

Liuwen Yu

Postdoc, Trustworthy AI Group, Luxembourg Institute of Science and Technology

liuwen.yu@list.lu yuliuwen02@gmail.com [Personal website](#) [Google Scholar](#) [DBLP](#)

Last updated: July 2026

Research Profile

I am an artificial intelligence researcher with expertise in AI reasoning, focusing on symbolic and hybrid approaches to human-level and human-centered AI, with applications in the digital transformation of law and finance.

Education and Key Qualifications

- **2023.11** Ph.D. in Computer Science / Joint Doctorate in LAST-JD-RIoE-MSCA-ITN, the Law, Science and Technology Joint Doctorate: Rights of Internet of Everything, funded under the EU Marie Skłodowska-Curie Action Innovative Training Network. Dissertation: *Distributed Argumentation Technology*.
- **2018.06** M.Sc. in Philosophy, Zhejiang University, China.

Professional Appointments

- **2025.11–present** Postdoctoral Researcher, Trustworthy AI Group, Luxembourg Institute of Science and Technology, Luxembourg.
- **2023.11–2025.10** Postdoctoral Researcher, CLAiM, University of Luxembourg, Luxembourg.

Participation in Research Grants

- **2025.11–2027.11** Postdoctoral Researcher, SERAFIN: Symbolic and Explainable Regulatory AI for Finance Innovation, Fonds National de la Recherche (FNR) CORE project, Grant No. C24/19003061/SERAFIN, PI: Prof. Dr. Leon van der Torre.
- **2024.11–2025.10** Postdoctoral Researcher, DISCREASON: Formal Analysis of Discretionary Reasoning – Deontic Logic and Formal Argumentation for Modeling Discretionary Decision-Making in Legal Cases, Marie Speyer Excellence Grant project, PI: Dr. Réka Markovich.
- **2023.11–2024.10** Postdoctoral Researcher, DELIGHT: Deontic Logic for Epistemic Rights, FNR OPEN research, Grant No. O20/14776480.
- **2019.11–2023.10** Doctoral Researcher, LAST-JD-RIoE-MSCA-ITN: Law, Science and Technology Joint Doctorate – Rights of Internet of Everything, Horizon 2020 Marie Skłodowska-Curie Action Innovative Training Network, EU Grant No. 814177.

Obtained Funding

- **2025** Main co-author, RECALL: Reasoning and Explaining CAusality and Liability through Legal AI, FNR CORE, PI: Réka Markovich, Co-PI: Amro Najjar.

- **2024** Main co-author, SERAFIN: Symbolic and Explainable Regulatory AI for Finance Innovation, FNR CORE, PI: Leon van der Torre, Co-PI: Ankit Arora, project budget: 598K euros.

Awards and Distinctions

- **2025** Carole Hafner Best Paper Award at ICAIL 2025 for “Addressing the Right to Explanation and the Right to Challenge through Hybrid-AI: Symbolic Constraints over Large Language Models via Prompt Engineering”.

Academic Service

- Corner Editor, *Journal of Logic and Computation – AI Logic*.
- Editor and PC Chair, Sixth International Workshop on Logic and Argumentation for New-generation AI, Barcelona, Spain, September 17, 2026.
- Editor and PC Chair, Fifth International Workshop on Logic for New-generation AI, Luxembourg, December 1–5, 2025.
- Editor, *Advances in Explainable Agentic AI and Large Language Models*, CALM 2025 proceedings, Communications in Computer and Information Science 2923, Springer, 2026.
- Editor, *Advances in Explainability, Agents, and Large Language Models*, CALM 2024 proceedings, Communications in Computer and Information Science 2471, Springer, 2025.
- PC Chair, First and Second International Workshops on Causality, Agents and Large Models (CALM 2024 and CALM 2025).
- Program Committee member / reviewer: IJCAI, AAMAS, EXTRAAMAS, JURIX, ICAIL, and other venues.

