Using AI to drive innovation in drug combination products

D. Koschel [1], W. Huang [1], C. Brunet[1] , F. Guely [2], Z.Rajil [2], M-M. El Dandachi [3], V. Vigneron [3], P. Pham [4] [1] Becton Dickinson PS R&D, Pont-de-Claix ; [2] Meetsys – Bassetti Group; [3] Laboratoire IBISC, Université d'Evry Paris-Saclay; [4] Atlas. SATT Paris Saclay

Artificial intelligence (AI) and its subsets, such as machine learning and deep learning, are increasingly capturing interest across various fields, especially in pharmaceutical research and development. These technologies assist in making decisions in complex scenarios. Numerous studies have shown their application in different aspects of pharmaceutical R&D [1] . By employing advanced modeling techniques, these computational tools can improve efficiency and accuracy, manage complex data, and enable novel discoveries. BD leverages the use of artificial intelligence in different ways and two specific examples of the utilization of artificial intelligence in R&D are showcased:

1 - Deep learning for Sub Visible Particulate (SVP) analysis

Background and objectives

Why do we perform SVP analysis and what's the challenge ?

As drug formulation technologies advance, a significant challenge remains in ensuring drug integrity due to interactions between drugs and pre-filled container components (Figure 1). Flow Imaging technologies like MicroFlow Imaging (MFI) (Figure 2) can count and capture images of subvisible particles (SVP) in solutions, helping assess drug stability. However, accurately distinguishing protein aggregates [2] from other particles (such as silicone oil droplets and air bubbles, see Figure 3)) continues to be difficult, despite various efforts, including manual annotation [3], and the use of the S factor [4], which have limitations in terms of time and accuracy.

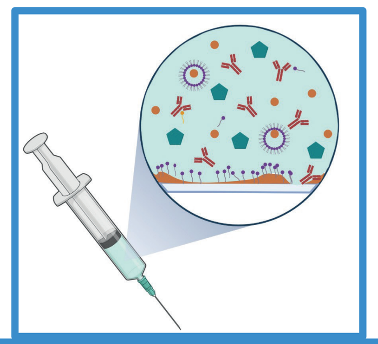


Figure 1: Complex drug formulation in prefilled container [5]

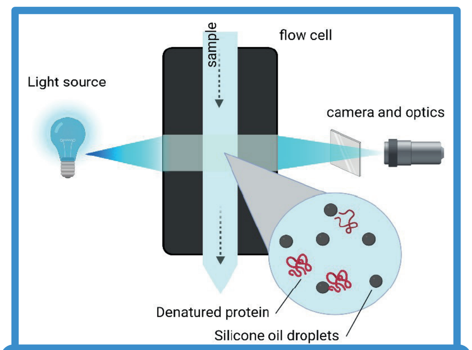


Figure 2: Schematic view of SVP analysis by MFI [5]

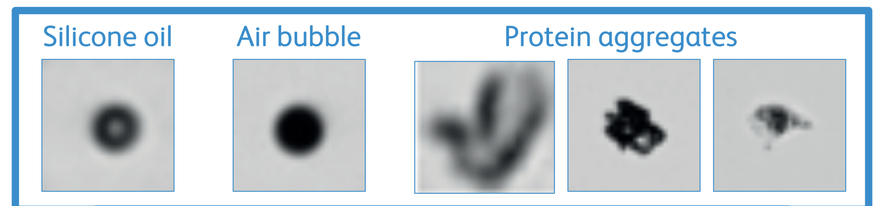


Figure 3: Example of SVP images captured by MFI

Why do we want to use deep learning for SVP analysis ?

With the increasing and inspiring compatibility of AI for imaging use, BD decided to set up a tripartite collaboration with IBISC lab at University of Évry Paris-Saclay and SATT Paris-Saclay, in order to deploy the state of art deep learning technology to resolve the problematic of SVP image classification [6].

