OPCW

Scientific Advisory Board

Fortieth Session
5 – 8 May 2026

SAB-40/WP.4
2 December 2025
ENGLISH only

SUMMARY OF THE THIRD MEETING OF THE SCIENTIFIC ADVISORY BOARD'S TEMPORARY WORKING GROUP ON ARTIFICIAL INTELLIGENCE 16 – 18 SEPTEMBER 2025

1. AGENDA ITEM ONE – Opening of the meeting and adoption of the agenda

1.1 The Temporary Working Group (TWG) on Artificial Intelligence (AI) of the Scientific Advisory Board (SAB) held its third meeting from 16 to 18 September 2025 at the OPCW Centre for Chemistry and Technology. As the Chairperson of the TWG, Dr Catharina Müller-Buschbaum, was unable to attend, the meeting was chaired by the Vice-Chairperson, Prof. Hajar Mousannif.

1.2 Prof. Mousannif and the Secretary to the SAB opened the meeting by welcoming the TWG members and external speakers. The Secretary subsequently provided the Group with additional logistical and organisational information. The Vice-Chairperson then invited all participants to introduce themselves; a list of participants appears in the Annex to this report.

1.3 As no objections or comments were raised in response to the proposed programme of work for the three-day meeting, the TWG adopted the following agenda for its third meeting:

1. Opening of the meeting and adoption of the agenda
2. Overviews and updates from subgroups[1]
3. The role of AI in nuclear verification
4. The disruptive potential of new technologies on training effectiveness – case studies on the development of XR and AI solutions for CBRN training
5. Out of the woods yet? Navigating AI and automation for resilient manufacturing
6. Plenary discussions: the end-of-mandate report[2]
7. The bright and dark sides of biologically relevant chemical space
8. Beyond the screen: integrating multi-method computer-aided drug design to accelerate drug discovery

[1] While agenda items 2, 11, and 13 were discussed separately during the meeting, they are reported herein under a single agenda item for clarity and to minimise duplication.

[2] While agenda items 6, 14, and 15 were discussed separately during the meeting, they are reported herein under a single agenda item for clarity and to minimise duplication.

9. Performance of AI-generated retrosynthesis routes for precursors of newly scheduled chemicals
10. Tour of the Technology and Training Hub
11. Subgroup discussions[1]
12. Data augmentation, representations, and agents in chemistry
13. Subgroup discussions[1]
14. Plenary discussions: drafting of recommendations[2]
15. Plenary discussions: drafting of the end-of-mandate report[2]
16. Closing remarks and any other business
17. Closure of the meeting

2. AGENDA ITEMS TWO, ELEVEN, AND THIRTEEN – Discussions on subgroup topics

2.1 Subgroup 1 discussed the current landscape of retrosynthesis tools and noted the growing availability of retrosynthesis and computational chemistry solutions, as well as their potential relevance to the verification and analytical activities of the OPCW. Subgroup members highlighted the importance of maintaining an up-to-date list of vendors and platforms offering such tools, noting that the landscape evolves rapidly and includes both commercial and open-source options. In particular, open-access platforms that can be used without registration or verification may present both opportunities and challenges for monitoring chemical activities of concern.

2.2 Subgroup 1 also examined the potential use of retrosynthesis and related computational methods to enhance the capability of the OPCW to detect and assess novel synthetic pathways for chemical weapons and their precursors. Subgroup members noted that AI-enabled retrosynthesis could help identify alternative routes that bypass traditional precursors or scheduled compounds, thereby supporting the ability of the Organisation to anticipate emerging risks. Integrating such tools into analytical workflows could improve the efficiency and comprehensiveness of verification activities. Subgroup 1 noted the relevance of the case study presented by Ms Mariella Steinöcker in this context (see agenda item 9).

2.3 In addition, the subgroup discussed the implications of advances in automated synthesis and self-driving laboratories. While these technologies are rapidly progressing, members agreed that their current applications remain largely confined to research and development contexts. Economic, safety, and scalability considerations still necessitate substantial human intervention and expertise for the synthesis and handling of toxic chemicals, thereby limiting the immediate risk of misuse. Nonetheless, the subgroup stressed the importance of continued vigilance.

2.4 Subgroup 2 found the presentation by Mr Paul Schneeweiss from the International Atomic Energy Agency (IAEA) particularly relevant and discussed various aspects with him in detail. These included digitisation, organisational capacity building and staffing, techniques for open-source information processing in a secured environment, and how the IAEA leverages its chatbot. The subgroup focused on the underlying principles of these key aspects to help form recommendations. In its discussions,

Subgroup 2 also considered the importance of the order of operations in terms of applying data cleaning and structuring efforts, and welcomed input from Prof. Carolina Horta Andrade (see agenda item 8). Since the last meeting of the TWG, Subgroup 2 members conducted background research on data curation, data protection, and data reliability—areas of the terms of reference assigned to this subgroup. The research was presented and further considered during subgroup discussions to help form and agree on recommendations. These recommendations were discussed at length and later presented to the wider TWG.

