OPCW

Scientific Advisory Board

SAB-40/WP.1
11 April 2025
ENGLISH only

SUMMARY OF THE FIRST MEETING OF THE SCIENTIFIC ADVISORY BOARD'S TEMPORARY WORKING GROUP ON ARTIFICIAL INTELLIGENCE 9 – 11 APRIL 2025

1. AGENDA ITEM ONE – Opening of the meeting

1.1 The Temporary Working Group (TWG) on Artificial Intelligence (AI) of the Scientific Advisory Board (SAB) held its first meeting from 9 to 11 April 2025. The meeting was chaired by Dr Catharina Müller-Buschbaum on behalf of the SAB, with Prof. Hajar Mousannif as Vice-Chairperson.

1.2 This first official meeting of the TWG followed an initial informal meeting that took place in January, shortly after the start of the Group's mandate. At this informal meeting, the Secretary to the SAB provided a brief overview of the mandate and work of the OPCW, and addressed the provisions related to the establishment of the SAB and related TWGs, as set out in the SAB Terms of Reference. He also commented on aspects of operation of the TWG to include points of contact, file storage and sharing, and reporting requirements. The TWG Chairperson provided an overview of the terms of reference, which are detailed in Annex 1 to this report.

1.3 Dr Müller-Buschbaum opened this first official meeting by welcoming the TWG members and expressing her pleasure and honour to be chairing this Group, with Prof. Mousannif as the Vice-Chairperson. She noted that the TWG on AI is attracting particular interest from both the Technical Secretariat (the Secretariat) and States Parties, given that AI is currently a prominent topic of global attention. Dr Müller-Buschbaum informed the Group that the initial process of determining the subgroups and their compositions was complete, and that substantive work would start during the first meeting. Prof. Mousannif also welcomed the TWG members, highlighting the increasing impact AI has on daily life. She recognised the breadth of skills and experience within the Group, which includes AI experts, chemists, and professionals working at the intersection of these two fields.

1.4 The Secretary to the SAB provided the Group with additional logistical and organisational information. He discussed the outputs of the TWG, namely summary reports for the meetings and an end-of-mandate report. As an example, the Secretary presented the end-of-mandate report from the TWG on the Analysis of Biotoxins to the Group, drawing attention to the series of recommendations and the detailed findings. He noted that the report has no fixed format or structure and can be adapted to suit the needs of the Group. The Secretary also underscored the important role external speakers play in informing the work of the Group and encouraged TWG members to identify relevant speakers for future meetings.

2. AGENDA ITEM TWO – Adoption of the agenda

As no objections or comments were raised in response to the proposed programme of work during the three days of the meeting, the TWG adopted the following agenda for its first meeting:

1. Opening of the meeting
2. Adoption of the agenda
3. *Tour de table*
4. Plenary discussion[1]
5. Multimodal and model-centric AI for chemistry: towards open and safe innovation
6. Overview of verification at the OPCW
7. Overview of inspection activities at the OPCW
8. Thinking models and safety considerations
9. Autonomous chemistry at a distance: real-time control and monitoring across borders
10. Unlocking chemistry automation requires superhuman models, not just robots
11. New catalytic strategies for chemical synthesis and biology
12. Breakout session for Subgroups 1 and 4
13. Detection and identification of toxic and high-risk chemicals with LC-HRMS[2] and machine learning
14. Power of authoritative data management and curated information
15. Breakout session for Subgroups 2 and 3
16. Plenary discussion[1]
17. Next steps, closing remarks, and any other business
18. Closure of the meeting

3. AGENDA ITEM THREE – *Tour de table*

All participants in the meeting were invited to introduce themselves to their colleagues. A list of participants appears in Annex 2 to this report.

4. AGENDA ITEMS FOUR AND SIXTEEN – Plenary discussion

4.1 In a similar approach to previous TWGs, the Chairperson and the Vice-Chairperson decided that the work of this TWG should be divided into four separate subgroups. This would ensure an efficient approach and a fair division of the workload. At future meetings of the TWG, subgroup leads will present the work of their respective subgroups, enabling everybody to gain insight and provide additional input into all four subgroups.

---

1 While agenda items 4 and 16 were discussed separately during the meeting, they are reported under agenda item 4 for clarity and to minimise duplication.

2 LC-HRMS = liquid chromatography-high resolution mass spectrometry.

4.2 An outline of the four subgroups for this TWG is as follows:

(a) Subgroup 1, led by Dr Michael Kuiper, will address questions 6(a) and 6(b) of the terms of reference, focusing specifically on synthesis and retrosynthesis prediction, and the automated and remote synthesis and production of chemicals.

(b) Subgroup 2, led by Ms Molly Strausbaugh, will address question 6(c) of the terms of reference, focusing on data curation, protection, and reliability.

(c) Subgroup 3, led by Prof. Keunhong Jeong, will address questions 6(d) and 6(e) of the terms of reference. This subgroup will focus on property, spectral, and data prediction and generation, in addition to data/sensor fusion for augmented detection and analysis.

(d) Subgroup 4, led by Col. Günter Povoden, will address question 6(f) of the terms of reference on simulation and training.

4.3 In addition to providing a summary of the current state of the art and expected near-term progress relating to their specific topic(s), each subgroup will also address the following three questions:

(a) What new capabilities are being enabled, that is, what can be done now that was not possible before? Consider both opportunities and risks.

(b) What are the current limitations and challenges to further progress, and which obstacles are likely to remain difficult or impossible to overcome?

(c) What external, non-technical factors exist that may accelerate or enable progress and/or technology adoption or slow it down?

4.4 Finally, the TWG will collectively consider six overarching questions from paragraph 8 of the terms of reference:

(a) What red flags or anomalies could help identify the potential misuse of AI systems?

(b) Which specific AI applications are sufficiently mature for the OPCW to utilise in augmenting its capabilities?

(c) What changes will be seen in industry in the coming years as AI becomes increasingly integrated into chemical production processes?

(d) How might AI impact verification efforts, either by increasing risks or by presenting opportunities?

(e) What existing guardrails and governance frameworks in the AI domain could be used, or further developed, to prevent the misuse of AI within the context of the Convention?

(f) How can the OPCW promote the responsible use of AI in relation to the Convention?

4.5 During the plenary discussion, members were invited to raise questions or reflections on the functioning of the Group and its terms of reference. The discussion included procedural aspects, such as the expected time frame for producing the meeting and end-of-mandate reports, and how the Group would manage feedback and contributions from members. Emphasis was placed on maintaining flexibility in timelines, while ensuring inclusive and transparent input, with the goal of achieving consensus on all outputs. The importance of clarity in how the TWG operates and how external contributions, such as those from external speakers or reference materials, are used in deliberations was also noted.

4.6 Time was provided for the subgroups to meet to discuss their initial ideas on their respective topics and approaches to their work. The subgroup leads then provided a brief overview of these discussions to the wider TWG during the second plenary discussion.

5. AGENDA ITEM FIVE – Multimodal and model-centric AI for chemistry: towards open and safe innovation

5.1 Dr Teodoro Laino from IBM Research Europe, Switzerland, provided an overview of recent advances in the application of AI, particularly foundation models, within the domain of chemistry. His presentation emphasised the growing importance of AI in overcoming scientific bottlenecks and accelerating discovery across chemical synthesis, spectroscopy, automation, data collection, and predictive modelling. The seminar was framed within the broader context of evolving scientific paradigms, from empirical to AI-assisted research, with a particular focus on augmenting—rather than replacing—the scientific method.

