

CALL FOR PAPERS TIME 2027

3rd International Workshop on Transformative Insights in Multi-faceted Evaluation

Co-located with **The Web Conference 2027 (WWW 2027)**

Dublin, Ireland | May 10 – 14, 2027

Lei Wang, Md Zakir Hossain, Zicheng Pan, Wenshuo Chen, Syed Mohammed Shamsul Islam, Tom Gedeon, Sabrina Caldwell, Rafiqul Islam, Mohammad A. Razzaque, and Arjumand Younus

The central question: How should we evaluate Web AI systems when capability alone is not enough? TIME 2027 invites work that moves beyond headline scores to examine how modern AI systems perform, adapt, fail, and affect people in realistic settings.

Important Dates

Workshop paper submission deadline	January 4, 2027 (AoE)
Workshop paper notification	February 1, 2027 (AoE)
Workshop paper camera-ready deadline	February 16, 2027 (AoE)
The Web Conference 2027	May 10–14, 2027, Dublin, Ireland

All deadlines are in the Anywhere on Earth (AoE) time zone. The submission portal will open soon; access details will be announced separately.

Call for Papers

Foundation models, generative and multimodal AI, retrieval-augmented generation (RAG), and AI agents are rapidly reshaping search, recommendation, knowledge access, and other Web applications. Yet their performance cannot be captured by task accuracy alone. Systems may produce convincing but incorrect answers, fail when information changes, misuse tools, behave inconsistently across users or contexts, or introduce safety, privacy, and fairness risks.

The **3rd International Workshop on Transformative Insights in Multi-faceted Evaluation (TIME 2027)** brings together researchers and practitioners to advance rigorous, responsible, and actionable evaluation of these emerging Web AI systems. Building on the **TIME 2025** [Wang et al.(2025)] and **TIME 2026** [Wang et al.(2026)] workshops, this edition places particular emphasis on foundation models, multimodal and generative systems, RAG, and agentic Web applications.

We seek to move evaluation beyond static benchmarks and single headline metrics toward evidence about how systems behave in realistic, interactive, and changing environments. We welcome work that examines not only what AI systems can do, but also how reliably, robustly, safely, fairly, transparently, and efficiently they do it, and whether their outcomes deliver meaningful real-world utility. Contributions that offer reproducible methods, reveal important failure modes, or translate evaluation evidence into practical recommendations are especially encouraged.

TIME 2027 welcomes perspectives from Web research, information retrieval, machine learning, multimodal learning, responsible AI, security, privacy, and human-centred AI. Our goal is to build an interdisciplinary community around a shared challenge: making the evaluation of modern Web AI more meaningful, comprehensive, and useful in practice.

Topics of Interest

Topics include, but are not limited to, the following areas:

1. Foundation Models and Generative AI

- Capability assessment across tasks, domains, and user needs; factuality, hallucination, and evidence-grounded generation.
- Robustness, generalisation, failure analysis, and evaluation under distribution shift.

- Multi-dimensional assessment beyond aggregate scores and conventional task metrics.

2. Multimodal AI and Web Systems

- Evaluation of systems that integrate text, images, video, audio, and other modalities.
- Cross-modal consistency, alignment, reasoning, and generation quality.
- Performance and reliability under heterogeneous inputs and changing Web content.

3. Retrieval-Augmented Systems and AI Agents

- RAG evaluation, information grounding, retrieval quality, and source attribution.
- Agent planning, tool use, multi-step execution, error recovery, and long-horizon tasks.
- Behaviour in interactive Web environments with incomplete, conflicting, or evolving information.

4. Trustworthy, Responsible, and Human-Centred Evaluation

- Safety, security, privacy, fairness, bias, and societal impact.
- Interpretability, transparency, uncertainty quantification, and calibration.
- Human-centred evaluation, usability, real-world utility, and cross-cultural or cross-domain differences.

5. Benchmarks, Metrics, and Reproducible Methodologies

- Evaluation frameworks, metrics, protocols, test suites, and benchmark design.
- Dataset quality, annotation reliability, validity, reproducibility, and statistical rigour.
- Open evaluation tools, transparent reporting, and cross-domain assessment.

6. Efficiency, Deployment, and Emerging Challenges

- Computational cost, latency, scalability, and sustainability.
- Evaluation, auditing, and monitoring in real-world deployment.
- Trade-offs among capability, reliability, safety, cost, and user utility; industry lessons, negative results, and overlooked failure modes.

What Contributions Are We Looking For?