Additional Information

Organisation of Scientific Activities

- **2026** Sixth International Workshop on Logic and Argumentation for New-generation AI, Barcelona, Spain.
- **2025** LUXLOGAI 2025 in Luxembourg: [website](#).
- **2025** LNGAI 2025 in Luxembourg: [website](#).
- **2025** CALM 2025 in Luxembourg: [website](#).
- **2024** CALM 2024 in Kyoto, Japan: [website](#).
- **2023** ICR & CLAiM Workshop on Distributed Argumentation Technology.
- **2023** Zhejiang University–University of Luxembourg Joint Lab on Advanced Intelligent Systems and REasoning (ZLAIRE) kick-off workshop.

Student Supervision

- **2025–present** Daily supervisor for Thomas Løye Skafte, doctoral researcher.
- **2024** Tutor for Velija Delikj’s bachelor semester project on autonomous cars.

- **2022–present** Advisor for Vincent de Wit, doctoral researcher.
- **2023–present** Advisor for master’s and Ph.D. students from Zhejiang University through ZLAIRE.

Teaching Profile and Activities

I enjoy both research-based teaching and foundational teaching. My research-based teaching focuses on advanced topics in knowledge representation and reasoning, including logic and formal argumentation in AI. I also contribute to foundational education by coordinating textbook drafting and teaching core subjects in computer science, such as discrete mathematics, propositional logic, and first-order logic.

- **2026.08** Course on *Logic and Argumentation for New-generation AI*, European Summer School for Logic, Language, and Information 2026 (ESSLLI 2026), Prague, upcoming.
- **2026.06** ZLAIRE joint doctoral course on *Logic and Argumentation for New-generation AI*, Zhejiang University and University of Luxembourg
- **2025.06** Course on *From Logic to Argumentation in AI*, North American Summer School in Logic, Language, and Information 2025 (NASSLI 2025), Washington University, United States.
- **2025.03** Doctoral course: *From Logic to Argumentation in AI*, University of Luxembourg, Luxembourg.
- **2024.09–2024.12** Master’s course: *Intelligent Systems: Agents and Reasoning*, University of Luxembourg, Luxembourg.
- **2024.04–2024.06** Doctoral course: *Introduction to Formal and Computational Argumentation*, University of Luxembourg, Luxembourg.
- **2024.04–2024.06** Master’s course: *Intelligent Agents 1*, University of Luxembourg, Luxembourg.

Academic Visits

- **2026.05, 2025.12, 2025.06, 2024.06, 2023.12, 2023.06** ZLAIRE, Institute of Logic and Cognition, Zhejiang University, China.
- **2024.10** Department of Information Science and Media Studies, University of Bergen, Norway.
- **2023.11** Computer Science Department, Cadi Ayyad University, Morocco.
- **2023.10** Artificial Intelligence and Data Engineering Department, Özyeğin University, Turkey.
- **2023.07** National Institute of Informatics, Tokyo, Japan.
- **2024.11, 2023.12, 2023.07** Department of Informatics, Kyoto University, Japan.
- **2022.10** Centre de Recherche en Informatique de Lens, France.

Talks

- **2026.06** Two-Layer Contract Governance: Object-Level CTD Rights and Meta-Level Contract Invalidity, at JURISIN 2026, Tokyo, Japan.

