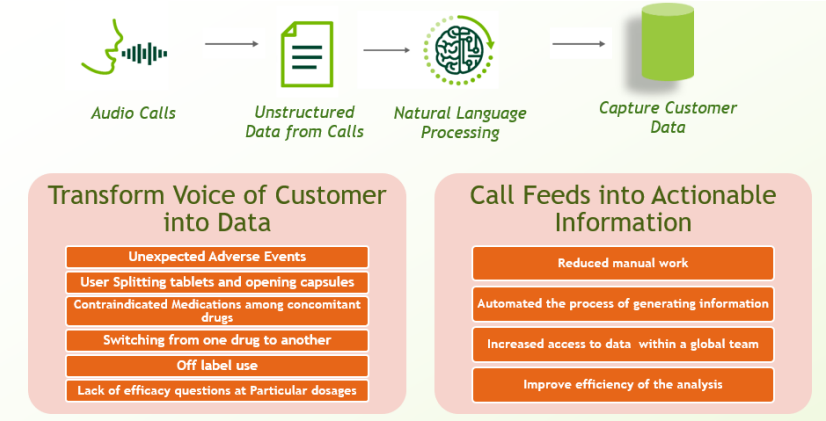
The challenge when processing pharmacovigilance data is that most of important information is not stored in clear fields identified by our pharmacovigilance database but is contained in free text. Data that don't have a pre-defined data model are called "unstructured data". Patient reports and call centre notes are examples of unstructured data. By opposition, data stocked on databases are called "structured data". This include informatics languages or lab values, among other things.

PV Data contains lots of useful information stored in unstructured data formats which are difficult to use. Due to this fact, it is not possible to solve data processing issues by using traditional programming, that's why we need to involve Artificial Intelligence systems based on Deep Learning. By applying AI methods like Natural Language Processing, we will be able to process unstructured data and transform it into structured data.



Speech-to-text to handle phone calls

Artificial Intelligence may be used to process calls and transform them into concrete data points, so that call data may be recorded in the PV database. More specifically, audio calls can be processed by a speech-to-text application for obtaining unstructured data. A Natural Language Processing algorithm will process unstructured data from calls and will capture customer data. Structuring the voice of the reporter enables easy and rapid search and visualization of the data.



Natural Language Processing (NLP)

As said earlier, Natural Language Processing is an algorithm which have the capacity to analyse a human voice (such as a phone call) or a text (such as an Individual Case Safety Report). NLP will apply for each relevant data point and capture customer data by transforming unstructured data into structured data. This way, with NLP, you can use written or spoken information in order to interpret the meaning of someone's symptoms. When processing PV cases, you must distinguish between Medical History and Adverse Event: both types of data may contain similar terms, but do not have at all the same meaning. Also, as discussed earlier, when you explore social media, you cannot just simply decide to consider each information you find as being a PV case just because it seems to be containing medical terms.

One NLP model can only work on one specific language. To be really useful, AI and specifically NLP need to confront all languages. So far, we were dealing with a text written in English but it is not always the case and reports can be written in many different languages. You must be clear that a model working on English text won't be able to extract information from another language. This is very important, as some languages are so different from English that you may really have to develop different algorithms to interpret your data. Thus, a solution is a model for each language for each feature. The languages for which this exercise will be difficult are languages where lots of common expressions contain terms that sound like a body system or a symptom, but are absolutely unrelated to health, or languages where the same succession of sounds may represent potentially various words having different meanings: because that's where voice recognition won't work well, and that's where people will make many typing errors. And of course, we wish the process to be consistent.

