Position Paper: **Validatory Visualization** for Human-GenAI Co-Creation





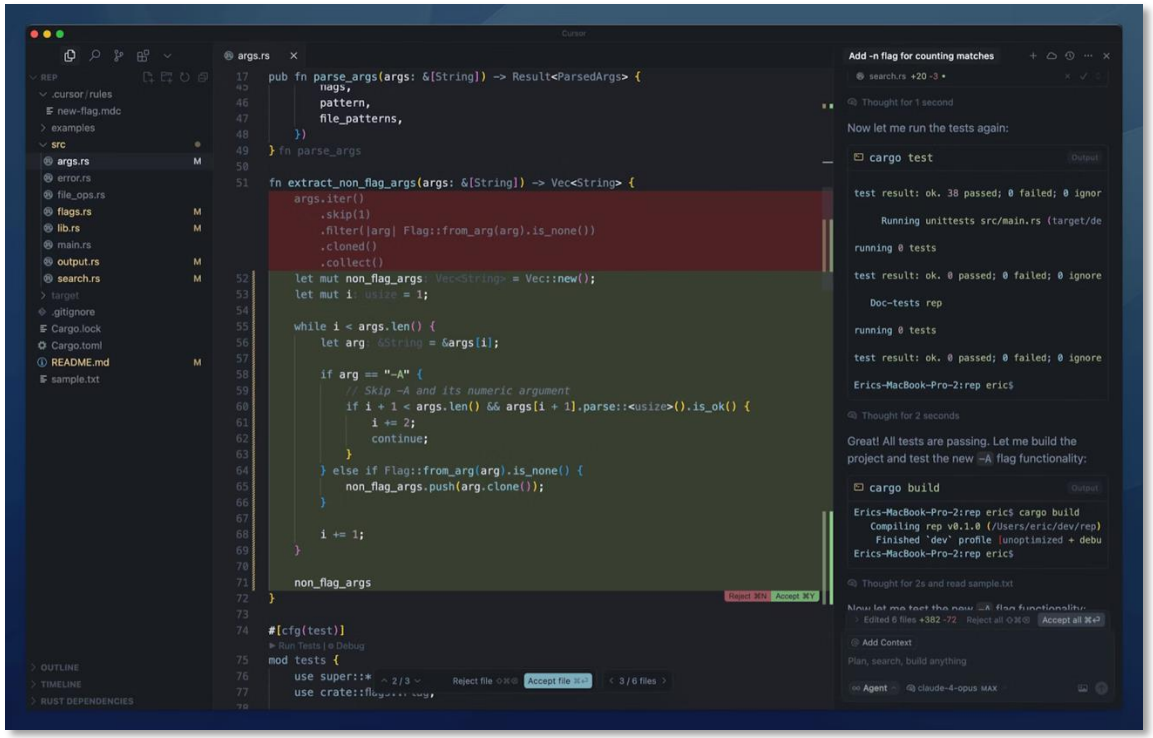

Yu Zhang[1] *presenter* Yuheng Zhao[2] Yuhan Guo[3] Guozheng Li[4] Siming Chen[2] Xiaoru Yuan[3]

[1]University of Oxford [2]Fudan University [3]Peking University [4]BIT

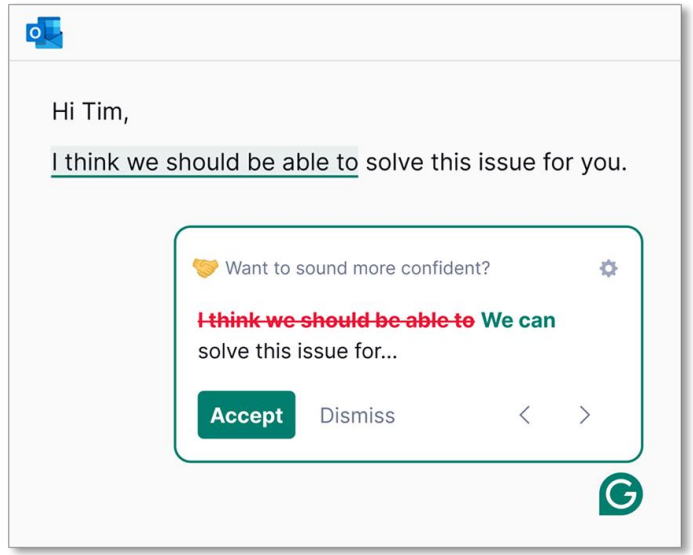
VIS2025

Background: Validatory VIS for AI

- AI in content creation apps: code editor, writing assistant, …
- **Validatory paradigm:** steer → generate → validate
- **Validatory VIS:** (*in-situ, minimalist*) VIS for validating AI output