5.2 Dr Laino began by highlighting the technological evolution from expert systems and deep learning to foundation models, noting the convergence of unsupervised learning, increased computational power, and large-scale multimodal data as key enablers. He illustrated how foundation models trained on data such as chemical structures, reaction sequences, and biological affinities can perform tasks including de novo molecule design, retrosynthesis, and materials discovery, while also raising concerns around dual-use implications.

5.3 Describing the work conducted by his group, Dr Laino noted their three key focus areas: compound design, synthesis, and evaluation. In compound design, generative language models were trained to suggest novel molecular scaffolds given target properties or biological receptors. In synthesis, his group developed high-performing transformer-based models trained on reaction SMILES,[3] achieving superior accuracy in forward and retrosynthesis prediction. These models were integrated into an AI platform (named RXN for Chemistry[4]) that, by using one of the very first scientific applications of language models, can automatically program robotic synthesis hardware, demonstrating real-world deployment since 2019.[5]

---

3 Simplified Molecular Input Line Entry System.

4 For more information, see: https://rxn.res.ibm.com/rxn/robo-rxn/welcome.

5 Schwaller, P., T. Laino, T. Gaudin, P. Bolgar, C. A. Hunter, C. Bekas, and A. A. Lee. “Molecular Transformer: A Model for Uncertainty-Calibrated Chemical Reaction Prediction.” *ACS Central Science* 5, no. 9 (August 30, 2019): 1572–83. https://doi.org/10.1021/acscentsci.9b00576.

5.4 In the evaluation phase, Dr Laino presented a novel framework for structural identification using infrared (IR) spectroscopy. The approach involves pre-training models on synthetic spectra and fine-tuning them with experimental datasets, enabling accurate structural prediction from IR spectra alone. This method was further extended to integrate multiple spectroscopic modalities (such as IR, nuclear magnetic resonance (NMR) spectroscopy, and mass spectrometry (MS)) using shared embeddings to boost prediction accuracy in a multimodal foundation model. The group's findings point towards a transformative shift in AI-assisted molecular characterisation.[6]

5.5 The seminar also highlighted an innovative project that applies video transformer models and wearable video acquisition systems to annotate laboratory procedures in real time. Trained on egocentric video data from the Ego4D dataset,[7] and using a visual transformer architecture that processes video streams from cameras embedded in safety goggles, the system captures experimental workflows and quantitative measurements. This facilitates reproducibility, procedural training, and automation of laboratory records. The approach requires minimal domain-specific fine-tuning and leverages pre-training on broad human activity datasets.[8]

5.6 Dr Laino then introduced a digital twin framework for modelling the dynamic behaviour of physical systems, such as batteries. These models integrate sensor time series data, textual specifications, and visual input, enabling predictive monitoring and diagnostics. While potential application to industrial-scale chemical plants remains limited by data availability and proprietary constraints, modular digital twins for unit operations represent a promising interim solution.

5.7 Dr Laino concluded by advocating for responsible AI governance, particularly in high-impact areas like chemical synthesis. He proposed a model of coordinated disclosure and ethical review akin to cybersecurity protocols, arguing for community standards in the deployment of high-capacity foundation models.

5.8 During the ensuing question-and-answer (Q&A) session, discussions among TWG members focused on the implications of these technologies, including limitations in laboratory automation, the challenge of data heterogeneity, and the prospects of AI in identifying complex chemical mixtures.

6. AGENDA ITEM SIX – Overview of verification at the OPCW

6.1 Ms Carolyn Browne, Director of Verification at the OPCW, gave an overview of the Organisation's approach to verification activities under the Chemical Weapons Convention (the Convention). She began by describing the main areas that fall within the scope of verification: the destruction of declared chemical weapons and related facilities (Articles IV and V of the Convention), and declarations concerning chemical industry activities (Article VI). Ms Browne emphasised that the verification system is fundamentally declaration-based, relying on transparency among States Parties.

---

6 Alberts, M., T. Laino, and A. C. Vaucher. "Leveraging Infrared Spectroscopy for Automated Structure Elucidation." *Communications Chemistry* 7, no. 1 (November 16, 2024). https://doi.org/10.1038/s42004-024-01341-w.

7 See: https://ego4d-data.org/.

8 Gabrieli, G., I. Espejo Morales, D. Christofidellis, M. Graziani, A. Giovannini, F. Zipoli, A. Thakkar, A. Foncubierta, M. Manica, and P. W. Ruch. "Activity Recognition in Scientific Experimentation Using Multimodal Visual Encoding." *Digital Discovery* 4, no. 2 (2025): 393–402. https://doi.org/10.1039/d4dd00287c.

6.2 States Parties are required to submit regular declarations covering chemical weapons stockpiles and related facilities, old and abandoned chemical weapons, and relevant industrial facilities. These facilities may produce, process, or consume Schedule 1, 2, or 3 chemicals as listed in the Annex on Chemicals to the Convention, or other discrete organic chemicals above defined thresholds. In addition, States Parties report aggregate national data on scheduled chemicals, transfers of scheduled chemicals, and holdings of riot control agents. Each declaration is reviewed for internal consistency, coherence with earlier submissions, and accuracy of reported transfers—an area that often reveals discrepancies between exporter and importer data.

6.3 The inspection process is central to the verification mechanism. A small proportion of declared facilities is selected for inspection each year, either by risk-based criteria or randomised methods, depending on the facility type. Inspections assess the accuracy and completeness of the declarations, verify the absence of Schedule 1 chemicals (at non-Schedule 1 facilities), and verify the destruction of declared chemical weapons. Following an inspection, the OPCW inspection team leader compiles a factual report, which is reviewed by the Verification Division. Depending on the findings, the Secretariat may request clarifications, revised declarations, or additional action from the State Party concerned.

6.4 Ms Browne outlined several ongoing challenges, including the fact that almost one-third of States Parties have yet to implement national legislation to enforce their obligations under the Convention. This has led to significant underreporting, particularly from States Parties with sizeable chemical industries. Additionally, there are differing interpretations among States Parties concerning which chemicals or facilities should be declared, and Ms Browne cited biomediated production and highly potent active pharmaceutical ingredients (HPAPIs) as specific examples. Lastly, she noted that the shift in the global chemical industry also challenges the Convention's requirements.

6.5 Looking ahead, the Director of Verification turned to the implications of emerging technologies, especially AI and synthetic biology. Advances in synthesis techniques may reduce barriers to producing chemical warfare agents and could allow production below current declarable thresholds. This raises concerns about whether the existing schedule-based system can remain effective in mitigating proliferation risks. As a potential response, the Secretariat has begun to explore a shift towards a more risk-based approach. Efforts are also under way to investigate the use of open-source data and blockchain technology to support certain aspects of the declaration process.

6.6 In the discussion that followed, Ms Browne addressed questions on the use of AI and risk-based approaches in verification. While acknowledging AI's potential, she noted current limitations due to classified data and the lack of consensus among States Parties on how to define or apply risk in inspections. She also discussed ongoing efforts to address discrepancies in chemical transfer reporting, including a pilot project exploring blockchain technology. However, she clarified that such tools would support existing processes rather than expand the Secretariat's mandate.

6.7 Concerns were raised about underreporting by States Parties lacking national legislation. Ms Browne confirmed that this remains a significant issue that contributes to regional disparities. She concluded by emphasising that while the Convention remains operational, its long-term effectiveness depends on whether States Parties are willing to adapt it to new technological realities.