- **2026.06** Reason-Aligned Normative Reasoning in LogiKEy: From Balancing Reasons to Epistemic Rights and Duties, together with Leon van der Torre and Luca Pasetto, at the pre-JURISIN Workshop 2026, Tokyo, Japan.
- **2026.05** CREDO-RAD: Reason-Aligned Reactive Choice and the Logic of Traceable Context Dependence, together with Leon van der Torre, at the 5th Anniversary of ZLAIRE & International Workshop on Logic in Artificial Intelligence, Hangzhou, China.
- **2026.05** Balancing for Agent Decision Making through Argumentation, at the 25th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2026), Paphos, Cyprus.
- **2026.01** Where Do Weights Come From? Weighting and Weighing by LLMs and Formal Argumentation in Discretionary Judicial Reasoning, together with Davide Liga, at the AAI Bridge Program on AI and Law, Singapore.
- **2026.01** SCALE: A Framework for AI and Law Alignment, at the AAI Bridge Program on AI and Law, Singapore.
- **2025.12** Contrary-to-Duty Rights: From Hohfeld to Agreement Revision, at the 38th International Conference on Legal Knowledge and Information Systems (JURIX 2025).
- **2025.12** A Principle-Based Robustness Analysis of Labeling-Based Bipolar Argumentation Semantics, at the Luxembourg Logic & AI Summit (LUXLOGAI 2025).
- **2025.11** Reasons and Arguments in Artificial Intelligence, invited talk together with Leon van der Torre, at the international research workshop *The Weight of Reasons and Normative Reasoning*.
- **2025.06** Addressing the Right to Explanation and the Right to Challenge through Hybrid-AI: Symbolic Constraints over Large Language Models via Prompt Engineering, at the 20th International Conference on AI and Law, Chicago, United States.
- **2025.06** Artificial General Intelligence and Law, invited talk co-presented with Leon van der Torre, at the Global Artificial Intelligence Technology Conference, Forum on Law and Artificial Intelligence, Hangzhou, China.
- **2025.03** Conceptualizations and Challenges for Distributed Argumentation Technology, invited talk at the Online Seminar on Computational Models of Argument (OSCMART).
- **2025.02** Generating Explanations with Hybrid AI to Support the Right to Challenge, at the 1st LoDEX Workshop on Logical Methods for Deontic Explanations, Ruhr-Universität Bochum, Germany.
- **2024.12** Formal and Computational Argumentation in AI & Law, invited talk, Ningbo University, China.
- **2024.10** Three Conceptualizations of Formal Argumentation, invited talk at the Logic & AI Seminar, University of Bergen, Norway.
- **2024.06** Weakest Link, Prioritised Default Logic, and Principles in Argumentation, invited talk at the Fourth International Workshop on Logics for New-Generation Artificial Intelligence (LNGAI 2024), Zhejiang University, China.
- **2024.03** Distributed Argumentation Technology, invited talk at the 6th Madeira Workshop on Belief Revision, Argumentation, Ontologies, and Norms, Madeira, Portugal.
- **2023.11** Legal and Moral Reasoning Capabilities in Intelligent Machines, invited talk at the Computer Science Department, Cadi Ayyad University, Morocco.

- **2023.11** Law, Science, and Technology in Luxembourg, invited talk at the Computer Science Department, Cadi Ayyad University, Morocco.
- **2023.09** From Distributed Argumentation Technology (DAT) to Dialogue Technology (DT), at the ZLAIRE kick-off workshop, Zhejiang University, China.
- **2023.05** Principles and Practice of Formal Argumentation: Argument Strength, Acceptance, and Storage, invited talk at Tsinghua University, China.
- **2022.05** Case-Based Reasoning via Comparing the Strength Order of Features, at the 4th International Workshop on EXplainable TRAnsparent AI and Multi-Agent Systems (EXTRAAMAS 2022), online.
- **2022.01** Enhancing Trust in Trust Services: Towards an Intelligent Human-input-based Blockchain Oracle (IHIBO), at the 55th Hawaii International Conference on System Sciences (HICSS 2022), online.
- **2021.12** A Principle-Based Analysis of Abstract Agent Argumentation Semantics, invited talk at the 5th IEEE International Conference on Agents (IEEE ICA 2021), online.
- **2021.11** A Principle-Based Analysis of Abstract Agent Argumentation Semantics, at the 18th International Conference on Principles of Knowledge Representation and Reasoning (KR 2021), Italy.
- **2020.12** Interpretations of Support Among Arguments, at the 33rd International Conference on Legal Knowledge and Information Systems (JURIX 2020), online.
- **2020.09** The Principle-Based Approach to Bipolar Argumentation, at the 18th International Workshop on Non-Monotonic Reasoning (NMR 2020), online.
- **2020.07** The Principle-Based Approach to Bipolar Argumentation, at an ICR seminar at the University of Luxembourg.
- **2019.12** On the Optimized Utilization of Smart Contracts in DLTs from the Perspective of Legal Representation and Legal Reasoning, at the 32nd International Conference on Legal Knowledge and Information Systems (JURIX 2019), Madrid, Spain.