Example – code editor: Cursor

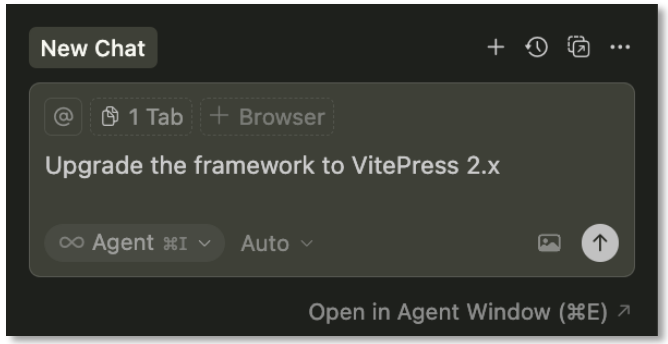


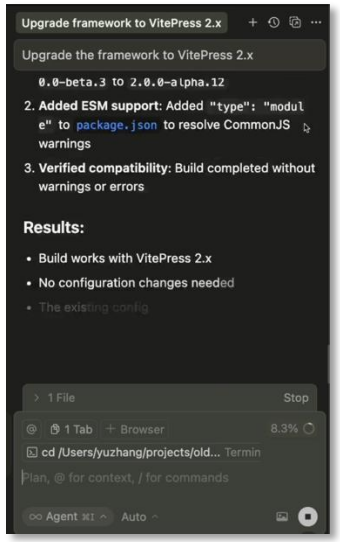
Example – writing assistant: Grammarly

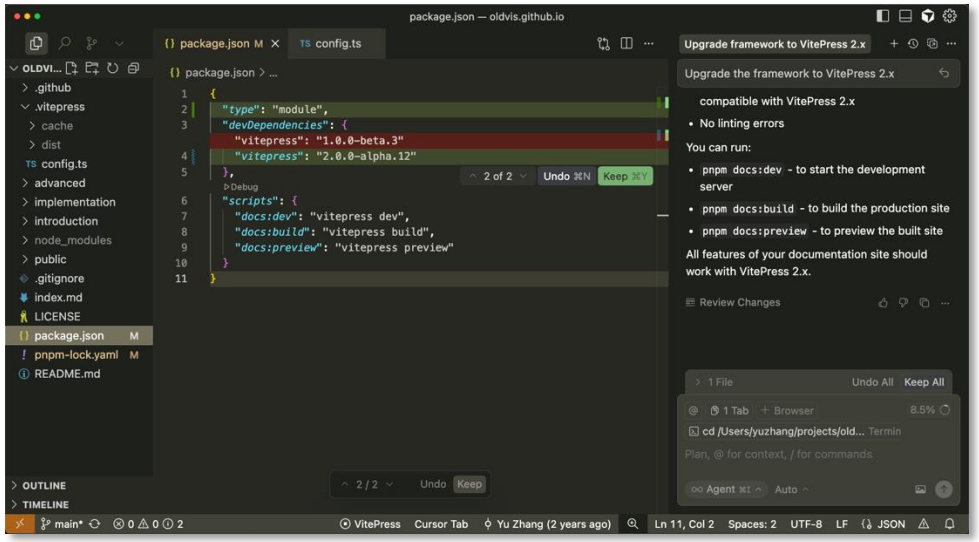
Background: Validatory VIS for AI

- Validatory paradigm: steer → generate → validate
- Validatory VIS: (*in-situ*, *minimalist*) VIS for validating AI output
- **Example:** Cursor

Steer: describe edit intent → **Generate:** edit code (and execute) → **Validate:** review code edits (in situ **color-encoding** + **text summary** in a separate panel)



