The Advanced Machine Learning for Innovative Drug Discovery (AIDD) project

Newsletter 31st August 2023

The **Advanced Machine Learning for Innovative Drug Discovery** (AIDD) project is a Marie-Skłodowska-Curie Innovative Training Network (ITN) for Early Stage Researchers (ESRs) funded by the European Commission under the Horizon 2020 Programme, Marie Skłodowska-Curie grant agreement No 956832. The project brings together fifteen academic and industry beneficiaries from ten European countries and the University of British Columbia (Canada) to train sixteen PhD students in close collaboration with partners from the USA, Australia, China, Israel, and other countries.

Project development

The AIDD project started on 01.01.2022 and will finish at the end of March 2025.

Fifth AIDD School – Summer School on Advanced Machine Learning in Gothenburg

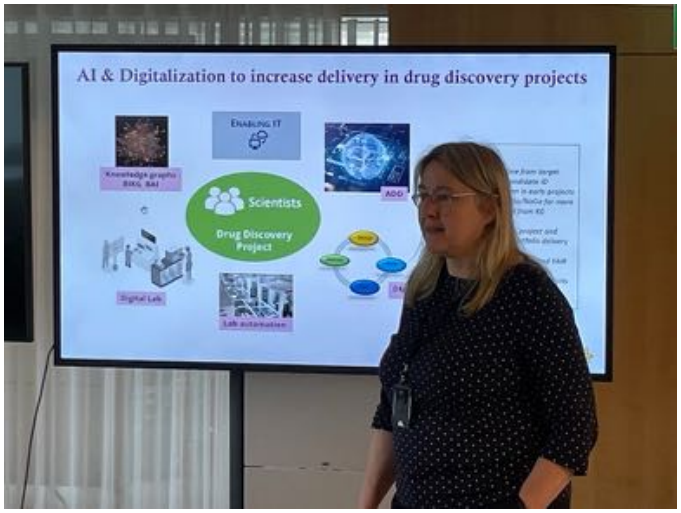
The 5th AIDD school was organised by AstraZeneca from July 3rd to 7th. The event took place at the AstraZeneca Gothenburg site just south of the city of Gothenburg in Sweden. The Gothenburg site is one of four AstraZeneca strategic R&D centres globally and has around 2,800 employees including scientists from more than 30 different countries. The Gothenburg site represents around 25% of the global research and development resources of the company.

Photo: AstraZeneca R&D Gothenburg / AstraZeneca image and broadcast library.

Prof. Ola Engkvist welcomed attendees including fellows and PIs from various locations around Europe and Canada. The program focused on ML/AI methods relevant to drug discovery and featured speakers from AstraZeneca, Swedish researchers, and various consortium partners. The school was complemented by a full schedule comprising an 'AmAZing Journey' site tour, iLab tour and visit to the nearby GoCo Health Innovation City. There were also short talks at the welcome session given by early career research staff working in the AstraZeneca Molecular AI department. Zoom broadcasts were organized for increased outreach, and almost all speakers presented lectures on-site. Audience members also included local PhD students who attended several talks.

Swedish researchers also provided short presentations around their research during the welcome session.

The school started on Monday 3rd July with a welcome by local organisers Prof. Ola Engkvist and Dr. Samuel Gehenden. The scientific program began with a talk on AI from a medicinal chemistry perspective by Head of Respiratory & Immunology Medicinal Chemistry, Dr. Werngard Czechtizky. She gave an overview of how AI has the potential to increase productivity and reduce timelines in drug discovery projects.



Dr. Werngard Czechtizky spoke about AI from a medicinal chemistry perspective.

This was complemented by a talk from Dr. Per-Ola Norrby which focused on how predicting reaction selectivity is used within drug discovery projects. In the afternoon there was the visit to the GoCo site, where fellows saw first-hand the new urban science centre where over 200,000 square meters are being developed for life science activities.

On the second day, the talks started with an overview of how to overcome safety liabilities with machine learning, provided by local staff scientist Vignesh Subramanian. The talks were continued by Data Scientist, Feng Gu, who described the journey of developing AI solutions for digital pathology data. After lunch, the AmAZing Journey site tour gave the opportunity for all fellows to explore the Gothenburg site. The day ended with two talks, the first provided by Associate Director, Ufuk Kirik, on the promise of graphs & graph-based learning in drug discovery, followed by the first external speaker of the event, Principal Investigator Lucie Delemotte, who talked about data-driven enhanced sampling of conformal changes in membrane proteins.