We encourage a broad range of submissions that deepen understanding of AI evaluation, including:

- **Research papers** presenting novel evaluation methods, empirical studies, or systematic investigations.
- **Surveys and systematic reviews** synthesising progress, exposing gaps, and identifying future directions.
- **Benchmark and metric analyses** examining the strengths, limitations, and validity of current evaluation practices.
- **Evaluation frameworks and tools** enabling practical, transparent, and reproducible assessment.
- **Negative results and failure analyses** revealing weaknesses, risks, and failure modes that conventional evaluation may overlook.
- **Position papers and industry case studies** proposing research priorities or sharing lessons from real-world use and deployment.

We particularly value contributions that go beyond reporting a performance gain to explain what an evaluation reveals, why it matters, and how it can inform better research or deployment decisions. Papers introducing a model or application are welcome when they include substantial, systematic evaluation relevant to the workshop.

Submission Requirements

TIME 2027 will use [OpenReview](#) for paper submissions and the review process. The submission portal will open soon, and the direct venue link will be announced separately. All papers must be written in English and use the [ACM conference double-column template](#). Submissions will be evaluated through a **double-blind peer-review process**; authors must anonymise their papers and follow the detailed instructions provided when the portal opens.

- **Long papers:** up to 8 pages of paper content, plus references and an optional appendix, with a maximum total length of 12 pages.
- **Short papers:** up to 4 pages in total, including references and any appendix.
- **Review:** double-blind expert peer review; see below for the review stages.
- **Publication:** accepted papers will be included in the *WWW 2027 Companion Proceedings*, subject to compliance with the camera-ready deadline and applicable ACM publication requirements.

All submissions should present original work, comply with relevant research-ethics requirements and applicable ACM policies, and clearly explain the evaluation question, methodology, findings, and significance to the TIME 2027 scope. Papers will be assessed for relevance, technical quality, clarity, methodological rigour, and the value of the evaluation insights they offer. Accepted contributions may be presented orally or as posters, as appropriate.

Review Process and Publication

Building on the process used for TIME 2025 and TIME 2026, TIME 2027 will use a multi-stage, double-blind review process designed to support rigorous, constructive, and fair evaluation.

1. **Initial screening.** The Program Chairs will check submissions for completeness and compliance with the required format and page limits. Submissions passing this screening will undergo automated checks for AI-generated content and potential reference hallucinations.
2. **Expert review and author response.** Submissions passing the initial checks will be reviewed by at least three expert reviewers, including at least one reviewer from industry. Preliminary review results will be shared with authors, including reviewer comments, AI-checking reports, and Program Chair feedback on formatting. Authors will have an opportunity to respond to the feedback and revise their submissions.
3. **Area Chair assessment and final decisions.** Area Chairs will assess revised submissions by considering reviewer comments and ratings, the paper's strengths and contributions, its relevance to the workshop, and the authors' responses. The Program Chairs will make the final decisions based on the reviews, Area Chair recommendations, overall paper quality, and fit with TIME 2027.

Paper awards. High-quality papers recommended by the Area Chairs will be considered by the Award Committee for the workshop's Best Paper and other paper awards.

Publication. Accepted papers will be included in the *WWW 2027 Companion Proceedings*, subject to meeting the camera-ready timeline and applicable ACM publication requirements. For examples of previous editions, see the ACM Digital Library collections for [TIME 2025](#) and [TIME 2026](#).

Why TIME at The Web Conference?

The Web is both a major setting for modern AI and a demanding environment in which to evaluate it. Information changes continuously, users have diverse needs, evidence may be incomplete or contradictory, and AI systems increasingly interact with external tools and carry out multi-step tasks. These conditions expose limitations that static benchmarks and isolated scores can miss.

TIME 2027 complements The Web Conference by bringing together communities that study AI capability, Web-scale information, trustworthy systems, and human experience. Through research presentations and interdisciplinary discussion, the workshop aims to identify shared evaluation challenges, strengthen practical evaluation methods, and help shape a more reliable, responsible, and human-centred future for Web AI.

We invite the community to help rethink how modern Web AI should be evaluated.

References

- [Wang et al.(2026)] Lei Wang, Md Zakir Hossain, Tom Gedeon, Zicheng Pan, Syed Islam, Rafiqul Islam, Yasmeen George, and Shreya Ghosh. 2026. TIME 2026: 2nd International Workshop on Transformative Insights in Multi-faceted Evaluation. In *Companion Proceedings of the ACM Web Conference 2026*. 1199–1202.
- [Wang et al.(2025)] Lei Wang, Md Zakir Hossain, Syed Islam, Tom Gedeon, Sharifa Alghowinem, Isabella Yu, Serena Bono, Xuanying Zhu, Gennie Nguyen, Nur Al Hasan Haldar, et al. 2025. TIME 2025: 1st international workshop on transformative insights in multi-faceted evaluation. In *Companion Proceedings of the ACM on Web Conference 2025*. 1848–1851.