2.5 Subgroup 3 proposed and discussed both short-term and long-term recommendations regarding the responsible application of AI in the chemical domain. In the short term, the group emphasised the need to evaluate AI-based interpretation systems for spectral data, enhance chemical detection through multi-sensor fusion, and assess pattern recognition performance in remote monitoring systems. Subgroup members also highlighted the importance of continuously updating AI systems in collaboration with designated laboratories, raising awareness among chemical AI developers of obligations under the Chemical Weapons Convention (the Convention), and promoting self-regulation and a safety-by-design approach through engagement with academia, industry, and publishers. For the long term, the subgroup discussed the utility in developing predictive models for unknown chemical threats and their key properties, creating detection models for variants of known chemical weapons using experimental and simulated data, and engaging with developers of generative AI to monitor for potential misuse, such as scheduled chemical generation or dataset abuse. These efforts should align with the work of emerging global AI safety institutes.

2.6 Subgroup 3 also explored how AI may impact verification efforts under the Convention. The subgroup noted that AI technologies present opportunities to enhance verification capabilities, such as through improved anomaly detection in monitoring systems, automated analysis of environmental samples, and fusion of diverse data sources for situational awareness. However, these same capabilities could be misused to evade detection, generate misleading synthetic data, or automate the design of novel toxic agents. In response, the subgroup further discussed the need to develop a Convention-relevant governance framework that incorporates transparency, auditability, and ethical design of AI systems. Subgroup members emphasised the importance of promoting explainable AI in verification contexts and highlighted the role of the OPCW in engaging AI developers, journal publishers, and private sector actors in setting technical standards. Additionally, the subgroup considered how elements of existing AI governance—such as industry self-regulation, AI safety institutes, and international collaborations—could be adapted or extended under the framework of the Convention to establish guardrails for high-risk AI applications, particularly generative models with dual-use potential.

2.7 Subgroup 4 members focused their discussions on the presentation by Mr Andrea D'Angelo (see agenda item 4) and how AI-supported training tools based on virtual reality (VR), augmented reality (AR), and mixed reality—collectively known as extended reality (XR)—could augment OPCW training capacities. While a broad spectrum of scenarios could be envisaged, the subgroup noted that the use of XR tools in inspector training for non-routine missions could be of particular interest. The subgroup noted the importance of being able to incorporate non-playable characters as they increase realism and provide additional elements to training scenarios. In addition,

any XR tools should be designed in such a way that the trainer can intervene during the training. Integration of AI will not only be useful for scenario generation, but can also play a role in automatic and very accurate evaluation, as well as in any generated after-action review.

2.8 There was agreement among the subgroup members that further work would be needed to improve realism and overcome certain hardware limitations. Data security should also be carefully considered, as the use of such technologies can lead to the creation of sensitive data. This includes scenario-related data, algorithms (including any dispersion models), and user-generated data. Subgroup members noted that cognitive bias may be introduced when developing specific training sessions with AI. While such simulation tools can be very useful, they should not completely replace real training and are best suited to augment existing training.

3. AGENDA ITEM THREE – The role of AI in nuclear verification

3.1 Mr Paul Schneeweiss, from the IAEA, introduced the work of the Agency, defined IAEA safeguards, described various AI use cases, and provided a summary of the 2025 “Emerging Technologies Workshop”.

3.2 IAEA safeguards (the main form of nuclear verification) are technical measures applied by the Agency in accordance with the terms of relevant safeguards agreements. Their purpose is to provide reassurance that nuclear material remains in peaceful use. Safeguards-relevant information that is provided by States and acquired from safeguards activities conducted by the IAEA, in addition to other relevant information, is collected and processed by the IAEA.

3.3 Within the Department of Safeguards at the IAEA, various AI projects have been implemented, with a focus on improving efficiency. Mr Schneeweiss highlighted several use cases. For example, AI is assisting in the digitisation of reports, which may be received as hard copies. Manual entry of reports is supported by an extraction tool with human-in-the-loop verification to ensure high accuracy. A similar approach is under development for the digitisation of complex site maps in diverse formats, with promising results from modern multimodal models trained on synthetic data. AI tools are already in daily use to support the review of open-source information, including transcription, translation, prioritisation, and categorisation. Additional prototypes are being developed to support the analysis of satellite imagery—focusing on detecting building footprints and infrastructure changes—and for surveillance data review, where deep learning models support the automatic flagging of safeguards-relevant events. Furthermore, a secure on-premises chatbot, ChatSG, has been built to assist staff with translation, summarisation, and information retrieval in a secured IT environment.