7. AGENDA ITEM SEVEN – Overview of inspection activities at the OPCW

7.1 Mr Ildefonso CamposVelarde, Director of the Inspectorate Division at the OPCW, and his colleague, Ms Tsholofelo Balci, an Inspection Team Leader, jointly delivered a presentation which provided an overview of the structure, roles, and operational responsibilities of their Division. The work of the Inspectorate Division encompasses three principal areas: routine inspections under Articles IV, V, and VI; ensuring preparedness for contingency operations; and supporting the provision of assistance to protect against the use of chemical weapons in accordance with Article X. The core business of the Division relates to inspections, and covers not only inspection planning activities, but also ensuring that inspectors are trained and equipped to conduct all types of missions. In collaboration with the International Cooperation and Assistance Division, the Inspectorate also plays a role in capacity-building efforts, including training of National Authorities and technical experts.

7.2 The Inspectorate works closely with other divisions across the Secretariat, particularly the Verification Division, and provides operational and training support across a range of mission types. Their work also intersects with demilitarisation and industry inspections. The Technology and Training Hub (TTH), based at the Centre for Chemistry and Technology, is primarily responsible for providing the OPCW inspection teams and other staff of the Secretariat with the equipment and material support required for their operations and to ensure the compliance of the inspection equipment with relevant provisions of the Convention.

7.3 Ms Balci described the inspection process and its three distinct phases: pre-mission, in-country, and post-mission. Pre-mission activities include planning, obtaining clearances for travel, and technical preparation. The inspection team prepares for each facility and coordinates with escort teams and facility personnel. In-country activities involve physical inspections, records review, and clarifications. Upon return to Headquarters, the inspection team finalises its report, holds debriefs, and ensures the secure handling of confidential materials.

7.4 Inspectors operate in a variety of environments, ranging from state-of-the-art research centres to abandoned industrial sites. Beyond inspection work, the Inspectorate is involved in training, research, administrative support, and mentoring. Training is essential due to technological change and the Secretariat’s tenure policy, which brings in new inspectors annually. Coaching and mentoring are vital for retaining institutional knowledge and ensuring continuity of expertise.

7.5 Four inspection regimes were discussed: Schedule 1, Schedule 2, Schedule 3, and other chemical production facilities. Schedule 1 chemicals pose the highest risk and are inspected every two to four years. Each inspection involves a range of tasks assigned by the Inspection Team Leader. A range of equipment is used in any given inspection, and includes laptops, personal protective equipment, and analytical tools.

7.6 During the ensuing discussion, participants raised questions about the integration of emerging technologies in inspection activities. In response, the presenters noted that while tools such as AI, drones, and satellite imaging are being explored, their use remains limited by budget constraints and confidentiality requirements. The TTH is expected to play a role in testing and deploying such technologies in the future.

7.7 Further questions addressed inspection prioritisation and data quality. The team clarified that plant site selection is based on factors such as risk profiles, product group codes, geographical distribution, and the time elapsed since the last inspection. They acknowledged challenges in maintaining consistency across inspections, particularly due to rotation of personnel, and emphasised the importance of mentoring and on-the-job training to retain institutional expertise.

7.8 Finally, TWG members also enquired about preparedness for challenge inspections and investigations of alleged use. The Inspectorate confirmed that it maintains operational readiness and conducts regular internal exercises to ensure teams remain capable of responding to such scenarios.

8. AGENDA ITEM EIGHT – Thinking models and safety considerations

8.1 Ms Joana Iljazi, from Google DeepMind in the United Kingdom of Great Britain and Northern Ireland, presented on the topic of thinking models and the safety considerations her team is addressing. Her presentation was structured around three key aspects: the capabilities of thinking models, the risks posed by their potential coupling with chemical weapons applications, and the safety measures being adopted.

8.2 Thinking models were introduced as frontier AI, representing a substantial and measurable step change in capability. These are the largest and most advanced models currently available, exhibiting doctorate-level knowledge across multiple scientific domains. In addition to synthesising information and formulating hypotheses, they are capable of reasoning, inference, and prediction. Their planning abilities allow them to decompose complex tasks into manageable subtasks and solve them using available tools and resources. These models are also multimodal, processing and producing a wide range of content types, including text, audio, video, and portable document format (PDF) files. Further strengths include long-context understanding, tool access and use, deployment within multi-agent frameworks, and applications in robotics.

8.3 This increase in capability has been observed consistently across leading frontier laboratories, including DeepMind, OpenAI, and Anthropic. Ms Iljazi highlighted that thinking models now exceed expert-level baselines in benchmarks such as graduate-level Google-proof Q&A (GPQA) and massive multitask language understanding (MMLU), and that the pace of improvement has accelerated markedly, with meaningful advancements now occurring within weeks between model releases. These performance gains extend into the scientific domain, with particular progress noted in the field of biology.

8.4 The components of thinking models were examined in detail. These include a pre-training phase with large corpora of text and a three-component post-training phase comprising supervised fine-tuning, reinforcement learning training, and preference optimisation. The models also undergo reasoning and chain-of-thought training, enhancing their ability to approach complex, multi-step problems. A distinctive feature of many such models is their “mixture of experts” architecture, wherein only a subset of parameters is activated per query. This specialisation enables more efficient processing and introduces a form of functional modularity. Inference-time compute was also discussed as a key capability, enabling dynamic allocation of computational resources depending on the complexity of the task. Additionally, search agents are employed to reformulate queries and retrieve grounded, factual answers, reducing the likelihood of hallucinations.

8.5 Particular attention was given to chain-of-thought prompting, both in its standard and iterative forms. These methods guide the model through stepwise reasoning processes, making its internal logic more transparent and interpretable. While this can improve performance, it also raises concerns when such mechanisms are used to bypass alignment safeguards. The use of these prompting techniques has also played a central role in the generation of synthetic training data—critical to the development of thinking models—by distilling high-quality model outputs into datasets suitable for fine-tuning.

8.6 Ms Iljazi then turned to the question of safety, outlining a three-step process for addressing emerging risks. This begins with understanding the threat landscape through red-teaming (an adversarial stress testing process) and threat modelling exercises. Next, evaluations are conducted against identified red lines, including adversarial testing and capability assessments. Finally, mitigations are implemented to slow or disrupt misuse. These include data filtering, safety-specific training, unlearning, and safety breaking. In addition, system-level tools such as classifiers and alert mechanisms are being explored to monitor patterns of malicious intent over time.

8.7 The potential use of thinking models in aiding the development of chemical weapons was addressed through a series of guiding questions that aim to define risk factors, identify capability thresholds, and evaluate the feasibility of certain actions. Ms Iljazi emphasised the importance of forward-looking risk evaluation, particularly in contexts where traditional safeguards may be undermined by the speed and scale of AI capabilities.

8.8 In closing, the need for coordinated, ongoing efforts across the AI safety community was underscored. TWG members were invited to engage with these efforts and contribute to the broader discourse on ensuring the safe and responsible advancement of thinking models.

8.9 Following the presentation, members of the TWG engaged in a wide-ranging discussion on the safety implications of thinking models and their potential misuse. Several members raised concerns about the ability of large language models to assist with the synthesis or procurement of chemical weapons and illicit drugs, particularly when models are used in multi-agent configurations or chained across platforms. Examples were given of jailbreaks that had been used, such as role play or creative writing prompts, to bypass alignment safeguards. Ms Iljazi and other participants noted that while recent models have improved in their ability to resist such prompts, gaps remain.