Attendees networking over lunch

On the third day, school began with a presentation by Engineering Lead, Michaël Ughetto, on an engineering view of biomedical knowledge graphs. The fellows continued with a talk from former AstraZeneca employee and now Principal Investigator at Guangzhou Laboratory, Hongming Chen, who provided insight into the new journey on AI modelling after REINVENT. Next, there was an excursion to the archipelago island Vrångö in the afternoon, which included beautiful scenery, classic seaside villages and lots of great beaches.

The school included an excursion to Vrångö / Hans Gamborg

The Thursday kicked off with a double session of hands-on tutorials by the REINVENT and CAZP software packages developed by the Molecular AI department. Fellows were given the opportunity to experience the day-to-day work of a computational chemist at AstraZeneca using the tools in an example discovery project. After the tutorial of both tools, fellows were given flexibility to focus more on either the *de novo* generation or synthetic route prediction aspect of the work. After lunch, an introductory talk and tour of the iLab facilities were given by Associate Principal Scientist, Tove Slagbrand. Fellows were able to see first-hand the wet-lab machinery used to generate the valuable datasets for modelling. After the iLab tour, a final virtual talk on the future of AI and LLM models was given by Principal Investigator, Mike Preuss.

Mike Preuss provided a virtual talk that focused on the future of AI and LLMs

The final day was hosted at AI Sweden in the Ericsson building on the north river embankment in Gothenburg. The day started with an introduction to AI Sweden (of which AstraZeneca is a partner) by Emma Ytterström, followed by a tour of the Edge Learning Lab. The day ended with a talk on privacy preserving ML at AI Sweden by Johan Östman. With this final presentation, the summer school was officially finished.

Voices about the school

"I really enjoyed my experience, and I appreciated the effort that went into organizing the event and curating the lecture content. The theme of the school, 'AI-driven molecular design in the pharmaceutical industry,' was a fitting choice given the current trends in the field. The lectures covered a wide range of topics, and I found them insightful, especially in shedding light on the practical applications of novel AI models in molecular design and other chemistry-related problem-solving. It was a great opportunity to get a comprehensive overview and delve into some detailed discussions. One of the standout aspects for me was reconnecting with other PhD students from our program. Seeing familiar faces and catching up on each other's progress is always good. The workshops and activities, like the trip to Gothenburg's archipelago and the visit to AI Sweden, the visit to experimental labs, added a nice mix of hands-on experiences to the theoretical aspects of the school. The REINVENT+CAZP workshop was particularly interesting, as it allowed us to step into the shoes of senior scientists and apply our knowledge to real-world scenarios. It was a practical exercise showcasing the practicality of our learning. In a nutshell, the AstraZeneca-hosted PhD Student School was a great experience. It offered a balanced blend of academic insights and networking opportunities, creating an environment where we could learn, engage, and collaborate."

- *Varvara Voinarovska, AIDD Fellow, AstraZeneca*

"As a graduate data scientist at AstraZeneca, it was very nice to be invited to the AIDD School hosted here at AstraZeneca. I especially appreciated the opportunity to network with the AIDD PhD students and hear more about their experiences and projects. There was a good variation of topics, many engaging talks and interesting site tours. Overall, I believe the school was very well planned, with plenty of variation and opportunities to learn something new. I'm grateful for the opportunity to take part!"

- *Emma Rydholm, Graduate Data Scientist, AstraZeneca*

Meet AIDD Partners (see continuously updated list on the AIDD website)