3.4 In concluding, Mr Schneeweiss noted that in some cases, readily available AI technologies are accessible and can assist in safeguards information analysis and nuclear verification. At the same time, many AI technologies require adaptation and fine-tuning to meet specific safeguards standards and needs. Responsible use remains the highest priority, with principles that include literacy, accountability, and above all the centrality of human decision-making, to ensure that decisions and responsibilities remain solely with human experts.

3.5 In January 2025, the Department of Safeguards held its third Emerging Technologies Workshop on AI for nuclear verification. This provided a valuable opportunity for learning and exchanges with AI experts, convening more than 300 staff and 30 external experts. The workshop examined the opportunities and challenges of AI for IAEA safeguards and enabled technical exchanges on natural language processing, computer vision, and emerging AI applications. Recommendations included tailoring solutions to safeguards needs, ensuring transparency and oversight, promoting AI literacy, and establishing governance and guidelines. The discussions at the workshop underscored the importance of cross-disciplinary engagement and preparing safeguards for an AI-empowered future.

3.6 The presentation elicited several questions from TWG members. Responding to questions on the organisational structure of AI teams and the future of AI at the IAEA, Mr Schneeweiss explained that ideas and concepts have been developed in a bottom-up manner by data scientists for management consideration, based on each division's needs. As experience has grown, the Department of Safeguards is shifting towards a more systematic, department-wide approach to considering and prioritising such initiatives, also driven by resource considerations. Mr Schneeweiss noted that AI has potential to be applied to easy tasks, increasing efficiency, rather than being used to go beyond human capabilities.

4. AGENDA ITEM FOUR – The disruptive potential of new technologies on training effectiveness – case studies on the development of XR and AI solutions for CBRN training

4.1 Mr Andrea D'Angelo from Fondazione Security and Freedom for Europe (SAFE), a non-profit organisation based in Italy that translates security innovation into operational tools, gave a presentation on his organisation's development of XR and AI solutions for chemical, biological, radiological, and nuclear (CBRN) training.

4.2 Over the past five years, SAFE has developed an XR training ecosystem that has evolved from stand-alone VR tools to scalable, immersive, multiplayer solutions. Its stand-alone VR training tool is used for portable, individual training across different scenarios, such as decontamination or an underground chemical laboratory. However, customisation is limited, and evaluation tools are simplified. Incorporating a scenario builder component has enabled the flexible, non-code creation of CBRN training environments for up to 10 participants (local or remote), supporting simultaneous training with complex scenarios for multiple people in different locations. Further developments resulted in an immersive, large-scale XR tool, which can support training for up to 20 participants in indoor and outdoor scenarios with real-time injects and weather simulation, creating dynamic systems. Other developments include XR equipment, including an XR-adapted integrated CBRN mask and an XR training vehicle, as well as an AR tabletop exercise that allows for strategic planning in simulated CBRN incidents.

4.3 Mr D'Angelo highlighted that SAFE is also researching and testing AI integration into XR training. AI is used to improve non-playable characters, such as simulated additional personnel or victims, making interactions with them more realistic and adaptive. Furthermore, models of dispersion, transport, and health effects are being explored. This allows for accurately simulating the spread of hazardous materials in the environment and modelling health effects in accordance with the period of exposure. These models will enhance the realism of XR training. AI can also drive behavioural

analysis and after-action reviews, which increases automation and helps improve training scenarios. Looking ahead, Mr D'Angelo commented on future challenges, such as data availability and biases, hardware limitations, and ethical concerns.

4.4 TWG members raised questions on the feasibility of expanding XR training with haptics and smell. The discussion further addressed how XR training compares to real-life training, the tools used for evaluation, the effects of repeated XR sessions, and issues of data confidentiality.

5. AGENDA ITEM FIVE – Out of the woods yet? Navigating AI and automation for resilient manufacturing

5.1 Dr Luke Rogers from On Demand Pharmaceuticals, United States of America, gave a presentation on his company's work to establish on-demand production and distribution of therapeutics, highlighting their automation approaches and the integration of AI technologies.

5.2 Using modular, reconfigurable manufacturing skids, On Demand Pharmaceuticals produces active pharmaceutical ingredients (APIs) on demand and at the point of care with the aim of establishing a drug manufacturing process capable of adaptively responding to emerging drug shortages. During a six-month collaborative compounding pilot project, On Demand Pharmaceuticals established its first "green node", which is a clean room facility at or near the point of care and capable of compounding and formulating medicines. Alongside this, the "blue node" supports API synthesis and the assembly of pre-measured raw material kits, called PARS-Paks (from pre-assembled raw material sequenced-packaging), which contain APIs and excipients (substances formulated together with active ingredients) for final formulation. These kits are transferred to the green node for compounding or formulation into patient-ready products. Together, these nodes form a distributed network designed to reduce vulnerability to supply chain disruptions. Dr Rogers highlighted that this compounding pilot project enabled delivery of approximately 12,000 patient-ready doses across a seven-drug formulary.