8.10 The discussion also touched on the distinction between safety and security terminology, and on whether current model evaluations are sufficient for identifying evolving threats. There was interest in the role of caching, embeddings, and inference-time compute as factors that influence model behaviour in high-risk scenarios. Finally, TWG members highlighted the need for long-term, session-spanning monitoring and international coordination on safety benchmarks, red-teaming approaches, and interface standardisation for agent-based model interactions.

9. AGENDA ITEM NINE – Autonomous chemistry at a distance: real-time control and monitoring across borders

9.1 Prof. Jason Hein from the University of British Columbia, Canada, delivered a presentation centred on the development and application of self-driving laboratories (SDLs) within chemical synthesis and analysis. He described an SDL as a closed-loop,

autonomous platform designed to integrate experimental planning, execution, and analysis.[9,10] His research group is part of the Acceleration Consortium, which aims to accelerate the discovery of materials and molecules needed for a sustainable future and enhances collaboration among international laboratories. While the group's research is not primarily focused on molecular discovery, Prof. Hein highlighted their role in automating and scaling up synthetic processes, with the broader goal of making scientific manufacturing more efficient and accessible.

9.2 Prof. Hein outlined the critical operational regimes of synthesis—design, make, and build—and explained how SDLs address each through modular and adaptable technology. His team has repurposed common tools, such as 3D printers and wireless syringes, to create low-cost, ergonomic automation systems that mimic human actions. Their platform can execute multi-step synthesis, monitor reactions, and generate hypotheses for further exploration. Prof. Hein highlighted how the "automated laboratory reactor", developed by modifying legacy equipment, enabled a first-year undergraduate student to perform a complex synthesis successfully. By providing feedback—such as indicating if the reaction addition speed is too fast or the colour is not as expected—the automated equipment significantly enhanced the student's capabilities.

9.3 Providing a variety of examples of his group's work, Prof. Hein showcased the capabilities and applicability of SDLs.[11,12,13] Notable case studies included a drug-checking system installed on campus, which allowed for anonymous, round-the-clock sample testing without involving healthcare professionals—an example of how SDLs can embed expert knowledge into accessible platforms.

9.4 The concept of adaptability recurred throughout Prof. Hein's presentation: SDLs can respond to dynamic experimental conditions and optimise workflows in real time, from fine-tuning chromatography methods, to building calibration curves. Prof. Hein described a broader digital transformation in synthesis through "chemportation": treating synthesis protocols as digital objects. His team developed a system that uses language models to translate expert procedures into executable extensible markup language (XML) code for SDLs, enabling reproducible and rapid synthesis across different laboratories.

---

9 Tom, Gary, Stefan P. Schmid, Sterling G. Baird, Yang Cao, Kourosh Darvish, Han Hao, Stanley Lo, et al. "Self-Driving Laboratories for Chemistry and Materials Science." *Chemical Reviews* 124, no. 16 (August 13, 2024): 9633–9732. https://doi.org/10.1021/acs.chemrev.4c00055.

10 Baird, Sterling G., and Taylor D. Sparks. "What Is a Minimal Working Example for a Self-Driving Laboratory?" *Matter* 5, no. 12 (December 7, 2022): 4170–78. https://doi.org/10.1016/j.matt.2022.11.007.

11 Strieth-Kalthoff, Felix, Han Hao, Vandana Rathore, Joshua Derasp, Théophile Gaudin, Nicholas H. Angello, Martin Seifrid, et al. "Delocalized, Asynchronous, Closed-Loop Discovery of Organic Laser Emitters." *Science* 384, no. 6697 (May 17, 2024). https://doi.org/10.1126/science.adk9227.

12 Zhang, Wenyu, Mason A. Guy, Jerrica Yang, Lucy Hao, Junliang Liu, Joel M. Hawkins, Jason Mustakis, Sebastien Monfette, and Jason E. Hein. "Leveraging GPT-4 to Transform Chemistry from Paper to Practice." *Digital Discovery* 3, no. 11 (2024): 2367–76. https://doi.org/10.1039/d4dd00248b.

13 Liu, Junliang, Yusuke Sato, Fan Yang, Andrew J. Kukor, and Jason E. Hein. "An Adaptive Auto-synthesizer Using Online Pat Feedback to Flexibly Perform a Multistep Reaction." *Chemistry Methods* 2, no. 8 (April 1, 2022). https://doi.org/10.1002/cmtd.202200009.

9.5 In closing his presentation, Prof. Hein highlighted the potential of SDLs to transform chemical synthesis and analysis through collaboration, digital twins, and vision-language models. By augmenting, rather than replacing, the chemist's role, SDL platforms can democratise access to advanced tools, improve reproducibility, and open new frontiers in chemical discovery.

9.6 Prof. Hein's presentation generated much discussion among the Group, exploring the integration of microfluidics with SDLs; Prof. Hein confirmed their utility for initial screening, though emphasised the significant preparatory work required. When asked about AI-driven SDL design, he acknowledged ongoing efforts and explained challenges such as robotic arm limitations and physical deck constraints, highlighting the future potential of digital twins and vision-language systems for optimal spatial planning.

9.7 Discussion of scalability and laboratory footprint led to the positioning of SDLs as specialised, modular platforms best suited for collaborative facilities, rather than ubiquitous laboratory equipment. Furthermore, Prof. Hein noted that SDLs are not to be viewed as replacements for chemists, but as tools that amplify human capabilities, particularly in routine or hard-to-access synthesis. Questions relating to safety concerns, system resilience, broader applications, and commercialisation efforts were also raised and discussed.

10. AGENDA ITEM TEN – Unlocking chemistry automation requires superhuman models, not just robots

10.1 Dr Stanisław Jastrzębski from molecule.one, Poland, gave a presentation on the automation of chemistry, with particular emphasis on the persistent challenges and the innovative strategies his company is adopting to address them. A key point of discussion was the inherent complexity of chemical processes, notably the high failure rates of reactions—estimated to range from 30% to 50%—which present a major obstacle to the development of automated workflows.

10.2 Despite widespread claims of autonomous laboratories, Dr Jastrzębski noted that the current state of automation remains limited. To illustrate this disconnect, he contrasted marketed depictions of laboratory automation with the realities of contemporary systems. Reference was made to a paper from Merck indicating a 35% failure rate in carbon–nitrogen coupling reactions, highlighting the practical difficulty of maintaining consistent success rates across multiple steps in chemical synthesis—an essential requirement for effective automation.

10.3 Molecule.one's approach targets what was described as the missing "brain" of automated chemistry. Drawing an analogy to autonomous vehicles, it was argued that the primary bottleneck lies not in the robotic platforms themselves, but in the planning algorithms that guide them. In chemistry, this translates to the need for highly accurate and precise synthesis planning models. Dr Jastrzębski underscored the need for superhuman models that can design accurate synthesis pathways.

10.4 Founded in 2019, molecule.one was the first company to integrate deep learning into commercial synthesis planning software. The firm's software, launched in partnership with CAS[14] in 2023, leverages a reaction database derived from the full breadth of published chemical literature. This comprehensive dataset enables molecule.one to offer synthesis planning software that is highly accurate and reliable.

10.5 Dr Jastrzębski highlighted the critical role of large-scale data generation in training deep learning models. He briefly presented a 60,000-reaction screen carried out over four weeks, with the reactions spanning key synthetic classes such as Suzuki couplings, backward couplings, and amide couplings. Overall, the company has generated approximately 200,000 microlitre-scale reactions to support its model development efforts.