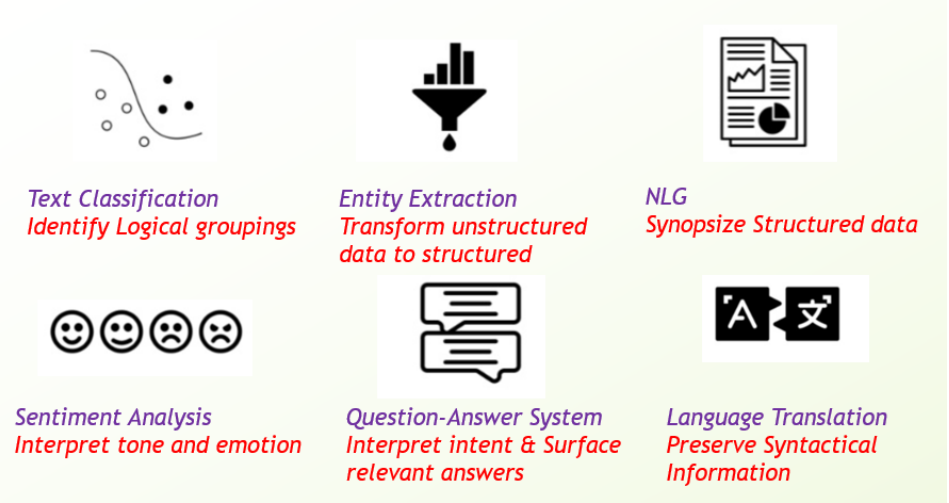
For facing these issues, many solutions: as explain earlier, we can build a model for each language for each feature. But in reality, this solution is not very efficient, especially for countries where so many different languages and dialects are spoken. Another solution is to keep human processing for some languages. But you will lose an important time in case processing. Finally, the solution recommended is an English model and a translator for each other language. Indeed, automated language processing is another favourite area of Artificial Intelligence, just like Google Translation or DeepL (linguee).

Context Matters and keywords are often largely irrelevant "BREAK A LEG !"

This said, even though NLP can be challenging, there exist different applications of natural language processing that allow the interpretation of data in the best possible way:

- Text classification: AI is able to identify logical groupings in order to process of classifying text into predefined categories based on their content.
- Entity extraction: AI transform unstructured data, like a text or an audio call, into structured data in order to capture customer data.
- Natural-language generation (NLG): Conversely, AI can also synopsize structured data, and transform it into natural language.
- Sentiment Analysis: Interpretation of tone and emotion based on words.
- Question answering (QA): System that automatically answers questions asked by a human in a natural language. QA system aim to interpret correctly the question and to bring appropriate answers.
- Language translation: As explained above, automated language processing is one of the favourite areas of AI.

Integration of a translator in a NLP model built in English is one of the most relevant use of language translation in IA.



Extract PV Data from Social Media

AI technology may allow us to identify potential new adverse drug reactions (ADR) in Social Media. Social Media surveillance is generally carried out in following steps:

- Text Mining for adverse events from Social media posts
- Filtering and Classification
- Detection of AE-drug association
- Extraction of ADRs
- Qualitative & Quantitative assessment for signal evaluation

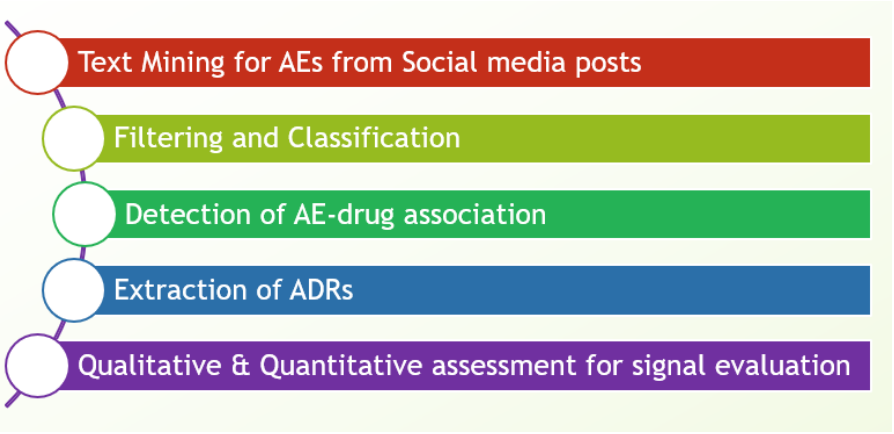
However, you need to keep in mind that even though Social Media may be a potential source for Adverse Reactions, there are lots of limitations depending on the media source, the language and the processing method, and the data permissions.