5.3 To assist in their work, staff at On Demand Pharmaceuticals use digital tools, some with AI integration, that they have either developed independently or adapted to their needs from existing products. In particular, Dr Rogers highlighted IntelliMeds, an in-house tool that compiles diverse datasets to resolve medicine manufacturing-related questions and anticipate shortages. Users interact with the database through a large language model (LLM) interface, which is ChatGPT-based, enabling complex queries, such as asking where a certain drug is synthesised and how many suppliers for it exist.

5.4 The blue and green nodes operate on a customised manufacturing execution system which is built on the commercially available Tulip platform. The system can be used to prescriptively guide users, including non-experts, through chemical synthesis or formulation and can capture inputs and outputs for full traceability of the process. Additionally, On Demand Pharmaceuticals is developing machine vision models for integration in the manufacturing execution system. These enable optical character recognition, object detection, or pose estimation to confirm that the correct materials or objects are used, and the appropriate actions are carried out in the formulation process. These tools strengthen quality oversight, reduce operator error, and expand the capabilities of non-expert staff.

5.5 Following the presentation, TWG members discussed the dual-use potential of prescriptively guided synthesis. It was noted that while synthesis becomes more approachable through the tool, the initial design and formulation still require expertise. Further questions revolved around chemical safety for new synthesis and formulations, the chemical synthesis complexity limit of the manufacturing approach, and the production cost of on-demand therapeutics.

6. AGENDA ITEMS SIX, FOURTEEN, AND FIFTEEN – Plenary discussions

6.1 The plenary discussions focused on preparing, considering, and agreeing on both the recommendations the TWG would like to propose and the approach it wants to take in drafting its end-of-mandate report. Prof. Mousannif indicated during the initial plenary session that subgroups should take time during the meeting to finalise their initial draft recommendations for consideration by the entire Group. During the plenary discussions on the last day of the meeting, each subgroup lead presented their draft recommendations. TWG members posed clarifying questions and shared their views. After discussion, the Group decided to move forward with the proposed recommendations and to remove and correct any duplications or conflicting language in the coming weeks. TWG members could then provide any additional comments they have on the revised recommendations.

6.2 It was also agreed that the Chairperson, Vice-Chairperson, and subgroup leads would meet again in person before the end of the year to draft the final report. In the interim period, each subgroup lead will work with their subgroup members to produce text in response to the questions posed in the terms of reference, as well as a refined list of draft recommendations. At the drafting meeting, the participants will merge the different texts into a cohesive document, which will then be shared for consideration by the entire TWG.

6.3 Lastly, the idea of submitting work conducted by the TWG to external journals and publishers was discussed. Many TWG members expressed interest and enthusiasm in this suggestion, and it was agreed that those interested should work with the Chairperson, Vice-Chairperson, and the Secretary to the SAB to determine the best approach.

7. AGENDA ITEM SEVEN – The bright and dark sides of the biologically relevant chemical space

7.1 Prof. José Medina-Franco, from the School of Chemistry of the National Autonomous University of Mexico, is the head of the DIFACQUIM research group, which focuses on cheminformatics and computer-aided drug design. His presentation introduced the concept of chemical space and highlighted a specific subset—biologically relevant chemical space (BioReCS)—which is central to many fields, including drug discovery, agrochemistry, natural products research, and toxicology.[3] Prof. Medina-Franco subsequently explored how this subset of chemical space contains not only beneficial drugs but also toxic compounds and potential chemical weapons.

---

[3] Medina-Franco, José L., Edgar López-López, Juan F. Avellaneda-Tamayo, and William J. Zamora. “On the Biologically Relevant Chemical Space: BioReCS.” *Frontiers in Drug Discovery* 5 (August 25, 2025). https://doi.org/10.3389/fddsv.2025.1674289.

7.2 Prof. Medina-Franco commenced by introducing his group and its research aims. These include cheminformatic analysis of chemical space, molecular modelling, property prediction, structure-activity relationships, analyses of structure-multiple activity relationships, the generation and curation of molecular databases, and the development of machine learning (ML) models.[4]

7.3 Prof. Medina-Franco then presented several definitions of chemical space provided in the literature and focused on a specific definition: M-dimensional Cartesian space in which compounds are located by a set of M physicochemical and/or cheminformatic descriptors.[5] The choice of descriptors depends on the aim of the study as well as the type and quantity of the molecules. Prof. Medina-Franco showcased various examples of approaches to visualising chemical space that complement its comprehensive and quantitative analysis. Novel approaches involve visual representations using networks of large and ultra-large chemical libraries. Highlighting the practical applications of chemical space, Prof. Medina-Franco outlined how diversity analysis and the comparison of chemical libraries could be used to analyse and characterise datasets, providing important additional information to the scientific community. In his example, he showed that the comparison of more than 50,000 compounds from 11 libraries could be effectively characterised using cheminformatics approaches.