10.6 The presentation also included an overview of a collaborative project with a major pharmaceutical company. The project involved the prediction and selection of building blocks for deoxyribonucleic acid-encoded (DNA-encoded) library synthesis. Due to the limited availability of data on DNA-encoded substrates, a workaround was implemented by first generating a large reaction dataset without DNA, followed by the application of transfer learning techniques to adapt the models. This approach enabled the successful completion of 10,000 reactions over a 12-week period.

10.7 A recurring theme throughout the presentation was the need for continued generation of reaction data to support the next generation of automated systems. This data-centric approach is considered foundational not only for enhancing automation, but also for enabling the discovery of new chemistries. The company's SpaceM1 platform, described as the first commercially available molecular space based on deep learning models trained on high-definition data, encompasses over one trillion compounds that are both inexpensive to procure and quick to synthesise.

10.8 In closing, Dr Jastrzębski emphasised that large-scale reaction data generation represents the most viable path forward in the quest to automate chemistry. By enabling more robust and predictive synthetic planning, such efforts hold the potential to unlock new realms of chemical discovery and, ultimately, to accelerate the development of novel therapeutics.

10.9 This presentation prompted a wide-ranging discussion. Questions and comments included concerns about the risk of false negatives when relying solely on liquid chromatography-mass spectrometry (LC-MS) analysis, particularly when principal products or by-products cannot be readily ionised. The Group also considered the limitations of building models on a narrow subset of known reactions, explored how transformer models predict reaction success, discussed reaction conditions discovery and optimisation, and reflected on the possible size of chemical space relevant to chemical weapons.

14 CAS is a division of the American Chemical Society.

11. AGENDA ITEM ELEVEN – New catalytic strategies for chemical synthesis and biology

11.1 Prof. Matthew Gaunt from the University of Cambridge, United Kingdom of Great Britain and Northern Ireland, opened his presentation by discussing the importance of high-throughput experimentation (HTE) in synthetic chemistry. While he acknowledged that HTE is not a universal solution, he underscored its transformative potential in accelerating reaction discovery, data generation, and methodological innovation. He noted that the ability to perform thousands of reactions is increasingly accessible, yet the subsequent data analysis remains a significant challenge.

11.2 Framing his talk within the broader ambition of predictive synthesis, Prof. Gaunt pointed to the limitations of current data quality and reproducibility in the literature, and the inability to replicate 100 years of chemistry in a short period. He argued that truly predictive chemical models remain decades away. He cautioned against the prevailing hype surrounding HTE, advocating instead for a rigorous and methodical approach to experimental design, data collection, and interpretation.

11.3 Prof. Gaunt described how his research group has collaborated with instrumentation manufacturers to adapt and repurpose industrial automation equipment for chemical experimentation. This bespoke platform enables large-scale reaction screening, but he emphasised that the true bottleneck lies in analysing the resultant data. His group has invested significant effort in developing robust calibration assays and utilising techniques such as LC-MS and NMR spectroscopy to ensure quantitative precision in data analysis.

11.4 A key theme in Prof. Gaunt's presentation was the importance of reaction diversity within HTE campaigns. He discussed the use of software tools to curate diverse chemical reaction sets, thereby increasing the scope and generalisability of the results. Operationally, he outlined workflows involving parallel liquid and solid dispensing and handling technologies, as well as technical solutions to mitigate solvent evaporation and cross-contamination in small-scale reactions.

11.5 Illustrative case studies were shared, including a copper-catalysed carbon–nitrogen bond formation reaction, highlighting the full experimental pipeline from reaction plate design to output analysis. Prof. Gaunt explored the technical challenges of miniaturisation and upscaling, emphasising the need for reproducibility when transitioning from microscale to preparative synthesis.

11.6 The presentation also addressed the application of HTE to biomolecular chemistry, particularly the development of novel modifications of nucleosides. Prof. Gaunt presented a recent example involving guanosine functionalisation to enhance oligonucleotide stability against exonuclease degradation. He noted the significant implications of such modifications for therapeutic development, particularly in improving the pharmacokinetic properties of oligonucleotide drugs.

11.7 Concluding his presentation, Prof. Gaunt advocated for deeper collaboration between synthetic chemists and data scientists. He called for the adoption of standardised data reporting formats to improve reproducibility and facilitate data integration across the scientific community. Ultimately, he stressed that the full potential of high-throughput chemistry can only be realised through a combination of experimental rigour, technological innovation, and interdisciplinary partnership.

11.8 Following Prof. Gaunt's presentation, questions from TWG members focused on two main themes: the integration of computational tools with HTE, and the challenge of standardising reaction data reporting. Prof. Gaunt acknowledged that while his group has explored limited in silico modelling, their lack of deep computational expertise has limited progress. He emphasised the value of collaboration with computational chemists and noted that while predictive models show promise, uncertainties in mechanistic understanding still hinder reliable AI-driven predictions.

11.9 The discussion also highlighted the need for consistent data reporting standards. Prof. Gaunt pointed to initiatives like the Open Reaction Database, but noted slow adoption due to fragmented journal requirements and incompatible electronic laboratory notebook systems. He suggested that translation tools already available for unstructured data might offer a more practical solution than enforcing universal standards, especially in academia where resources are limited.

12. AGENDA ITEM TWELVE – Breakout session for Subgroups 1 and 4

12.1 Subgroup 1, led by Dr Kuiper, is focused on synthesis and retrosynthesis prediction, as well as the automated and remote production of chemicals—both broad and technically complex areas. To support these discussions and ensure meaningful progress, it is the largest of the subgroups and includes several members with deep expertise in these domains.

12.2 During its breakout session, Subgroup 1 discussed these areas in detail. Members agreed that the synthesis and retrosynthesis of chemical compounds is already well supported by reliable software tools and is largely a solved problem. In contrast, autonomous synthesis—particularly via remote laboratories—remains more challenging and is not yet considered a reliable means for remotely synthesising chemical products. The subgroup discussed AutoMech—open-source software, available in GitHub—which calculates high-level ab initio thermochemical and kinetic data for chemical kinetic models. It noted the availability of open-source projects in its AutoChem module, which focuses on cheminformatics and coordinating transformations.

12.3 Subgroup 1 recommended that an AI-assisted review of the technical areas covered under question 6 of the terms of reference be conducted and regularly updated to monitor developments and highlight relevant advances. The findings from this review and subsequent updates will be shared with all TWG members at regular intervals to support knowledge exchange. Subgroup 1 also proposed evaluating selected AI tools, including Google's NotebookLM, to assess their potential for helping TWG members digest large volumes of information, as well as producing non-technical summaries for a lay audience.

12.4 Subgroup 4, led by Col. Povoden, is examining AI-supported simulation and training tools. During its breakout session, the subgroup focused on identifying existing state-of-the-art technologies in this area, particularly where AI could support tasks relevant to the OPCW.

12.5 The subgroup discussed how AI-supported training tools, such as those used for inspection tasks in complex environments, are commonly based on extended reality (XR), a term encompassing augmented reality, mixed reality, and virtual reality. Subgroup 4 noted that XR allows for the creation of "digital twin" environments in which users can interact with simulated scenarios, offering a cost-effective, safe, and flexible alternative to physical training environments. The ability to replicate industrial facilities, laboratories, or hazardous scenarios was seen as a particular advantage, especially when combined with performance assessment and evaluation features.

12.6 The potential beneficiaries of such tools could include the Inspectorate Division, non-routine missions, and interested States Parties. The subgroup also highlighted the role of the International Cooperation and Assistance Division in supporting the implementation of such training technologies. It was noted that States Parties may also benefit from exploring AI-supported tools for verification data management, provided that suitable training is available.