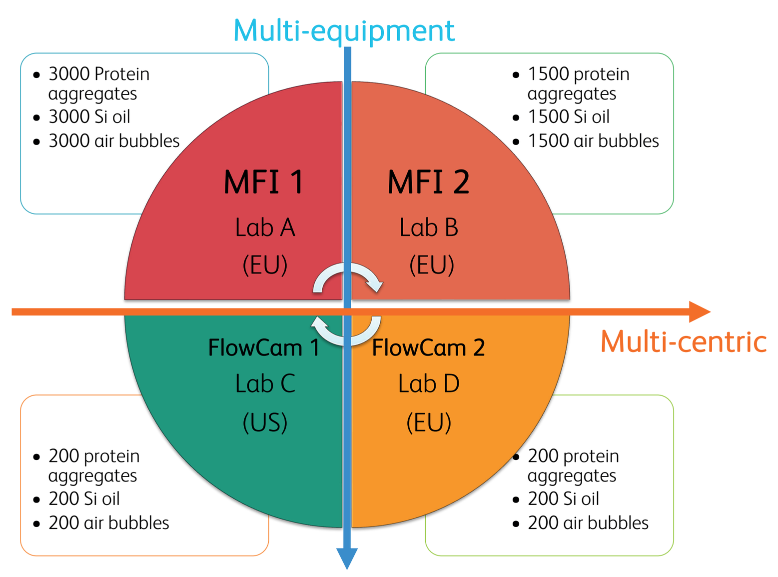
Our objectives

- Help improve interpretation of subvisible particulate
- Allows a screening of drug degradation to support primary container selection & secure drug development timeline
- Accelerate development of suitable primary container for new drug formulation

Methodology

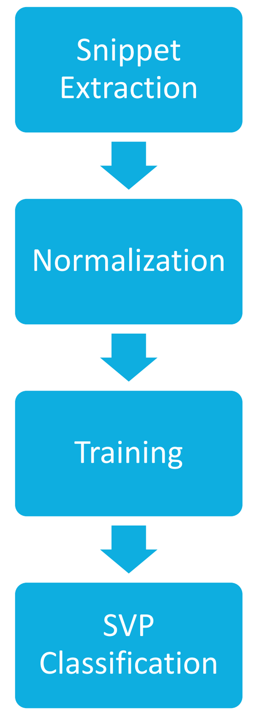
Data base

Specific SVP image databases are created in BD with reference sample solution on different equipment and in different labs. The goal is to target three classes of SVP: protein aggregates, silicone oil droplets and air bubbles.



Method

- Utilization of a Convolutional Network (CNN) for image classification
- Inputs: Normalized 32x32 snippets
- CNN model, 6 layers and ~25k parameters
- Outputs: 2 (or 3) classes for classification prediction



What are the main results so far?

We aim to achieve 98% accuracy in distinguishing between the three classes mentioned above. The following table presents our results compared three existing methods: S-Factor, Linear Discriminant analysis (LDA) [7] and Support Vector Machines (SVM) [8].

Classification method	S-Factor	ADL	SVM	OUR METHOD
Mono-class	71.1%	N/A	N/A	97.6%
Multi-class	N/A	62.5%	91.1%	97.1%

Conclusions and next steps

- Better results than the State of the Art methods and therefore better interpretation of subvisible particulate data
- "Easy to implement" innovation

- Missing physical characterization between different nature of protein agregates
- Some difficulties of SVP annotation with MFI images

[1] P. Suriyaamporn et al. The Artificial intelligence Powered New Era in Pharmaceutical Research and Development: A review article, *AI in Phramaceutical Development and Technology,* 25, 188, 2024
[2] C. Probst. Characterization of protein aggregates, silicone oil droplets, and protein-silicone interactions using imaging flow cytometry. *Journal of Pharmaceutical Sciences*, 109(1) :364–374, 2020.
[3] Rio et al. Sub-Visible Particle classification and label consistency analysis for flow imaging microscopy via machine learning methods, *Pharmaceutical Biotechnology*, 113, 880-890, 2024

2 - Knowledge capitalization

Background and objectives

Capturing knowledge is essential in an R&D organization. Retaining valuable knowledge helps avoid past mistakes, foster innovation and provides a strategic edge by leveraging unique insights and expertise [9]. For our pharmaceutical customers or partners, this results in quickened and derisked medical device developments.