7.4 Transitioning to the chemical multiverse, Prof. Medina-Franco described this concept as a group of numerical vectors that describe chemical space differently from the same set of molecules.[6] In other words, it is the analysis of chemical compounds using various vectors with distinct sets of descriptors, where each creates a separate chemical space. Chemical multiverses can be used to study structure-activity relationships, diversity analysis, and virtual screening, among other things. Prof. Medina-Franco noted the development of the Multiple ActivitY Analyzer (MAYA),[7] which is an open-access tool to automatically construct a chemical multiverse, and he highlighted a number of case studies.

7.5 Regarding the robust exploration of BioReCS, Prof. Medina-Franco identified key challenges, including the difficulty of using universal descriptors that can span structurally diverse chemical subspaces and ensuring true diversity in generated libraries. He called for the development of more advanced computational methodologies and hybrid workflows tailored to novel or challenging domains.

---

4 Saldivar-González, Fernanda I., Diana L. Prado-Romero, Raziel Cedillo-González, Ana L. Chávez-Hernández, Juan F. Avellaneda-Tamayo, Alejandro Gómez-García, Luis Juárez-Rivera, and José L. Medina-Franco. "A Spanish Chemoinformatics GitBook for Chemical Data Retrieval and Analysis Using Python Programming." *Journal of Chemical Education* 101, no. 6 (May 28, 2024): 2549–54. https://doi.org/10.1021/acs.jchemed.4c00041.

5 Saldívar-González, Fernanda I., and José L. Medina-Franco. "Approaches for Enhancing the Analysis of Chemical Space for Drug Discovery." *Expert Opinion on Drug Discovery* 17, no. 7 (June 5, 2022): 789–98. https://doi.org/10.1080/17460441.2022.2084608.

6 Medina-Franco, José L., Ana L. Chávez-Hernández, Edgar López-López, and Fernanda I. Saldívar-González. "Chemical Multiverse: An Expanded View of Chemical Space." *Molecular Informatics* 41, no. 11 (August 23, 2022). https://doi.org/10.1002/minf.202200116.

7 Espinoza-Castañeda, J. Israel, and José L. Medina-Franco. "Maya (Multiple ActivitY Analyzer): An Open Access Tool to Explore Structure Multiple Activity Relationships in the Chemical Universe." *ChemRxiv*, October 8, 2024. https://chemrxiv.org/engage/chemrxiv/article-details/6703066e12ff75c3a1a561b0.

7.6 Summing up, Prof. Medina-Franco underlined that BioReCS is a more realistic, application-driven framing of chemical space that emphasises both useful and undesirable molecules. A proper understanding of BioReCS is key to developing AI tools that distinguish toxic chemicals and chemical weapons from chemicals with desirable and beneficial properties. He also emphasised that growth in database size alone does not guarantee increased diversity or biological relevance, so future strategies must be thoughtful about direction, coverage, and functional meaning with safe and desirable properties.

7.7 In the question-and-answer session that followed, the availability of databases containing toxic and dangerous chemicals, which can be mapped and clustered together, was discussed at length.

8. AGENDA ITEM EIGHT – Beyond the screen: integrating multi-method computer-aided drug design to accelerate drug discovery

8.1 Prof. Carolina Horta Andrade, from the Federal University of Goiás, Brazil, provided an overview of her current work on the drug discovery process and usage of ML models to accelerate it. She outlined how computational and data-driven methods can reduce the time, cost, and uncertainty traditionally associated with developing new drugs.

8.2 Prof. Andrade depicted the current drug life cycle—from discovery through trials to obtaining approvals—as lengthy, costly, and uncertain. The entire process typically takes between 10 and 15 years and costs billions of dollars. Even then, the success rate remains less than 12%. She highlighted that the integration of innovating approaches, including AI and ML models, could shorten this timeline and reduce costs. For successful integration, the reliability and predictive capability of ML models depend on the quality, quantity, and relevance of the training data. She noted that better-curated data can yield more accurate models than complex algorithms.

8.3 Prof. Andrade and her research group are using quantitative structure-activity relationship (QSAR) models,[8] a computational method that uses statistical and ML models to predict the biological activity or properties of chemical compounds. She underscored that the quantity, quality, and relevance of data are critical for building effective predictive models. Data curation is a fundamental process in computational modelling, particularly in the field of QSAR, where it is considered a major bottleneck.[9] Careful data curation to remove errors, duplicates, and unreliable information is important for robust predictive models. Indeed, for a successful model, better data are of greater value than a more complex algorithm.

