

Peer Review

Review of: "Comprehensive Classification of Web Tracking Systems: Technological Insights and Analysis"

Dr. Soumi Ghosh¹

1. Information Technology, Guru Gobind Singh Indraprastha University, Delhi, India

Based on the content and structure observed in the manuscript "Comprehensive Classification of Web Tracking Systems: Technological Insights and Analysis", here are the detailed evaluations for each criterion, with ratings out of 10:

Title and Abstract: Accuracy of title reflection: 8/10

- The title clearly indicates that the paper offers a comprehensive classification of web tracking systems, aligning well with its content.

Abstract clarity and informativeness: 9/10

- The abstract succinctly summarizes the research focus, methodology (classification scheme and survey), and key findings regarding user awareness and technological mechanisms. Slightly more emphasis on primary results could enhance clarity.

Introduction: Clarity of research question/problem: 8/10

- The introduction articulates the motivation for studying WT systems from a technological perspective, though a more explicit research question or hypothesis could improve clarity.

Adequacy of background and context: 9/10

- Provides good context on the evolution of WT systems, their importance, and the gaps in existing literature, setting a solid foundation.

Objectives and significance: 8/10

- Clearly states that the goal is to classify WT systems and explore user awareness, emphasizing the novelty of a technical classification scheme.

Methodology: Appropriateness of methods: 8/10

- The combination of a technical classification framework and a user survey is suitable for the study's aims.

Data collection and analysis validity: 7/10

- The survey design is well-explained, targeting a broad online audience with simple questions. However, potential biases (e.g., self-selection, limited demographic scope to Greece) are acknowledged as limitations. The classification scheme is based on observable server-side criteria, supporting validity.

Limitations: 7/10

- The paper recognises sampling bias and the need for more advanced statistical models (e.g., regression) for future work.

Results: Clarity and logic: 8/10

- The presentation of user awareness levels, familiarity with WT tools, and their relationship with background factors is straightforward.

Support for claims: 8/10

- Results align with the conclusions about user awareness being basic and independent of concerns, supported by survey data.

Visuals and tables: 7/10

- Tables and figures are referenced and organised but (based on available pages) could be enhanced for clarity with more detailed labels or graphical summaries.

Discussion: Interpretation and implications: 8/10

- Effectively discusses the findings' significance, especially regarding user awareness gaps and the need for technical education.

Link to research objectives: 8/10

- Connects well with the objectives of classifying WT systems and understanding user perceptions.

Explanation of significance: 7/10

- Highlights the importance of understanding WT mechanisms for regulation and user education but could elaborate more on practical implications.

References: Accuracy and completeness: 9/10

- The references cited seem accurate and relevant, including recent studies on WT mechanisms and user privacy.

In-text citation alignment: 9/10

- Properly cited, with consistent formatting and relevant source representation.

Clarity and Writing: Overall clarity: 8/10

- The manuscript is clearly written and generally free of grammatical errors. Some sections could benefit from more concise language.

Organization and flow: 8/10

- Logical structure; sections flow well. Minor improvements in transitions might enhance readability.

Suggestions for improvement:

- Clarify research questions explicitly in the introduction.
- Use more visual aids to consolidate survey findings.
- Address potential bias more thoroughly in methodology discussions.

Overall Impression: Quality and contribution: 8/10

- The paper offers a valuable, detailed technical classification of WT systems and highlights user awareness issues, filling a gap in the literature.

Recommendation:

- **Accept with minor revisions** — primarily to improve clarity, explicitly state research questions, and enhance the visual presentation of survey data.

Additional Comments:

- The study is timely given increasing concerns over user privacy and tracking. Incorporating more statistical rigour in analysing survey data (e.g., regression modelling) would strengthen the findings. Future iterations could also broaden the demographic scope beyond Greece to improve generalisability.

Assessment of the strengths and weaknesses of the article:

Strengths:

Novelty and Technical Depth: The paper presents a comprehensive, technically grounded classification scheme for web tracking systems based on their architecture and operational mechanisms, filling a notable gap in the literature.

Clear Focus on Technology: By emphasising the technological foundations rather than organisational or privacy aspects, the study provides valuable insights into how tracking systems function, which is often underexplored.

Integration of User Awareness Study: Including a survey of over 1,000 users offers a practical perspective on public understanding, highlighting important gaps in awareness that can inform policy and education efforts.

Methodological Transparency: The paper clearly details the classification criteria and survey methodology, allowing for reproducibility and critical assessment.

Relevance to Contemporary Issues: The research addresses current concerns about online privacy, user tracking, and regulations like GDPR, making it timely and impactful.