You cannot simply jump into Facebook and export data from there. You must ensure that you are authorized to collect the data and that the data you collect matches the basic requirements of a PV case. Facebook is actually not a great source of information as you do not know who the real reporter is.

From a global standpoint, forums initiated by your company, or other medical organizations are more interesting sources of information than Twitter or Instagram.

The other limitation of social media information is that the author of a post is not always medical expert. The medical terminology used may not be appropriate, and it is sometimes

difficult to ensure that the event described really happened. Also, it is good to keep in mind that some "massive reactions" on social networks in one specific country may be related to specific events occurring in that country: for example, if in a country you have one serious ADR occurring, there are big chances that you will see more cases reported in this country, and also more discussions about this drug in Social Medias. But this unique case, won't have any impact on the rest of the world. So, browsing Social Media can be useful, but you will need to process many messages until you can identify a real case for pharmacovigilance.



Searching for Adverse Events in Social Media is quite challenging, but it works.

Let's see a concrete example of deep learning approach for detecting adverse drug reactions in social media. The objective of a study published by Oxford University Press in 2017 was to develop a deep learning model that labels word in an output sequence with adverse drug reaction membership tags in Twitter posts. The best performing model developed was better for adverse drug reaction identification than a baseline lexicon system and a conditional random field model.

It is interesting to note that this model reached optimal performance with fewer training examples than other models. The conclusion of the study was that adverse drug reaction detection performance in social media is significatively improved by using deep learning approach [3].

A concrete example: WEB-RADR

WEB-RADR is a European project launched in 2014 regrouping many partners (regulatory agencies, pharmaceutical industries, academia, patient groups, etc.) through 9 European countries (Ireland, United Kingdom, France, Switzerland, Belgium,

Netherlands, Germany, Sweden and Croatia). The key challenge is to develop mobile applications enabling patients and healthcare professionals to reports Adverse Drug Reactions by using the power of social media [4].

WEB-RADR performs Named Entity Recognition (NER) in order to identify products, despite ambiguous product names, by relating the text with a dictionary of medication names [5]. The application also offers a solution to the duplication challenge. The presence of duplicate records is problematic because it may lead to an overestimation of the amount of evidence, and the use of social media in safety Signal Detection means that we are at risk that two posts refer to the same Adverse Event. For WEB-RADR, a method called VigiMatch was implemented to screen for duplicates based on comparing the textual content between posts as “bags of words” [6]. Finally, the last challenge consists in the identification, extraction and mapping of product-event and product-indication combinations in free-text. The WEB-RADR project established a publicly available benchmark reference dataset that can be used to test and compare the performance of entity recognition methods targeted at the automated identification and mapping of personal experiences of AEs and indications reported in social media, especially Twitter [7].

In conclusion, we can say that the application WEB-RADR is contributing to the advancement of pharmacovigilance. It hopefully contributes to the improvement of existing methods and systems, and the development of new ones.

Individual Case Safety Report

In this section, we will focus on case processing. The processing of individual case safety reports (ICSR) is interesting because there are various sources of information for reports (e-mails, phone calls, letters from healthcare professionals or patients, posts on dedicated web sites, health forums, literature, …) and the source of information varies from case to case. Consequently, the quantity and quality of information can vary a lot between two cases and the quality of received reports is not always satisfying (lack of information, lack of consistency). Even though there is a huge inconsistency in data, the authority requirements stay the same for each case.

Artificial Intelligence & ICSR

Regardless of the source of data, the case processing workflow requires lots of manual work.