8.4 Noting the increasing applications of AI in drug discovery, Prof. Andrade highlighted four case studies from her research group. First, she covered the "OpenZika" project, which uses massive virtual screening and AI-powered QSAR models to screen over 30 million compounds to find potential drug candidates to treat the Zika virus. Second, she spoke about a multidisciplinary effort that uses molecular docking and ML models

---

8 Alves, Vinicius, Rodolpho Braga, Eugene Muratov, and Carolina Andrade. "Quimioinformática: Uma Introdução." *Química Nova*, 2017. https://doi.org/10.21577/0100-4042.20170145.

9 Fourches, Denis, Eugene Muratov, and Alexander Tropsha. "Trust, but Verify: On the Importance of Chemical Structure Curation in Cheminformatics and QSAR Modeling Research." *Journal of Chemical Information and Modeling,* 50, no. 7 (June 24, 2010): 1189–1204. https://doi.org/10.1021/ci100176x.

to identify potential anti-SARS-CoV-2 drugs by targeting key enzymes and host proteins. The third case study addressed "LabMol Insight-AI", an open-science platform that uses ML to accelerate drug discovery and enhance chemical toxicity assessments. Lastly, Prof. Andrade discussed the "Sensorial Olfactory Framework Immersive with AI (SOFIA)" project, which integrates computational chemistry, AI, and neuroscience to create emotionally adaptive olfactory VR experiences, with the aim of incorporating the sense of smell to influence memory and emotion.

8.5 Prof. Andrade concluded her presentation by further highlighting the importance of high-quality data for the drug discovery process. She underlined that progress in drug discovery depends on the combination of rigorous data practices, collaboration, and the integration of human and computational intelligence.

8.6 During the discussion, members of the TWG enquired about the laboratory's projects and the practical application of AI in drug discovery efforts. Prof. Andrade noted that academic institutions have limited access to large datasets, which slows their capacity to drive substantial progress in the drug discovery process.

9. AGENDA ITEM NINE – Performance of AI-generated retrosynthesis routes for precursors of newly scheduled chemicals

9.1 Ms Mariella Steinöcker from the Vienna University of Technology, Austria, recently completed a master's thesis in cooperation with the Austrian Armed Forces CBRN Defence Centre, which assessed the capabilities of non-State actors to produce chemical warfare agents using AI and low-tech laboratories. The goals of this work were threefold: to evaluate and compare different AI retrosynthesis tools, to experimentally test selected synthetic routes, and to provide insights that may help regulatory authorities anticipate misuse.

9.2 To conduct this work, Ms Steinöcker selected two scheduled precursors (*N*,*N*-dimethylformamidine and *N*,*N*-dimethylacetamidine) of two recently scheduled chemicals. These specific precursors were selected as they are small, synthetically accessible molecules, making them ideal for this test case. She then compared six AI retrosynthesis tools (IBM RXN, AiZynthFinder, Reaxys, Synthia, ASKCOS, and ChemAIRS) with each other and applied them to generate synthetic routes to the two precursors. These AI-based retrosynthesis tools aim to optimise efficiency and employ three main algorithmic approaches: transformer models, template-based models, and heuristic search methods.

9.3 In total, the retrosynthetic tools generated 19 routes to *N*,*N*-dimethylformamidine and 15 routes to *N*,*N*-dimethylacetamidine. The routes generated by the six tools were subsequently assessed for feasibility, safety, cost, and the requisite laboratory skills. Their synthetic complexity scores (SCScores),[10] which evaluate how complex the starting molecules are compared to the products, were also calculated. Ms Steinöcker noted that many of the proposed routes involved hazardous or toxic chemicals. She highlighted that several routes could, in principle, be carried out with only basic laboratory skills, raising security concerns.

---

10 Coley, Connor W., Luke Rogers, William H. Green, and Klavs F. Jensen. "SCScore: Synthetic Complexity Learned from a Reaction Corpus." *Journal of Chemical Information and Modeling* 58, no. 2 (January 26, 2018): 252–61. https://doi.org/10.1021/acs.jcim.7b00622.

9.4 Based on the synthetic route assessment and scoring, two routes for each precursor were selected and performed experimentally, and the products were then analysed using nuclear magnetic resonance spectroscopy. Laboratory validation yielded mixed results: some reactions produced the desired precursor product, while others failed due to a lack of reproducibility, low technical standards, or incorrect AI suggestions.

9.5 Ms Steinöcker concluded that AI-assisted retrosynthesis holds potential, but its applicability is limited by challenges such as purity issues, reproducibility concerns, and technical requirements. She noted that increasing accessibility and improvements to these tools are reducing the barriers to potential misuse. Within the scientific community, it is considered that AI will not replace chemists, but rather support them, leaving the final responsibility for evaluation and safe implementation with the human researcher.