A few insights about the difference between explicit and tacit knowledge:

Explicit Knowledge
Recorded and structured data, accessible and communicable in documents

Tacit Knowledge
Mental representations deeply rooted in people and their experiences and intuition

Capturing and transferring tacit knowledge within an organization can be challenging due to its unspoken and experiential nature.
However, by implementing it in a database, valuable insights, tips, best practices and skills are retained and shared across the team.

About the data base

MeetSYS, now part of the BASSETTI Group, has developed innovative databases for knowledge management through a combination of expert methodologies and advanced software solutions. BD uses one of their key solutions: i²Kn (Intelligence to Knowledge Network) platform.

MeetSYS's methodology used together with BD involves several steps illustrated in Figure 4.

Since the knowledge management roll out in R&D, around 1500 pages were created in BD. There are around 200 contributors to the data base and more than 100 readers. In 2023, 184 new pages were created on important R&D topics and 20 to 25 new comers learning curves were therefore accelerated while reducing oversight effort for senior experts.

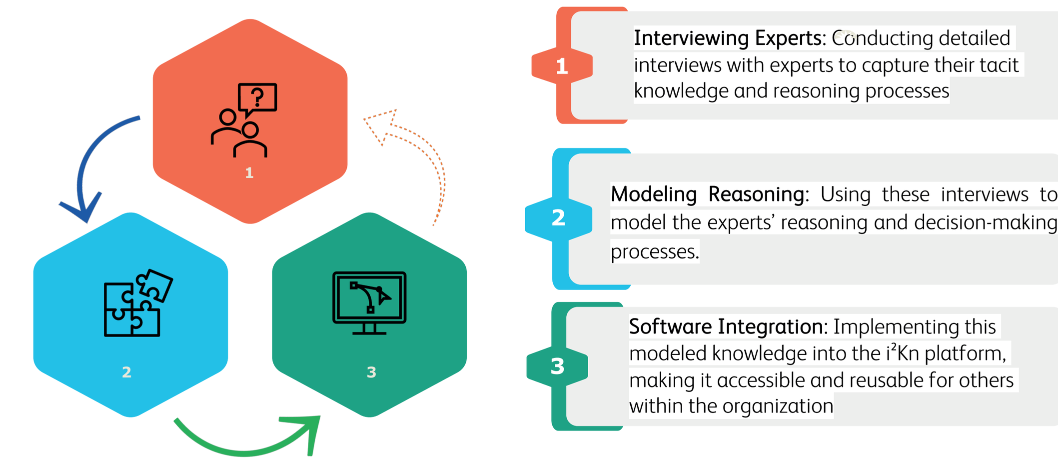


Figure 4: Model of knowledge management steps

What are the advantages to use AI in the knowledge management (KM)?

Benefits of KM before AI implementation
In most companies, employees lose up to 2h/day looking for information [10]. This time can be greatly reduced with the implementation of strong KM tools. This results in a company capable of reducing delivery delays and striking out any production anomalies. Moreover, it is known that proved the introduction of an efffective Knowledge management in an organization has an effect on the generation of innovation distinctive competences [11] leading to constantly exploring new and more effective solutions.

AI implementation
KM efficiency resides in its adoption and the use by the employees.
AI greatly facilitates the adoption of KM tools, as it introduces the possibility of querying the company's knowledge in natural language and providing reliable and confidential answers.

The chosen system is based on the Large Language Model (LLM) - a type of Natural Language Processing model. This model tends to produce hallucinations - answers that appear plausible but are not, due to AI errors or incorrect data.