12.7 Subgroup 4 recognised that AI-supported training could be extended beyond field simulations to include mission preparation, where it could enhance situational awareness and decision-making in non-routine missions. The subgroup agreed that future training could strategically incorporate a broader range of emerging technologies.

12.8 Several examples of XR training tools were reviewed by the subgroup, including the results of the VERTIgO Project,[15] which provides a virtual reality training ecosystem for chemical, biological, radiological, and nuclear (CBRN) operators and is currently used by the Italian Joint NBC Defense School. It was noted that an advanced version of this tool will also be used in Project 104,[16] supporting crime scene investigation training within simulated illegal laboratories. The system includes performance evaluation and can operate in both single- and multi-user modes.

12.9 Col. Povoden informed the subgroup of an ongoing North Atlantic Treaty Organization (NATO) study, due for completion in 2025, which is investigating AI-supported dynamic chemical dispersion models and ways to enhance simulation realism. A summary of the final report will be shared with the TWG.

12.10 The subgroup also noted a tool developed by the University of Vienna and Fondazione SAFE, which adds an AI component to XR training by enabling interaction with non-player characters. These characters respond dynamically based on the trainee's tone and phrasing—such as providing incorrect answers in response to aggressive questioning—offering a novel approach to interview training in chemical incident investigations.

---

15 Virtual Enhanced Reality for inTeroperable traIning of CBRN military and civilian Operators (https://defence-industry-space.ec.europa.eu/system/files/2021-06/EDIDP2020_factsheet_SVTE_VERTIGO.pdf).

16 "Strengthening crime scene forensics and prosecution capabilities in investigating CBRN incidents in the Middle East Region", a project of the European Union CBRN Risk Mitigation Centres of Excellence.

12.11 The broader application of AI in tabletop exercises was considered, particularly for simulating chemical dispersion and the spread of misinformation or disinformation on social media. This was seen as a valuable addition to training aimed at developing rapid, effective responses.

12.12 Finally, the subgroup noted that AI-supported simulation tools could also be of interest to the chemical industry, particularly in fulfilling obligations under regulatory frameworks such as the Seveso III Directive,[17] or similar national regulations.

13. AGENDA ITEM THIRTEEN – Detection and identification of toxic and high-risk chemicals with LC-HRMS and machine learning

13.1 Prof. Anneli Kruve from Stockholm University, Sweden, opened with an overview of her research group's interdisciplinary focus, combining analytical chemistry with historical context and modelling approaches. She noted that their work on environmental monitoring and toxic compound identification has relevant links to the detection of chemical weapon markers.

13.2 One ongoing project involves collaboration with the Swedish Food Agency on an initiative to monitor drinking water production facilities for signs of contamination or unauthorised interference. Prof. Kruve noted that a key challenge lies in the complexity of environmental water matrices, especially in regions influenced by pharmaceutical manufacturing or elevated infection rates, which introduce a wide variety of pharmaceutical residues. The presence of consumer product residues and other compounds further increases this matrix complexity.

13.3 In Sweden, wastewater from water treatment plants is often discharged into surface waters, such as lakes, that may also serve as drinking water sources. Although regulations govern the removal of specific substances during the wastewater treatment and drinking water purification processes, these frameworks typically address parent compounds and may not fully account for transformation products generated through the purification or environmental processes. In certain cases, these transformation products—present in trace quantities—may exhibit similar or even greater toxicity than their corresponding parent compounds.

13.4 Detection of these compounds presents analytical difficulties, particularly those that evade treatment or form as by-products during oxidative purification processes. Prof. Kruve highlighted that advanced treatment methods such as ozonation or the introduction of reactive species can generate a multitude of unknown compounds, raising further concerns about the unintended formation of toxic substances.

---

17 Directive 2012/18/EU of the European Parliament and of the Council of 4 July 2012 on the control of major-accident hazards involving dangerous substances, amending and subsequently repealing Council Directive 96/82/EC (https://eur-lex.europa.eu/eli/dir/2012/18/oj/eng).

13.5 To address these complexities, Prof. Kruve's research group employs high-resolution mass spectrometry in tandem with separation techniques, such as liquid and gas chromatography.[18,19,20] A central aim is to maximise information yield from measurements by prioritising the most analytically and toxicologically relevant compounds. This involves methodical triaging, both pre- and post-acquisition, to optimise resource allocation during analysis.

13.6 A significant limitation in traditional identification workflows is the reliance on direct spectrum-to-compound matching, which often fails for unknown or novel substances. In response, the group is developing alternative approaches using molecular formula-based screening and fragmentation trees to identify toxic structural features, or moieties, within molecules—regardless of whether they exist in current databases.[21,22]

13.7 In the last part of her presentation, Prof. Kruve addressed the challenge of quantifying detected chemicals in the absence of reference standards. She explained that similarly sized peaks in mass spectra can represent vastly different concentrations due to variations in ionisation efficiency. To address this, her group has developed several quantification models, recently validated in an interlaboratory comparison study.[23,24] The models using ionisation efficiency predictions produced the most accurate results with fewer outliers. Finally, Prof. Kruve highlighted the limitations of de novo identification from annotated spectra, which, while promising, still suffers from low accuracy.

13.8 TWG members discussed the importance of the quality of the data used to train models, and Prof. Kruve stated that it is not possible to use any synthetic data in this quantification approach. Activity cliff analysis, matrix effects, and the presence of—and variation in—ion suppression effects were also discussed.

---

18 Szabo, Drew, Travis M. Falconer, Christine M. Fisher, Ted Heise, Allison L. Phillips, Gyorgy Vas, Antony J. Williams, and Anneli Kruve. "Online and Offline Prioritization of Chemicals of Interest in Suspect Screening and Non-Targeted Screening with High-Resolution Mass Spectrometry." *Analytical Chemistry* 96, no. 9 (February 21, 2024): 3707–16. https://doi.org/10.1021/acs.analchem.3c05705.

19 Peets, Pilleriin, Wei-Chieh Wang, Matthew MacLeod, Magnus Breitholtz, Jonathan W. Martin, and Anneli Kruve. "MS2Tox Machine Learning Tool for Predicting the Ecotoxicity of Unidentified Chemicals in Water by Nontarget LC-HRMS." *Environmental Science & Technology* 56, no. 22 (October 21, 2022): 15508–17. https://doi.org/10.1021/acs.est.2c02536.

20 Rahu, Ida, Meelis Kull, and Anneli Kruve. "Predicting the Activity of Unidentified Chemicals in Complementary Bioassays from the HRMS Data to Pinpoint Potential Endocrine Disruptors." *Journal of Chemical Information and Modeling* 64, no. 8 (March 25, 2024): 3093–3104. https://doi.org/10.1021/acs.jcim.3c02050.

21 Dührkop, Kai, Markus Fleischauer, Marcus Ludwig, Alexander A. Aksenov, Alexey V. Melnik, Marvin Meusel, Pieter C. Dorrestein, Juho Rousu, and Sebastian Böcker. "SIRIUS 4: A Rapid Tool for Turning Tandem Mass Spectra into Metabolite Structure Information." *Nature Methods* 16, no. 4 (March 18, 2019): 299–302. https://doi.org/10.1038/s41592-019-0344-8.

22 Stravs, Michael A., Kai Dührkop, Sebastian Böcker, and Nicola Zamboni. "MSNovelist: De Novo Structure Generation from Mass Spectra." *Nature Methods* 19, no. 7 (May 30, 2022): 865–70. https://doi.org/10.1038/s41592-022-01486-3.