9.6 In the following question-and-answer session, TWG members raised questions about further purification steps, personal protective equipment, and the SCScore. Further discussion revolved around the topic of the safety and accessibility of information.

10. AGENDA ITEM TEN – Tour of the Technology and Training Hub

10.1 The Head of the Technology Hub and Logistics, Mr Ardalan Zargham, provided an overview of the activities conducted at the Technology and Training Hub, which entail the systematic monitoring, testing, validation, and integration of relevant equipment and technologies for ensuring that the operational capabilities of the Organisation are iteratively expanded and optimised. Mr Zargham provided an overview of the tools being utilised to support these efforts; these tools include, but are not limited to, an AI-powered foresight tool for continuously scouting and mapping the technological landscape.

10.2 Following the overview, Mr Zargham summarised the routine activities performed by the team to provide logistical and operational support for the conduct of verification activities. The TWG was taken on a tour of the facilities, and viewed some of the important equipment that is taken on missions, including personal protective equipment, detection equipment, and medical countermeasures. The Group appreciated having an up-close look at how OPCW staff prepare for missions and the types of equipment at their disposal.

11. AGENDA ITEM ELEVEN – Subgroup discussions

See agenda item 2.

12. AGENDA ITEM TWELVE – Data augmentation, representations, and agents in chemistry

12.1 Dr Robert Pollice, from the University of Groningen, the Netherlands, provided an overview of data augmentation, representations, and agents in chemistry—three research topics that are current areas of focus at the Artificial Organic Chemistry Laboratory at the university.

12.2 Starting with data augmentation, Dr Pollice introduced large action models, which plan and perform actions based on user prompts. He then presented current advances by his research team regarding the development of a large action model capable of performing quantum chemistry simulations. The aim of this research is to create systematic approaches for

automated training data generation to fine-tune existing LLMs for chemical simulation tasks and to establish an open-source framework connecting LLMs with quantum chemistry software. Dr Pollice noted that while base models perform poorly, fine-tuning on even small datasets of computer-generated models can significantly improve performance. Work is ongoing to further enhance the accuracy of the model.

12.3 Moving to representations, Dr Pollice provided an overview of the development and application of the Simplified Molecular Input Line Entry System (SMILES),[11] highlighting that for a given molecule, there can be many possible SMILES strings. His team has been developing an algorithm capable of finding all possible equivalent SMILES identifiers for a given SMILES input. Dr Pollice noted that the SMILES randomisation algorithm implemented in RDKit, which is the standard open-source Python package for cheminformatics, does not randomise all aspects of a SMILES string. Consequently, a large fraction of alternative SMILES representations is missed when this algorithm is used. By implementing their own algorithm that accounts for all modes of randomisation, Dr Pollice and his team found that this leads to a substantial increase in the number of alternative SMILES generated for even relatively small organic molecules with rings. By combining this algorithm with a mutation of SELF-referencing Embedded Strings (SELFIES),[12] an improved SMILES randomisation algorithm has been developed, which can be used to explore molecular space more comprehensively and find new molecules that would have been missed with the RDKit randomisation algorithm.[13]

12.4 Finally, Dr Pollice explained that all common molecular string representations rely on Lewis structures.[14] However, these structures have a number of disadvantages as they do not accurately represent the electronic structure of a molecule, nor do they take non-covalent interactions or transition states into account. In light of this, Dr Pollice and his research team have developed a new string representation that is based on the quantum theory of atoms in molecules (QTAIM).[15] This approach connects the critical points of electron densities in a well-defined hierarchy, producing QTAIM graphs. These graphs can subsequently be converted into strings in exactly the same way that molecular graphs are converted into SMILES. A software package is currently being built, which will support the application of this new string representation. In concluding

---

11 Weininger, David. "SMILES, a Chemical Language and Information System. 1. Introduction to Methodology and Encoding Rules." *Journal of Chemical Information and Computer Sciences* 28, no. 1 (February 1, 1988): 31–36. https://doi.org/10.1021/ci00057a005.

12 Krenn, Mario, Florian Häse, AkshatKumar Nigam, Pascal Friederich, and Alan Aspuru-Guzik. "Self-Referencing Embedded Strings (SELFIES): A 100% Robust Molecular String Representation." *Machine Learning: Science and Technology* 1, no. 4 (October 28, 2020): 045024. https://doi.org/10.1088/2632-2153/aba947.

13 Nigam, AkshatKumar, Robert Pollice, Mario Krenn, Gabriel dos Gomes, and Alán Aspuru-Guzik. "Beyond Generative Models: Superfast Traversal, Optimization, Novelty, Exploration and Discovery (STONED) Algorithm for Molecules Using SELFIES." *Chemical Science* 12, no. 20 (2021): 7079–90. https://doi.org/10.1039/d1sc00231g.