Many aspects of case processing can be made easier by AI technologies. Artificial intelligence helps reducing the manual effort of data entry and increase the quality of the data collected.

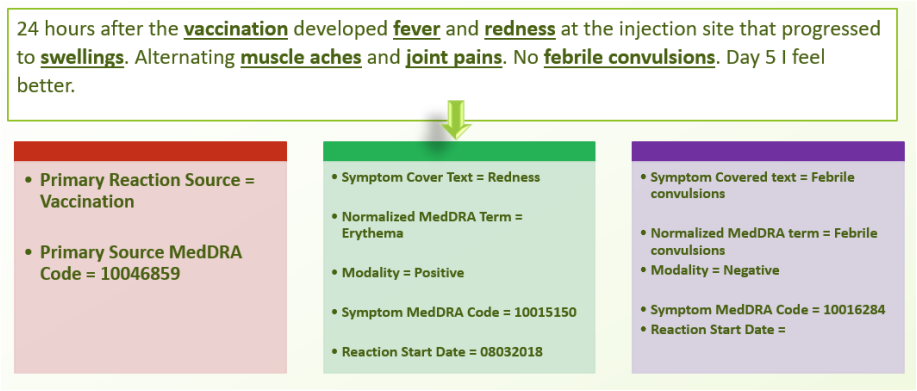
Of course, we all dream of a world where all cases are complete and sent from Authorities to be imported perfectly in our PV Database, so we just need to review and evaluate the information provided. However, in real life we receive mainly incomplete cases and must ensure we search for all missing information, fill-in all gaps, but still report cases on time.

And as we permanently look for new sources of cases, we unfortunately end up in new sources of troubles: new data sources, new types of inconsistencies, new types of incomplete cases…

Let's see now a concrete example of applying artificial intelligence in ICSR. In 2018, a study called "Sorting Through the Safety Data Haystack: Using Machine Learning to Identify Individual Case Safety Reports in Social-Digital Media" published in the journal "Drug Safety" had an objective to develop a Machine Learning model for classifying ICRS from social digital medias and compared his performance with that of human pharmacovigilance experts [8]. Three ICSR classifiers model were developed using a random sampling containing social digital media posts that mentioned Roche products and brands in combination with medical and scientific terms sourced from social media and news media blogs. Agreement with human pharmacovigilance experts was evaluated by calculating the Gwet AC1 statistic (gKappa), which is used for assessing the degree of agreement between two rates with binary outcomes [9]. The best performing model was tested against the Roche global pharmacovigilance expert using a blind dataset. The results of the study showed that the best Machine Learning model had an accuracy of 83% and a gKappa of 78%. It also showed that the model developed is nearly 1000 time faster than a human expert for performing the same task [8]. This confirms that AI could be a valuable solution to the challenge of ICSR detection.

Breaking down Complex Topics

The following example shows you how data may be automatically populated in your PV database via an algorithm driven by Artificial Intelligence.



In our example, we want to process the following text that describes the adverse events reported by a person who was vaccinated against tetanus a few days before. The algorithm driven by AI is able to detect that the primary reaction source is a vaccination. It is also able to make a difference between the presence and the absence of symptoms. By example, the presence of redness is detected as a positive modality, and the absence of febrile convulsions is detected as a negative modality. But it's not all: for each of these information, the algorithm is also able relate the data with the MedDRA term and the MedDRA code corresponding.

Signal Detection

We talked a lot so far about case processing, even though Signal Detection is a key activity in which using AI really makes sense. In fact, Signal Detection is one of the most important objectives of Pharmacovigilance and the reason for that is just because case processing was a better example to expose the components of Artificial Intelligence and discuss both the benefits and the challenges. So why is AI so important for Signal Detection? Right now, we use mainly two methods:

- Either we review cases individually and search for potential signals. But Qualitative Review is time-consuming, and some signals may be undetected.
- Or we use statistical evaluations that may not produce always the best results, particularly with a small number of cases, or with a large number of cases having lots of missing information where the relationship between the product and the reactions cannot clearly be determined. However, Quantitative Review requires enough data to make reliable statistical analysis.