Publications

Books, Edited Volumes, and Proceedings

1. Liuwen Yu. *Textbook on Formal and Computational Argumentation*. In preparation.
2. Beishui Liao, Antonino Rotolo, Leendert van der Torre, and Liuwen Yu, editors. *Logics for New-Generation AI*. Fifth International Workshop, Luxembourg, December 1–5, 2025. College Publications, 2025.
3. Yazan Mualla, Liuwen Yu, Hui Zhao, Amro Najjar, and Davide Liga, editors. *Advances in Explainable Agentic AI and Large Language Models – Second International Workshop, CALM 2025, Luxembourg City, Luxembourg, December 1–5, 2025, Proceedings*. Communications in Computer and Information Science 2923, Springer, 2026.
4. Yazan Mualla, Liuwen Yu, Davide Liga, Igor Tchappi, and Réka Markovich, editors. *Advances in Explainability, Agents, and Large Language Models – First International Workshop on Causality, Agents and Large Models, CALM 2024, Kyoto, Japan, November 18–19, 2024, Proceedings*. Communications in Computer and Information Science 2471, Springer, 2025.

Journal Articles and Book Chapters

1. Liuwen Yu, Xu Li, and Leendert van der Torre. “Agent Defense in Abstract Argumentation: Semantics and Principle-Based Analysis.” *Argument & Computation*, 17(2):141–182, 2026.
2. Tjitze Rienstra, Leendert van der Torre, and Liuwen Yu. “Reasoning Alignment for Agentic AI: Argumentation, Belief Revision, and Dialogue.” *Journal of Applied Logics – IfCoLog Journal*, 12(6):1683–1712, 2025.
3. Liuwen Yu, Leendert van der Torre, and Réka Markovich. “Thirteen Challenges in Formal and Computational Argumentation.” *Journal of Applied Logics – IfCoLog Journal*, 12(3):652–746, 2025. Also appears in *Handbook of Formal Argumentation*, Vol. 3, College Publications, 2024.
4. Pere Pardo, Liuwen Yu, Chen Chen, and Leendert van der Torre. “Weakest Link, Prioritised Default Logic and Principles in Argumentation.” *Journal of Logic and Computation*, 35(4), 2025, exaf007.

Conference and Workshop Papers

1. Luca Pasetto, Thomas Skate, Leendert van der Torre, and Liuwen Yu. “The Libra Dialogue Protocol for Comparative Deliberation.” In: *11th International Conference on Computational Models of Argument (COMMA 2026)*. Upcoming.
2. Xu Li, Leendert van der Torre, and Liuwen Yu. “Revealed Epistemic Trust.” In: *Proceedings of the 23rd International Conference on Principles of Knowledge Representation and Reasoning (KR 2026)*. Upcoming.
3. Davide Liga, Liuwen Yu, and Ankit Arora. “Retrieval-Augmented Generation for Financial Regulation via LLM Orchestration and Hybrid Symbolic-Subsymbolic AI.” In: *Proceedings of the 21st International Conference on Artificial Intelligence and Law (ICAIL 2026)*. Upcoming.
4. Liuwen Yu, Réka Markovich, Nico Potyka, and Marco Cuccarini. “Modeling Judicial Discretionary Decision Making with Higher-Order Quantitative Bipolar Argumentation.” In: *Proceedings of the 21st International Conference on Artificial Intelligence and Law (ICAIL 2026)*. Forthcoming.
5. Marija Slavkovik, Liuwen Yu, Leendert van der Torre, and Réka Markovich. “Whistleblowing and the Machine – Towards a Considered Position.” In: *AAMAS-26 Workshop on Rebellion and Disobedience in AI (RAD-AI 2026)*. Preprint, arXiv:2606.21201.
6. Huimin Dong, Réka Markovich, Leendert van der Torre, and Liuwen Yu. “Hierarchical Institutions for Contract Invalidity.” In: *Local Proceedings of the 20th International Workshop on Juris-Informatics (JURISIN 2026)*. Published.
7. Davide Liga, Réka Markovich, and Liuwen Yu. “Where Do Weights Come From? Weighting and Weighing by LLMs and Formal Argumentation in Discretionary Judicial Reasoning.” In: *Logical and Symbolic Reasoning in Language Models @ AAI 2026*. Published.
8. Liuwen Yu, Leendert van der Torre, Réka Markovich, and Chenyang Cai. “Balancing for Agent Decision Making through Argumentation.” In: *Proceedings of the 25th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2026)*, pp. 2876–2885. Published.
9. Marija Slavkovik, Liuwen Yu, Leendert van der Torre, Réka Markovich, and Beishui Liao. “Disobedience in Normative Multi-Agent Systems.” In: *Proceedings of the 25th International*