14 Krenn, Mario, Qianxiang Ai, Senja Barthel, Nessa Carson, Angelo Frei, Nathan C. Frey, Pascal Friederich, et al. "SELFIES and the Future of Molecular String Representations." *Patterns* 3, no. 10 (October 2022): 100588. https://doi.org/10.1016/j.patter.2022.100588.

15 Shahbazian, Shant. "Why Bond Critical Points Are Not 'Bond' Critical Points." *Chemistry – A European Journal* 24, no. 21 (February 5, 2018): 5401–5. https://doi.org/10.1002/chem.201705163.

his presentation, Dr Pollice showed an example of how this approach can be applied to transition states and noted that this new string representation can successfully propose transition-state structures based on the associated reactant structures.

12.5 During the ensuing discussion, questions were raised regarding the prevention of user data accumulation and whether training data size eventually reached a point of convergence. Dr Pollice noted that models were used through application programming interfaces, which can be reinitialised each session to avoid data retention. Furthermore, no convergence has been observed, although applications such as user-guided fine-tuning with reinforced learning or similar approaches could help address this challenge in future work.

13. AGENDA ITEM THIRTEEN – Subgroup discussions

See agenda item 2.

14. AGENDA ITEM FOURTEEN – Plenary discussions: drafting of recommendations

See agenda item 6.

15. AGENDA ITEM FIFTEEN – Plenary discussions: drafting of the end-of-mandate report

See agenda item 6.

16. AGENDA ITEM SIXTEEN – Closing remarks and any other business

Prof. Mousannif commended the members on their impactful work during this meeting and throughout the mandate of the TWG. Members were encouraged to continue drafting their subgroup contributions to the end-of-mandate report. While the subgroups are drafting individual sections of this final report, it will be shared for review and agreement by consensus. Prof. Mousannif concluded the meeting by expressing appreciation to the external speakers for their valuable presentations and fresh perspectives. No additional points were raised.

17. AGENDA ITEM SEVENTEEN – Closure of the meeting

The Chairperson ended the meeting at 16:10 (CET) on 18 September 2025.

ACKNOWLEDGEMENTS

The TWG members thank the guests and members of the Secretariat who participated in discussions. The TWG also wishes to acknowledge Ms Ernesa Ademagić of the OPCW Office of Strategy and Policy for her support and contributions to the meeting and its preparations. Lastly, the TWG thanks the OPCW Director-General for his establishment and support of the TWG, and acknowledges the generous contributions of the European Union and the United States of America that help to cover the costs of the Group's work.

Annex: List of Participants at the Third Meeting of the Scientific Advisory Board's Temporary Working Group on Artificial Intelligence

Annex

LIST OF PARTICIPANTS AT THE THIRD MEETING OF THE SCIENTIFIC ADVISORY BOARD'S TEMPORARY WORKING GROUP ON ARTIFICIAL INTELLIGENCE

	Participants	**Institution**
1.	Dr Roy Forbes	University of Witwatersrand, South Africa
2.	Dr Anya Gryn'ova	University of Birmingham, United Kingdom of Great Britain and Northern Ireland
3.	Prof. Keunhong Jeong*	Korea Military Academy, Republic of Korea
4.	Prof. Anneli Kruve	Stockholm University, Sweden
5.	Dr Michael Kuiper	Google DeepMind, United Kingdom of Great Britain and Northern Ireland
6.	Mr Arthur Li	Chemical.AI, Canada
7.	Prof. Imee Su Martinez*	University of Philippines-Diliman, Philippines
8.	Prof. José L. Medina-Franco	National Autonomous University of Mexico, Mexico
9.	Prof. Hajar Mousannif	Cadi Ayyad University, Morocco
10.	Col. Günter Povoden	CBRN Defence Centre, Austrian Armed Forces, Austria
11.	Ms Molly Strausbaugh	CAS, United States of America
12.	Dr Tongning Wu	China Academy of Information and Communications Technology, China
	External Speakers	**Institution**
13.	Prof. Carolina Horta Andrade	Federal University of Goiás, Brazil
14.	Mr Andrea D'Angelo	Fondazione Security and Freedom for Europe, Italy
15.	Dr Robert Pollice	University of Groningen, Netherlands
16.	Dr Luke Rogers	On Demand Pharmaceuticals, United States of America
17.	Mr Paul Schneeweiss	International Atomic Energy Agency, International
18.	Ms Mariella Steinöcker	Vienna University of Technology, Austria
	Technical Secretariat Staff	**Division**
19.	Dr Peter Hotchkiss (Secretary to the SAB)	Office of Strategy and Policy
20.	Dr Ardalan Zargham	Inspectorate Division

* Member of the SAB.

- - - o - - -