There is a clear need to find new methods that allow us to identify as early as possible new risks associated with a drug. By

implementing AI in Signal Detection, you won't just simply turn text fields into PV database fields, but really propose potential candidates for Signal Detection Analysis. To exemplify, technologies using machine learning and NLP can, with the right training, conduct rapid and sophisticated analyses, with the aim of identifying signals that may hint at a hidden problem. In fact, these algorithms can be able, as time goes by, to "read" any source information. And that's really where your role of PV specialist will be involved: you can teach to your AI application where you consider that potential signs and symptoms make you worry, and AI can help you to verify that your hypothesis is realistic. So, you probably wonder whether AI can help you even if you have a small database or lots of missing information. In fact, yes. Nowadays, the health authorities provide you access to large databases, and these sources can be used by AI to help you detect a signal.

We talked earlier about machine learning methods, such as Artificial Neural Networks, which are used for processing Individual Case Safety Reports and for labelling adverse drug reactions on social networks. But ML is also useful for highlighting complex dependencies between drugs and adverse events. That was the goal of the study "Machine Learning guided association of adverse drug reactions with in vitro target-based pharmacology" published in 2020. A model that predicts ADR occurrences from in vitro secondary pharmacology of common targets for 2134 marketed drugs was developed. They identified 211 target-ADR associations (including established relations, which valid the model) occurring to a greater extent than expected by chance. These associations provide a comprehensive resource to support drug development and human biology studies [10].

Security & Regulations

What are Health Authorities thinking about that?

The most important point is of course to ensure that AI is really compliant with what authorities want us to do. Since many years, the Health Authorities push you to find more cases, and you do your best to do so, which means right now the first people who suffer from an increase in the number of PV cases are not Pharma Companies, but Health Authorities! And this fact brings them to this conclusion: they need to ensure their PV specialists focus on the most important part of their work, and AI can help them to reach this objective. So that's good news, as it proves that AI is perfectly compliant with PV Regulations.

However, from a practical standpoint, implementing software based on Artificial Intelligence is more difficult than implementing software based on traditional programming. For example, how do we validate AI systems? With traditional programming, we can clearly define our requirements and test what the program does. With AI, we want the program to come with suggestions for situations we were not able to plan in advance… So how does the user artificial intelligence test looks like in such a situation? That's the question you need to answer before moving forward. But keep in mind that the AI world is waiting for you…

Artificial Intelligence creates new way to receive healthcare and lead Health Authorities to raise new questions. One of them is the question of reimbursement of Healthcare for people treated by medical process using AI algorithms. The mission is to define criteria for determining how the application is beneficent for the patient and for the public health. Another question arises of how approve medical devices using AI. Some results and prediction provide by AI application are simply not verifiable and sometimes, patients and healthcare professionals have no choice other than trust the application. Clinical studies for validation of an AI application are not the same as clinical studies for validation of a new drug.

Protection of Personal Data

In the introduction, we asked ourselves about security as long as we transfer PV responsibilities to software using AI methods. When you process data, you have the obligation to ensure the protection of personal data. The application of some of the traditional principles of data protection may be challenging when using Artificial Intelligence [11].

- The "Convention for the protection of individuals with regard to automatic processing of personal data", also called "Convention 108", is an international legal framework regarding the right to privacy of individuals. It was accepted in 1981 and modernized in 2018. 55 countries around the world ratified the Convention 108.
- The General Data Protection Regulation (GDPR) is a regulation on data protection and privacy. It applies in the European Union and in the European Economic Area since 2018. It also applies to collaborators outside EU and EEA if they are processing data from subject within the EEA.