Conference on Autonomous Agents and Multiagent Systems (AAMAS 2026), pp. 2438–2446. Published.

10. Xu Li, Leendert van der Torre, and Liuwen Yu. “A Logical Analysis of an Information Filtering Architecture Based on Epistemic Trust Inference.” In: *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI 2026)*, pp. 19259–19266.
11. Huimin Dong, Leendert van der Torre, and Liuwen Yu. “Contrary-to-Duty Rights: From Hohfeld to Agreement Revision.” In: *Proceedings of the 38th International Conference on Legal Knowledge and Information Systems (JURIX 2025)*, pp. 61–73.
12. Davide Liga and Liuwen Yu. “Which Neurons Nudge Normative Stance? Causal Tests and Mechanistic Evidence via Contrastive Last-Token Steering.” In: *Proceedings of the 38th International Conference on Legal Knowledge and Information Systems (JURIX 2025)*, pp. 110–120.
13. Liuwen Yu, Leendert van der Torre, Réka Markovich, Beishui Liao, and Chenyang Cai. “DiSCo–RAD: Reasoning Alignment for Judicial Discretion.” In: *Proceedings of the Fifth International Workshop on Logics for New-Generation Artificial Intelligence (LNGAI 2025)*, College Publications, 2025, pp. 86–105.
14. Caren Al Anaissy, Chen Chen, Srdjan Vesic, Leendert van der Torre, and Liuwen Yu. “A Principle-Based Robustness Analysis of Labeling-Based Bipolar Argumentation Semantics.” In: *Proceedings of the Fifth International Workshop on Logics for New-Generation Artificial Intelligence (LNGAI 2025)*, College Publications, 2025, pp. 35–50.
15. Liuwen Yu, Davide Liga, and Réka Markovich. “Addressing the Right to Explanation and the Right to Challenge through Hybrid-AI: Symbolic Constraints over Large Language Models via Prompt Engineering.” In: *Proceedings of the 20th International Conference on Artificial Intelligence and Law (ICAIL 2025)*, ACM, pp. 239–247. Carole Hafner Best Paper Award.
16. Liuwen Yu and Leendert van der Torre. “The A-BDI Metamodel for Human-Level AI: Argumentation as Balancing, Dialogue and Inference.” In: *Proceedings of the 6th International Conference on Logic and Argumentation (CLAR 2025)*, pp. 361–379.
17. Xu Li, Liuwen Yu, and Leendert van der Torre. “Rational Monotony in Input/Output Logic.” In: *Proceedings of the 17th International Conference on Deontic Logic and Normative Systems (DEON 2025)*, pp. 255–272.
18. Vincent de Wit, Liuwen Yu, Amro Najjar, and Réka Markovich. “Balancing Normative Reasons for the Intelligent Human-Input-Based Blockchain Oracle (IHiBO).” In: *Advances in Explainability, Agents, and Large Language Models*, Communications in Computer and Information Science 2471, Springer, 2025, pp. 33–47.
19. Benoît Alcaraz, Aria Nourbakhsh, and Liuwen Yu. “Assessing the Robustness of LLMs in Predicting Supports and Attacks.” In: *Advances in Explainability, Agents, and Large Language Models*, Communications in Computer and Information Science 2471, Springer, 2025, pp. 88–93.
20. Aleks Knoks, Muyun Shao, Leendert van der Torre, Vincent de Wit, and Liuwen Yu. “A Principle-Based Analysis for Numerical Balancing.” In: *Proceedings of the Fourth International Workshop on Logics for New-Generation Artificial Intelligence (LNGAI 2024)*, College Publications, 2024, pp. 33–46.
21. Liuwen Yu, Caren Al Anaissy, Srdjan Vesic, Xu Li, and Leendert van der Torre. “A Principle-Based Analysis of Bipolar Argumentation Semantics.” In: *Proceedings of the 18th European*