- *6th Artificial Intelligence in Chemistry Symposium*, 4th- 5th September Cambridge, UK, ***AIDD presenter***: *Samuel Genheden*
- *Conferentia Chemometrica 2023, 10-13th September, Siófok, Hungary,* ***AIDD participant:*** *Igor Tetko*
- *EUROTOX 2023, 10-13 September, Ljubljana, Slovenia,* ***AIDD presenter****: Djork-Arné Clevert*
- *PHYMOL Conference, 12-15 September 2023, Luxembourg****, AIDD participants:*** *Alexandre Tkatchenko and Leonardo Medrano Sandonas*
- *Symposium of the Polish Bioinformatics Society, 13-15 September 2023, Gliwice, Poland,* ***AIDD keynote speaker****: Djork-Arné Clevert*
- *ICANN2023, 23-26 September, Cyprus,* ***AIDD participants:*** *Igor Tetko, Michael Wand*
- *13th Global Summit on Regulatory Science, 27-18 September 2023, Parma, Italy,* ***AIDD speaker****: Djork-Arné Clevert*
- *Congress IEEE NanoPeru 2023: Trends in Nanoscience and Nanotechnology, 2-3 October 2023, Lima, Peru (hybrid event****), AIDD speaker****: Leonardo Medrano Sandonas*
- *BiotechX Europe, 4-6 October 2023, Basel, Switzerland,* ***AIDD speaker****: Djork-Arné Clevert*
- *Crash TEsting machine learning force fields: Applicability, best practices, limitations, 23-25 October 2023, Luxembourg,* ***AIDD participants:*** *Alexandre Tkatchenko and Leonardo Medrano Sandonas*
- *Broad Institute Machine Learning in Drug Discovery Symposium, 27th October 2023, Cambridge, Massachusetts U.S. (hybrid event),* ***AIDD speaker****: Djork-Arné Clevert*
- *The Scheele Symposium 2023: Adventures in Chemical Biology – basic science translating into patients, 9th November 2023, Stockholm, Sweden,* ***AIDD Speaker****: Ola Engkvist*
- *17. International Conference on Learning Representations (ICLR), 2-3 December, Sydney, Australia****, AIDD participant:*** *Jürgen Schmidhuber*
- *NeurIPS 2023, 10-16 December, New Orleans, USA,* ***AIDD participant:*** *Adam Arany, Günter Klambauer, Alan Kai Hassen*
- *ASEAN HPC School, 11 – 16 December 2023, Bogor, Indonesia,* ***AIDD participant****: Samuel Genheden*

Project publications with participation from AIDD fellows

Svensson, E.; Hoedt, P.-J.; Hochreiter, S.; Klambauer, G. Robust Task-Specific Adaption of Models for Drug-Target Interaction Prediction; 2022., https://openreview.net/forum?id=dIX34JWnIAL

Hassen, A. K.; Torren-Peraire, P.; Genheden, S.; Verhoeven, J.; Preuss, M.; Tetko, I. V. Mind the Retrosynthesis Gap: Bridging the Divide between Single-Step and Multi-Step Retrosynthesis Prediction; 2022. https://openreview.net/forum?id=LjdtY0hM7tf

Andronov, M.; Voinarovska, V.; Andronova, N.; Wand, M.; Clevert, D.-A.; Schmidhuber, J. Reagent Prediction with a Molecular Transformer Improves Reaction Data Quality. *Chemical Science*, 2023. https://doi.org/10.1039/D2SC06798F.

Sarkis, M., Fallani, A., Tkatchenko, A. Modeling Non-Covalent Interatomic Interactions on a Photonic Quantum Computer. arXiv:2306.08544, 2023. https://arxiv.org/abs/2306.08544

Radaeva M., Ban F., Zhang F., LeBlanc E., Lallous N., Rennie P.S., Gleave M.E., Cherkasov A. Development of Novel Inhibitors Targeting the D-Box of the DNA Binding Domain of Androgen Receptor. *International Journal of Molecular Sciences*. 2021; 22(5):2493. https://doi.org/10.3390/ijms22052493

Cremer, J., Medrano Sandonas, L., Tkatchenko, A., Clevert, D.A., de Fabritiis, G. Equivariant Graph Neural Networks for Toxicity PRediction. ChemRxiv, 2023. https://doi.org/10.26434/chemrxiv-2023-9kb55-v2

Kopp, A., Hartog, P., Šícho, M., Godin, G., Tetko, I. openOCHEM consensus model wins Kaggle First EUOS/SLAS Joint Compound Solubility Challenge. ChemRxiv, 2023. https://doi.org/10.26434/chemrxiv-2023-p8qcv

Torren-Peraire, P.; Hassen, A. K.; Genheden, S.; Verhoeven, J.; Clevert, D.-A.; Preuss, M.; Tetko, I. V. Models Matter: The Impact of Single-Step Retrosynthesis on Synthesis Planning; 2023. https://arxiv.org/abs/2308.05522

Additional information

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie Actions, grant agreement No 956832.

Follow us at Twitter: https://twitter.com/AiddOne