23 Malm, Louise, Emma Palm, Amina Souihi, Merle Plassmann, Jaanus Liigand, and Anneli Kruve. "Guide to Semi-Quantitative Non-Targeted Screening Using LC/ESI/HRMS." *Molecules* 26, no. 12 (June 9, 2021): 3524. https://doi.org/10.3390/molecules26123524.

24 Malm, Louise, Jaanus Liigand, Reza Aalizadeh, Nikiforos Alygizakis, Kelsey Ng, Emil Egede Frokjær, Mulatu Yohannes Nanusha, et al. "Quantification Approaches in Non-Target LC/ESI/HRMS ANALYSIS: An Interlaboratory Comparison." *Analytical Chemistry* 96, no. 41 (October 1, 2024): 16215–26. https://doi.org/10.1021/acs.analchem.4c02902.

14. AGENDA ITEM FOURTEEN – Power of authoritative data management and curated information

14.1 Ms Molly Strausbaugh from CAS, United States of America, focused on the ethical considerations and data management practices within the context of AI applications in chemistry. She began by referencing the American Chemical Society's ethical statement on AI, stressing the necessity of peer review and legal procedures before accepting AI work for clients. Ms Strausbaugh outlined CAS's central role in collecting, harmonising, and organising chemical data—ranging from reactions and structures to properties, spectra, methods, and commercial sources—and described the CAS Registry, highlighting the challenges of chemical nomenclature.

14.2 Ms Strausbaugh underscored the importance of workforce development, focusing on the need to equip scientists with stronger technical literacy. This can be achieved through the "triangle for success" where domain, algorithm, and content experts work in partnership to harness the power of AI. Ms Strausbaugh then described the process of structuring data, noting the critical considerations, which include understanding who will be using the data and how, determining important attributes and their relationships, extracting data, and connecting data points.

14.3 The presentation also covered the use of natural language processing for extracting insights from documents and the challenges posed by translations and scientific language. Ms Strausbaugh shared how CAS manages these issues, often by pairing language experts with skilled technical writers to ensure accuracy and readability. She discussed the connection of data points, including chemical substances, regulations, literature, and formulations, and the use of lexicons to normalise terms and enhance searchability.

14.4 Ms Strausbaugh concluded by addressing the enforcement of data integrity, particularly within the context of supplier data and the automated decision-making process to reject or remove substances that do not meet CAS standards. She highlighted the importance of governance and legal oversight in ensuring ethical practices and the challenges of balancing oversight with commercial activity.

14.5 The presentation elicited several questions, principally on the role of AI in supporting chemical substitution and reformulation efforts. TWG members were interested in how AI could assist in identifying functionally equivalent substances to replace restricted or hazardous chemicals. There was discussion around the challenge of ensuring data coverage for niche or specialised substances, and the importance of data provenance when assessing alternatives. The need for curated, structured data to support substitution decisions was emphasised, particularly within the context of regulatory compliance and safety assessments. Other questions related to the challenges posed by Harmonized Tariff Schedule (HTS) codes and their updates, and chemical supply chain mapping.

15. AGENDA ITEM FIFTEEN – Breakout session for Subgroups 2 and 3

15.1 Subgroup 2, led by Ms Strausbaugh, is focusing on data curation, protection, and reliability. During its breakout session, the subgroup discussed critical aspects of data governance relevant to both the Secretariat and States Parties, with a focus on the pillars of people, process, and technology. The discussion covered several key areas, including roles and skill sets, processes, policies, knowledge management, tools, digitalisation and transformations, data analysis, and data accessibility.

15.2 The subgroup considered the roles and skill sets that are essential for effective data governance, and identified domain experts with scientific data expertise, data stewards responsible for structuring data, and algorithm experts such as data scientists, as particularly relevant. The processes discussed emphasised the importance of establishing rules for data access, considering current workflows, and identifying areas for improvement. Subgroup 2 members underscored the importance of effective policies, including the need for clear data definitions and role-appropriate training to ensure that all personnel understand the required skill levels and receive general basic AI training tailored to their daily work.

15.3 Knowledge management tools were highlighted, with a focus on AI systems and commercial off-the-shelf data governance software. The discussion also included digitalisation and transformations, such as converting hard-copy work to digital formats, securing data storage with best-in-class infrastructure, and harmonising data collection and submission procedures for declarations. The potential for automated decision-making processes was listed as a potential consideration to raise in the recommendations. The use of trained AI systems (hosted on site to ensure security) for translations, supported by human oversight, was suggested as a potential item for consideration.

15.4 During the session, the subgroup also focused on data analysis, with discussions on AI-based analysis methods, including dashboards, descriptive statistics, historical perspectives, inferential statistics (where permitted), and data visualisations. Understanding risk through data analysis was emphasised. Data accessibility was also covered, focusing on search and retrieval, protections and restrictions, maintaining classification status, and terms for data access within the Secretariat.

15.5 The session concluded with recommendations to build a graphic illustrating the interplay of people, processes, and technology. Potential speakers for future meetings were also identified.

15.6 Subgroup 3, led by Prof. Jeong, is focusing on property, spectral, and data prediction and generation, along with data and sensor fusion for augmented detection and analysis. During its breakout session, in addition to considering aspects of the overarching questions in paragraph 8 of the terms of reference and in order to address technical areas 6(d) and 6(e), Subgroup 3 organised its work into three primary areas to address the technical topics listed below:

(a) assessment of the current capabilities of AI in chemical analysis and verification;

(b) identification of opportunities and risks presented by AI technologies; and

(c) development of governance frameworks and guardrails to prevent misuse.

15.7 For the first area, the subgroup discussed the substantial advances in AI-driven property prediction, spectral analysis, and data generation that enable enhanced chemical identification. The subgroup noted that AI systems can demonstrate capabilities in predicting physico-chemical and toxicological properties, interpreting complex spectral data from various analytical techniques, and synthesising hypothetical chemical compounds with predicted behaviours. These developments have implications for both verification procedures and potential dual-use concerns.

15.8 The subgroup examined how AI integration with sensor fusion and data analysis provides augmented capabilities in chemical threat detection and monitoring. This includes multimodal data fusion from various detection technologies, real-time threat assessment capabilities, and automation of chemical analysis that reduces human error while enhancing decision-making efficiency. Members recognised the need to identify specific research examples and recent technological breakthroughs to strengthen their understanding of this topic before the second meeting of the TWG.

15.9 For the second area, the subgroup identified several new capabilities enabled by AI that present both opportunities and risks. Opportunities include enhanced detection and monitoring through AI-driven remote sensing, predictive analytics for risk mitigation, and accelerated development of countermeasures for chemical threats. Corresponding risks include dual-use dangers where AI tools for legitimate purposes could be repurposed for illicit chemical weapons development, vulnerabilities to adversarial attacks and data manipulation, and regulatory challenges as AI evolution outpaces current frameworks.

15.10 For the third area, the subgroup explored various ways in which a comprehensive approach to strengthening the role of the OPCW in AI governance could be developed. Ideas included establishing a permanent AI unit within the SAB, creating specialised monitoring mechanisms for dual-use applications, and developing ethical AI guidelines specific to chemical security. Members also considered modernising verification capabilities through AI-based tools, establishing technical assistance programmes to address capacity gaps among States Parties, developing regulatory initiatives for AI applications in chemistry, and fostering international cooperation through multi-stakeholder governance forums.