Weaknesses:

Limited Generalisability of Survey Data: The survey was conducted exclusively with Greek participants via online sharing, which may introduce bias and limit applicability to broader populations.

Lack of Advanced Statistical Analysis: The analysis relies primarily on descriptive statistics, with suggestions for future work to include more sophisticated methods like regression analysis, which would

strengthen the validity of conclusions.

Insufficient Explicit Research Questions: While the objectives are implied, explicitly stating research questions or hypotheses at the outset would improve clarity and focus.

Potential Oversimplification in Classification: The classification scheme, though technically detailed, might not fully capture emerging or hybrid tracking techniques, possibly limiting its comprehensiveness.

Visual and Structural Improvements Needed: Some tables and figures could be more detailed or clearer to better communicate complex findings.

Limited Discussion of Ethical and Privacy Implications: The paper emphasises the technical aspects but does not deeply explore the ethical, legal, or societal impacts of the tracking mechanisms, which are crucial to the topic.

In Summary: The article is a substantial technical contribution that innovatively classifies WT systems, complemented by empirical user data. Its main strengths lie in its depth and relevance; its weaknesses relate to scope (geographically limited survey), analytical rigor, and broader contextual discussions. Overall, it's a valuable piece that could be strengthened with broader data, deeper analysis, and a more explicit framing of research questions.

These are specific suggestions, with detailed explanations, to enhance the presentation of survey findings and deepen the discussion on ethical and privacy implications:

Specific Examples of Visual Aids to Enhance Survey Findings

1. Bar Charts or Histograms Showing Awareness Levels:

- **What:** Visualize the percentage of respondents who are aware of various WT tools (e.g., cookies, GPS, fingerprinting).
- **Why:** Makes it easier to grasp the relative familiarity with different tracking mechanisms at a glance.
- **Example:** A bar chart illustrating the percentage of users familiar with cookies (96.9%), GPS (68.9%), and fingerprinting (30%).

2. Pie Charts Depicting Demographic Distributions and Awareness Correlations:

- **What:** Show the breakdown of demographic groups (e.g., age, gender, tech background) and their respective levels of awareness or concern.
- **Why:** Helps identify which groups lack awareness and could benefit most from targeted education.

3. Cross-Tabulation Heatmaps:

- **What:** Display relationships between variables, such as "Computer Science Background" versus "Understanding of WT systems" or "Privacy Concerns."
- **Why:** Highlights patterns or gaps, making correlations visually intuitive.

4. Line Graphs Showing Change Over Time (if applicable):

- **What:** If longitudinal data or comparisons with prior studies are available, plot trends in awareness or concern over time.
- **Why:** Demonstrates how awareness evolves, informing policy or educational interventions.

5. Infographics Summarizing Key Findings:

- **What:** Combine key statistics in a concise visual format to be included as a summary slide or section highlight.
- **Why:** Facilitates quick comprehension for diverse audiences, including non-technical stakeholders.

Additional Details to Address Ethical and Privacy Implications in the Discussion

1. Impact of Increasing Tracking on Privacy Rights:

- Discuss how pervasive WT systems challenge individual rights to privacy and control over personal data, especially when user awareness remains limited.

2. Potential for Data Misuse and Harm:

- Explore scenarios where tracking could lead to unauthorized profiling, discrimination, or identity theft, emphasizing the ethical necessity for transparency and user consent.

3. Regulatory and Legal Challenges:

- Analyze how current frameworks like GDPR, CCPA, and emerging regulations attempt to regulate WT technologies, but face enforcement challenges, especially with new and opaque tracking techniques.

4. Technological Arms Race and Ethical Dilemmas:

- Address how developers of advanced tracking (e.g., fingerprinting, AI-enhanced profiling) raise questions about the ethics of invasive data collection versus business interests.

5. Recommendations for Ethical Design and User Empowerment:

- Suggest the importance of privacy-by-design principles, informed user consent mechanisms, and tools that empower users to understand and control tracking.

6. Societal Implications of Eroding Privacy Norms:

- Reflect on broader societal impacts, such as the normalization of surveillance, erosion of anonymity, and the potential chilling effects on free speech and behaviour.

Edited Summary

To **visualize survey findings more effectively**, incorporating bar charts for tracking tool familiarity, demographic-based heatmaps correlating awareness and concern, and infographics summarizing key data points would greatly improve clarity and impact.

In **expanding the discussion**, a more thorough examination of the ethical dilemmas posed by opaque tracking techniques, potential harms, regulatory gaps, and societal consequences would deepen the paper's relevance. Addressing these aspects can foster a holistic understanding of WT systems, emphasizing the importance of ethical design and policy measures.

Declarations

Potential competing interests: No potential competing interests to declare.