Thus, Pharmacovigilance professionals have the obligation to be compliant with Convention 108 and GDPR. But what means exactly "being compliant with regulations"? It means that you have the obligation to ensure the protection of data privacy. Indeed, when you talk about artificial intelligence, you talk about collecting and processing massive databases, which can include personal data. For example, regulations oblige you to first ensure that the data collection is useful and necessary. Convention 108 provides specific safeguards that can be applied to AI algorithms used for automated decision-making systems [12]. You also have to test strongly your model, in order to detect unexpected bias that can sometime be surprising… In 2016, Microsoft launched 'Tay', an artificial intelligence chatter boot on Twitter. However, it caused controversy, after posting offensive and racist tweets [13]. More recently, we learned that an Artificial Intelligence tool used by Amazon was judged misogynist. The tool's function was to analyse hundreds of CV in order to determine the most interesting and it showed bias against woman [14].

Which perspective should we adopt for being compliant? A solution can be transparency. When using an AI application, we should give a description of the way the algorithm structure access to data during the learning stage. We should also detail theoretical mathematical methods used, in order to explain clearly how data are processed, and what we aim to prove. Transparency can also address the issue of people self-determination in data processing. Indeed, the complexity and the obscurity that often characterize AI algorithms necessarily affect data subjects consent in terms of self-determination [12]. Given this limit, data protection regulations are increasingly emphasising the role of risk assessment. Risk assessment allows controllers to plan measures to control risks and ensure the appropriate processing of personal data [15]. Carried out by data controllers and a safe AI environment, risk assessment have a positive impact on people trust and on their willingness to use AI applications [12]. Examples of risk that can results from personal data processing:

- Social disadvantage (refusal of access to benefits, discrimination, damage to reputation…)
- Unauthorised reversal of pseudonymisation, identity theft, defamation
- Financial losses, fraud
- Spam, targeted advertising

Finally, companies using Artificial Intelligence may set up an ethics committee to improve the risk assessment process and transparency. The review made by ethics committees can provide a significant support to AI developers to design rights oriented algorithms. Moreover, the necessary dialogue between the developers and the committees experts can positively boosts the creation of more transparent and less risky data processing procedures [12].

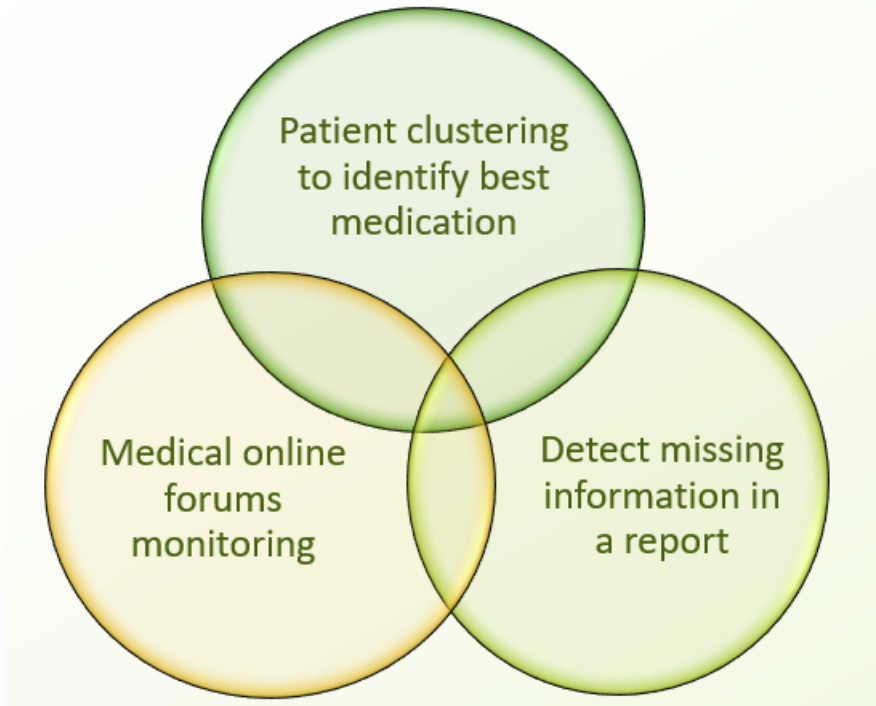
Conclusion

Finally, what can Artificial Intelligence do for you? We saw earlier some benefits of Artificial Intelligence. AI is scalable, improve quality and accuracy of programs and reduce human burden of repetitive tasks.