- Conference on Logics in Artificial Intelligence (JELIA 2023)*, Lecture Notes in Computer Science 14281, Springer, pp. 209–224.
22. Chen Chen, Pere Pardo, Leendert van der Torre, and Liuwen Yu. “Weakest Link in Formal Argumentation: Lookahead and Principle-Based Analysis.” In: *Proceedings of the 5th International Conference on Logic and Argumentation (CLAR 2023)*, Lecture Notes in Computer Science 14156, Springer, pp. 61–83.
 23. Liuwen Yu, Caren Al Anaissy, Srdjan Vesic, Xu Li, and Leendert van der Torre. “Exploration des sémantiques d’argumentation bipolaire : une analyse basée sur les principes.” In: *Journées d’Intelligence Artificielle Fondamentale (JIAF 2023)*, pp. 39–41.
 24. Pere Pardo, Leendert van der Torre, and Liuwen Yu. “Advanced Intelligent Systems and Reasoning: Standardization, Experimentation, Explanation.” In: *Joint Proceedings of the Third International Workshop on Logics for New-Generation Artificial Intelligence and the International Workshop on Logic, AI and Law (LNGAI 2023)*, College Publications, pp. 3–20.
 25. Liuwen Yu and Dov M. Gabbay. “Case-Based Reasoning via Comparing the Strength Order of Features.” In: *Proceedings of the 4th International Workshop on Explainable, Transparent Autonomous Agents and Multi-Agent Systems (EXTRAAMAS 2022)*, Lecture Notes in Computer Science 13283, Springer, pp. 143–151.
 26. Liuwen Yu, Mirko Zichichi, Réka Markovich, and Amro Najjar. “Intelligent Human-Input-Based Blockchain Oracle (IHIBO).” In: *Proceedings of the 14th International Conference on Agents and Artificial Intelligence (ICAART 2022)*, pp. 515–526.
 27. Liuwen Yu, Mirko Zichichi, Réka Markovich, and Amro Najjar. “Enhancing Trust in Trust Services: Towards an Intelligent Human-Input-Based Blockchain Oracle (IHIBO).” In: *Proceedings of the 55th Hawaii International Conference on System Sciences (HICSS 2022)*, pp. 1–10.
 28. Liuwen Yu, Mirko Zichichi, Réka Markovich, and Amro Najjar. “Argumentation in Trust Services within a Blockchain Environment.” In: *Proceedings of the 33rd Benelux Conference on Artificial Intelligence and the 30th Belgian Dutch Conference on Machine Learning (BNAIC/BENELEARN 2021)*, pp. 418–438.
 29. Liuwen Yu, Dongheng Chen, Lisha Qiao, Yiqi Shen, and Leendert van der Torre. “A Principle-Based Analysis of Abstract Agent Argumentation Semantics.” In: *Proceedings of the 18th International Conference on Principles of Knowledge Representation and Reasoning (KR 2021)*, pp. 629–640.
 30. Lisha Qiao, Yiqi Shen, Liuwen Yu, Beishui Liao, and Leendert van der Torre. “Arguing Coalitions in Abstract Argumentation.” In: *Logics for New-Generation AI 2021*, College Publications, pp. 93–106.
 31. Liuwen Yu, Réka Markovich, and Leendert van der Torre. “Interpretations of Support Among Arguments.” In: *Legal Knowledge and Information Systems (JURIX 2020)*, IOS Press, pp. 194–203.
 32. Liuwen Yu and Leendert van der Torre. “A Principle-Based Approach to Bipolar Argumentation.” In: *NMR 2020 Workshop Notes*, Vol. 227, 2020.
 33. Liuwen Yu. “On the Optimized Utilization of Smart Contracts in DLTs from the Perspective of Legal Representation and Legal Reasoning.” In: *Proceedings of the JURIX 2019 Doctoral Consortium*, 2019.