To limit this problem, Meetsys decided to use the LLM only for the representation learning and form generation part (Figure 5), leaving the content and substance to verified expertise: this is where BD experts play a major role.
The implementation of new knowledge in i²Kn can also be made more efficient by introducing AI summarization of large documents.

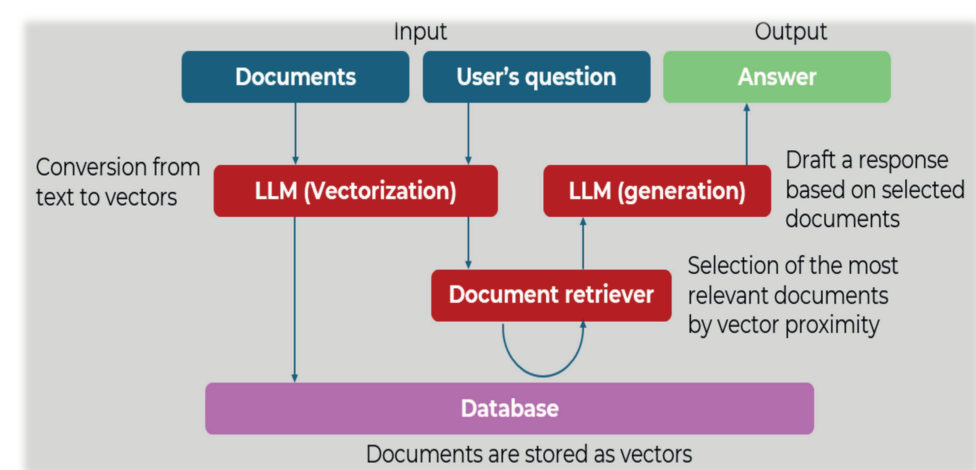


Figure 5: Utilisation of the LLM for vectorisatin and generation only

Conclusions and next steps

Meetsys and BD are currently working on increasing the reliability rate of the AI answers which is the main challenge for most AI. This will be done by setting a proximity threshold to the document retriever rather than systematically selecting the 5 closest vectors. AI is also being explored by the rest of Bassetti Group for other uses such as automatic classification.

[4] R. Strehl, et al. Discrimination between silicone oil droplets and protein aggregates in biopharmaceuticals: a novel multiparametric image filter for sub-visible particles in microflow imaging analysis. *Pharm Res*, 29(2), 594-602. 2012
[5] M. Rodriguez. Interactions between biological drugs and materials for medical use : a study at the interface of the silicone oil layer and the formulation in pre-filled syringes *[Université Grenoble Alpes*]. 2023 https://theses.fr/2023GRALV059
[6] Vitrine technologique - ATLAS. (2024). https://satt-paris-saclay.fr/vitrine-technologique/atlas/
[7] D. Choubey, M. Kumar, V. Shukla, S.Tripathi, V. Dhandhania Comparative Analysis of Classification Methods with PCA and LDA for Diabetes, *Current Diabetes Reviews*, 16, 833 -850, 2020
[8] G. Brereton, G.R. Lloyd, Support Vector Machines for classification and regression, 135 (2) *The Analyst*, 230 – 67, 2009
[9] C. Li , S. Ashraf, F. Shahzad, I. Bashir, M. Murad, N.Syed, M. Riaz. Influence of Knowledge Management Practices on Entrepreneurial and Organizational Performance: A Mediated-Moderation Model. *Front Psychol.* 11 – 2020
[10] McKinsey Global Institute, The social economy: Unlocking value and productivity through social technologies, 2012
[11] D. Palacios, I. Gil, F. Garrigos,. The impact of knowledge management on innovation and entrepreneurship in the biotechnology and telecommunications industries. *Small Bus Econ* 32, 291–301, 2009

BD and the BD Logo are trademarks of Becton, Dickinson and Company. © 2024 BD. All rights reserved.