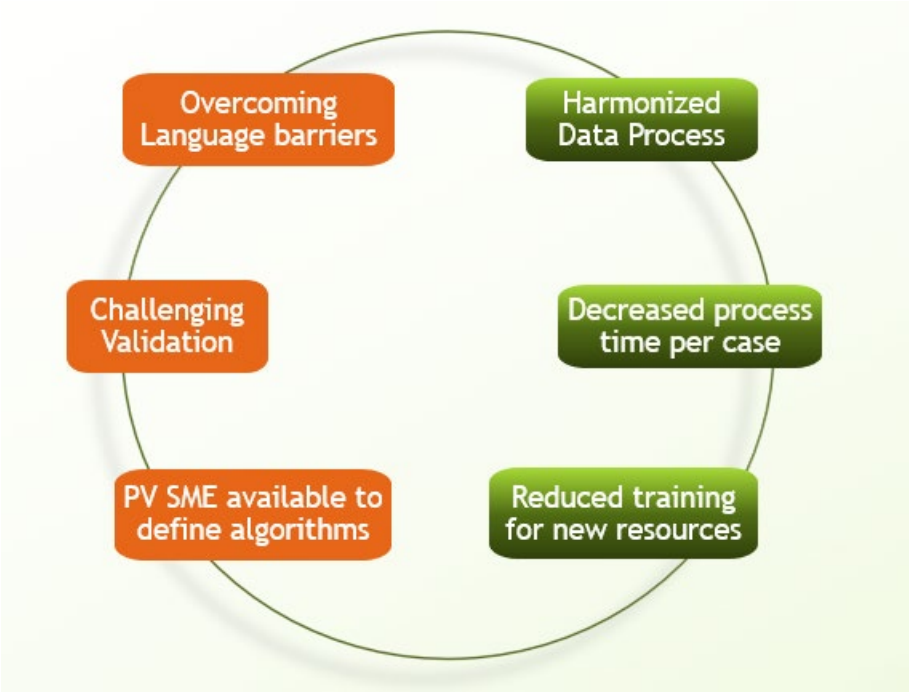
For Individual Case Safety Reports, Artificial Intelligence can clearly reduce the time allocated to case processing by automating manual and routine tasks, like case follow-ups to verify information and capture any missing data. AI support activities that require medical knowledge and expertise, and advanced analytical skills, speed literature searches for relevant information and transform scanned documents on AEs into actionable information. AI features allow you to determine whether a case requires expedited processing and reporting. Then the information processed by AI is proposed to the drug safety specialist for review and confirmation or correction. This information, once validated by the PV specialist, serves as input for subsequent machine learning rounds, improving the algorithms over time.

Besides case processing, AI may be involved in other PV Activities, for example:

- Patient clustering to identify the best medication: this can be achieved by analysing a patient profile versus reported adverse events for all drugs relevant to that case
- Detection of missing information in a report can be used in real-time when a report is registered. Imagine a speech to text algorithm able to detect all required piece of information and highlight which is missing so that the operator may ask any further relevant question
- Automatic review of social media to detect if people report adverse events publicly on Social Networks. As many adverse events are not reported, it is a nice way to address the lack of received data.



AI has really a lot to offer to Pharmacovigilance, however, PV has also a lot to offer to Artificial Intelligence. We are far away from a world where you just simply hire a data scientist and consider the work is done. You must ensure your AI system matches your needs and the regulatory needs. Implementing this technology requires good planning and lots of testing.