15.11 The subgroup agreed that more specific examples, recent research citations, and technological case studies would strengthen its analysis. The subgroup also discussed potential presenters for the second meeting of the TWG and noted that Prof. Jeong would present recent research on the topic related to AI and chemical verification.

16. AGENDA ITEM SIXTEEN – Plenary discussion

See agenda item four.

17. AGENDA ITEM SEVENTEEN – Next steps, closing remarks, and any other business

The TWG Chairperson and Vice-Chairperson commended the members on a very productive first meeting and invited them to work on their respective questions and meet during the intersessional period, with a view to presenting the preliminary results of their deliberations at the second meeting of the TWG, scheduled from 24 to 26 June 2025. The Secretary to the SAB recalled that any additional proposals for speakers at future meetings should be shared with him, the Chairperson, and the Vice-Chairperson as soon as possible.

18. AGENDA ITEM EIGHTEEN – Closure of the meeting

The Chairperson ended the meeting at 16:24 (CET) on 11 April 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.

Annexes:

Annex 1: Terms of Reference of the Scientific Advisory Board's Temporary Working Group on Artificial Intelligence

Annex 2: List of Participants at the First Meeting of the Scientific Advisory Board's Temporary Working Group on Artificial Intelligence

Annex 1

TERMS OF REFERENCE OF THE SCIENTIFIC ADVISORY BOARD'S TEMPORARY WORKING GROUP ON ARTIFICIAL INTELLIGENCE

1. Artificial intelligence (AI) is emerging as a powerful enabling technology that is increasingly being integrated into many other disciplines and technologies, including biotechnology, robotics, and drones. The resultant synergistic effect can significantly enhance capabilities beyond what each technology can achieve in isolation. AI is accelerating progress in chemistry and related fields, in addition to making associated processes cheaper, faster, and more effective.

2. AI, particularly in relation to risks, governance, and regulation, is continuing to capture significant attention, both nationally and regionally, in addition to coming to the fore within a range of international organisations, including the United Nations and the OPCW. In its recent comprehensive scientific report submitted to the Fifth Review Conference (RC-5/DG.1, dated 22 February 2023), the Scientific Advisory Board (SAB) identified a number of potential risks posed by the misuse of AI, including its use and integration in other technologies, as well as opportunities that this technology may afford the OPCW in its implementation of the Chemical Weapons Convention (the Convention).

3. Given the novelty of AI, its unprecedented pace of development, and its rapid inclusion in many fields, it behoves the OPCW to identify and understand the potential impacts AI might have on its mission to achieve a world free of chemical weapons, to prevent the re-emergence of chemical weapons, and to promote the peaceful uses of chemistry. Consequently, in accordance with paragraph 9 of the terms of reference of the SAB (Annex to C-II/DEC.10/Rev.1, dated 2 December 2004), the Director-General has decided to establish a Temporary Working Group (TWG) on Artificial Intelligence and has appointed Dr Catharina Müller-Buschbaum as the Chairperson of the Group.

4. Through a review of current AI capabilities and technology adoption, the objective of the TWG is to understand the impact of the technology on the object and purpose of the Convention and identify the risks to and opportunities for its implementation. The findings will be considered by the SAB and recommendations will be provided to the Director-General.

5. The TWG will consist of individuals who have expertise in AI, especially in the context of the chemical sciences. Group members may have expertise in a range of subfields of AI, including machine learning, deep learning, natural language processing, robotics, and computer vision; the application of AI in research and/or industry relating to chemical sciences or data analytics; AI ethics and governance; or experience of implementation of the Convention. The TWG will comprise qualified members of the SAB as well as representatives from the chemical industry and relevant academic and scientific organisations. Guest speakers from all geographical regions will be invited to assist the TWG in its collection of data and information, and formulation of advice.

6. The TWG will provide a summary of the current state of the art, and expected near-term progress to be made, in the following areas:

   (a) synthesis and retrosynthesis prediction;

   (b) automated and remote synthesis and production of chemicals;

   (c) data curation, protection, and reliability;

   (d) property, spectral, and data prediction and generation;

   (e) data/sensor fusion for augmented detection and analysis; and

   (f) simulation and training.

7. While considering the six technical areas set out in question 6, the TWG should ensure that the following questions are also addressed:

   (a) What new capabilities are being enabled, that is, what can be done now that was not possible before? Consider both opportunities and risks.

   (b) What are the current limitations and challenges to further progress, and which obstacles are likely to remain difficult or impossible to overcome?

   (c) What external, non-technical factors exist that may accelerate or enable progress and/or technology adoption or slow it down?

8. The TWG is also requested to consider how advances in AI will impact the implementation of the Convention and the work of the OPCW by considering the following questions:

   (a) What red flags or anomalies could help in identifying the potential misuse of AI systems?

   (b) Which specific AI applications are sufficiently mature for the OPCW to utilise in augmenting its capabilities?

   (c) What changes will be seen in industry in the coming years as AI becomes increasingly integrated into chemical production processes?

   (d) How might AI impact verification efforts, either by increasing risks or by presenting opportunities?

   (e) What existing guardrails and governance frameworks in the AI domain could be used, or further developed, to prevent the misuse of AI within the context of the Convention?

   (f) How can the OPCW promote the responsible use of AI in relation to the Convention?

9. The TWG will also highlight and consider any other application areas of AI that may be relevant within the context of this work.

10. On the basis of this in-depth review and assessment, the TWG will provide a list of recommended short- and long-term actions to ensure that AI can be harnessed for good and that its associated risks can be mitigated or, as a minimum, closely monitored.

11. The Director-General might pose additional, related questions to the TWG, through the SAB.

12. The TWG will exist for a period of one year starting on 1 January 2025. Thereafter, its work will be reviewed by the SAB and the Director-General, and a decision will be made as to whether it should continue its work and, if so, whether these terms of reference should be revised.

Annex 2

LIST OF PARTICIPANTS AT THE FIRST 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	Prof. Matthew Gaunt	University of Cambridge, United Kingdom of Great Britain and Northern Ireland
3	Prof. Jason Hein	University of British Columbia, Canada
4	Prof. Keunhong Jeong*	Korea Military Academy, Republic of Korea
5	Prof. Anneli Kruve	Stockholm University, Sweden
6	Dr Michael Kuiper	Google DeepMind, United Kingdom of Great Britain and Northern Ireland
7	Mr Arthur Li	Chemical.AI, Canada
8	Prof. José L Medina-Franco	National Autonomous University of Mexico, Mexico
9	Prof. Hajar Mousannif	Cadi Ayyad University, Morocco
10	Dr Catharina Müller-Buschbaum*	Accenture, Germany
11	Col. Günter Povoden	CBRN Defence Centre, Austrian Armed Forces, Austria
12	Ms Molly Strausbaugh	CAS, United States of America
13	Dr Tongning Wu	China Academy of Information and Communications Technology, China
	External Speakers	**Institution**
14	Ms Joana Iljazi	Google DeepMind, United Kingdom of Great Britain and Northern Ireland
15	Dr Stanisław Jastrzębski	molecule.one, Poland
16	Dr Teodoro Laino	IBM Research Europe, Switzerland
	Technical Secretariat Staff	**Division**
17	Ms Tsholofelo Balci	Inspectorate Division
18	Dr Carolyn Browne	Verification Division
19	Dr Ildefonso CamposVelarde	Inspectorate Division
20	Dr Peter Hotchkiss (Secretary to the SAB)	Office of Strategy and Policy
21	Dr Jim Ottele	Inspectorate Division

* Member of the SAB.

- - - o - - -