Sources:

1. Wouters, O.J., McKee, M., Luyten, J., (2020) Estimated Research and Development Investment Needed to Bring a New Medicine to Market 2009-2018, *JAMA*. 2020;323(9):844-853. doi:10.1001/jama.2020.1166
2. Mockute, R., Desai, S., Perera, S., Assuncao, B., Danysz, K., Tetarenko, N., Gaddam, D., Abatemarco, D., Widdowson, M., Beauchamp, S., Cicirello, S., Mingle, E. (2019) Artificial Intelligence Within Pharmacovigilance: A Means to Identify Cognitive Services and the Framework for Their Validation, *Springer Link*
3. Cocos, A., Fiks, A.,G., Masino, A., J., (2017) Deep learning for pharmacovigilance: recurrent neural network architectures for labeling adverse drug reactions in Twitter posts, J Am Med Inform Assoc. 2017 Jul 1;24(4):813-821. doi: 10.1093/jamia/ocw180. PMID: 28339747; PMCID: PMC7651964.
4. *WEB-RADR* [Online], Recognising Adverse Drug Reactions (accessed December 15, 2020) WEB-RADR (web-radr.eu)
5. Ellenius, J., Bergvall, T., Dasgupta, N., Hedfors, S., Pierce, S., Norén, G.N., (2016), Medication Name Entity Recognition in Tweets Using Global Dictionary Lookup and Word Sense Disambiguation, *PHARMACOEPIDEMIOLOGY AND DRUG SAFETY* 25, 414-415
6. van Stekelenborg, J., Ellenius, J., Maskell, S. et al., (2019), Recommendations for the Use of Social Media in Pharmacovigilance: Lessons from IMI WEB-RADR, *Drug Safe*, 42, 1393–1407, https://doi.org/10.1007/s40264-019-00858-7
7. Dietrich, J., Gattepaille, L.M., Grum, B.A., Jiri, L., Lerch, M., Sartori, D., Wisniewski, A., (2020), Adverse Events in Twitter-Development of a Benchmark Reference Dataset: Results from IMI WEB-RADR, *Drug Safe*, 43, 467–478, https://doi.org/10.1007/s40264-020-00912-9
8. Comfort, S., Perera, S., Hudson, Z., Dorrell, D., Meireis, S., Nagarajan, M., Ramakrishnan, C., Fine, J. (2018) Sorting Through the Safety Data Haystack: Using Machine Learning to Identify Individual Case Safety Reports in Social-Digital Media, *Drug Safety*
9. Ohyama, T. (2019) Statistical inference of Gwet's AC1 coefficient for multiple raters and binary outcomes, *Taylor & Francis Online*, https://doi.org/10.1080/03610926.2019.1708397
10. Ietswaart, R., Arat, S., Chen, A., Farahmand, S., Kim, B., DuMouchel, W., et al., (2020), Machine learning guided association of adverse drug reactions with in vitro target-based pharmacology, *EBioMedicine*, https://doi.org/10.1016/j.ebiom.2020.102837
11. *Council of Europe* [Online], (2017), Guidelines on the protection of individuals with regard to the processing of personal data in a world of Big Data, https://rm.coe.int/CoERMPublicCommonSearchServices/DisplayDCTMContent?documentId=09000016806ebe7a
12. Mantelero, A., (2018), Artificial Intelligence and data protection: Challenges and envisaged remedies, Report on Artificial Intelligence - A. Mantelero (coe.int)
13. *The Guardian* [Online], Tay, Microsoft's AI chatbot, gets a crash course in racism from Twitter (24 March 2016), Tay, Microsoft's AI chatbot, gets a crash course in racism from Twitter | Artificial intelligence (AI) | The Guardian
14. *Reuters* [Online], Amazon scraps secret AI recruiting tool that showed bias against women (11 October 2018), Amazon scraps secret AI recruiting tool that showed bias against women | Reuters
15. *Office of the data protection Ombudsan*, Risk assessment and data protection planning