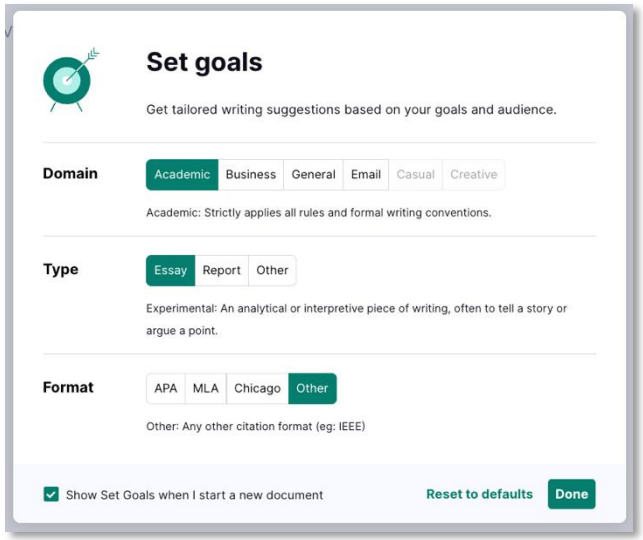


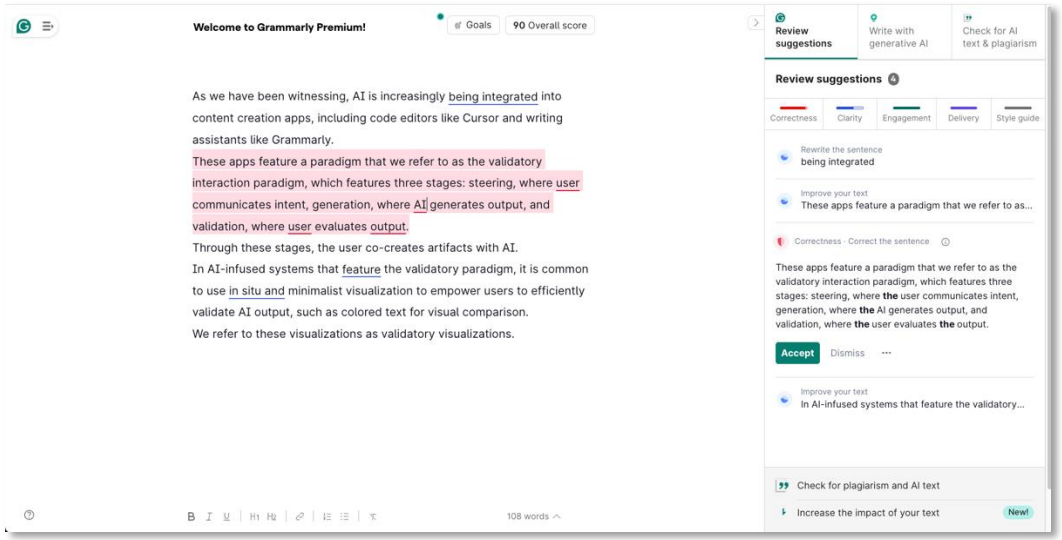


Background: Validatory VIS for AI

- Validatory paradigm: steer → generate → validate
- Validatory VIS: (*in-situ*, *minimalist*) VIS for validating AI output
- **Example:** Grammarly

Steer: config writing goals → **Generate:** check language issues and suggest fixes → **Validate:** review suggestions (in situ **color-encoding** + **list** in a separate panel)





Argument Chain

① GenAI **cannot fully automate** many generative tasks.

① ⇒ ② **Human validation** is usually required in GenAI applications.

③ **Many everyday tasks are generative** (e.g., writing) in nature and benefit from GenAI.

②, ③ ⇒ ④ Many **ordinary users need to validate** AI-generated content.

⑤ GenAI produces content at a **high rate**.

⑥ **Visualization amplifies human information processing abilities.**

④, ⑤, ⑥ ⇒ ⑦ **Validatory visualization is valuable for many ordinary users.**

The Value of Visualization **focus of validatory visualization** **widely studied in VIS**

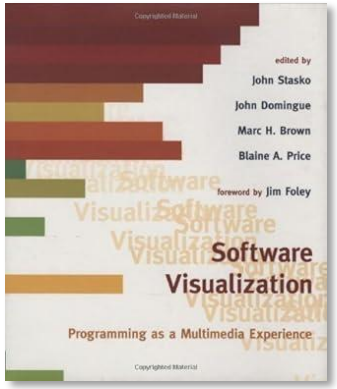
van Wijk, VIS, 2005:

*"a great visualization method is **used by many people**, who use it routinely to **obtain highly valuable knowledge**, **without having to spend time and money** on hardware, software, and effort."*

Opportunities

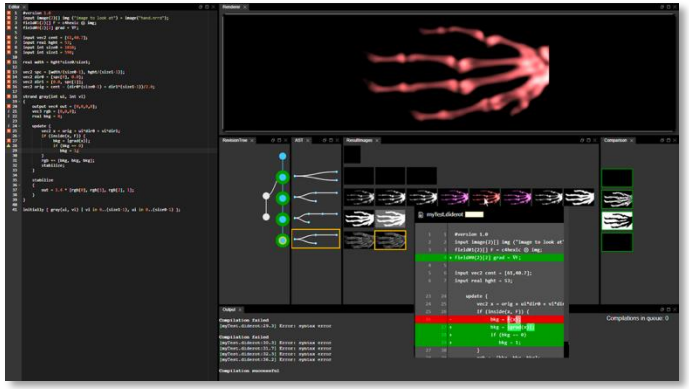
Validatory visualization for productivity tools

- **Q:** Can we improve Cursor's design for showing AI code edits?
- Relevant VIS topics: software visualization, visual comparison



Software Visualization, Stasko et al., 1998

Augmenting Code with In Situ Visualizations to Aid Program Understanding, Hoffswell et al., CHI, 2018



Vis-a-Vis: Visual Exploration of Visualization Source Code Evolution, Bolte and Bruckner, TVCG, 2021

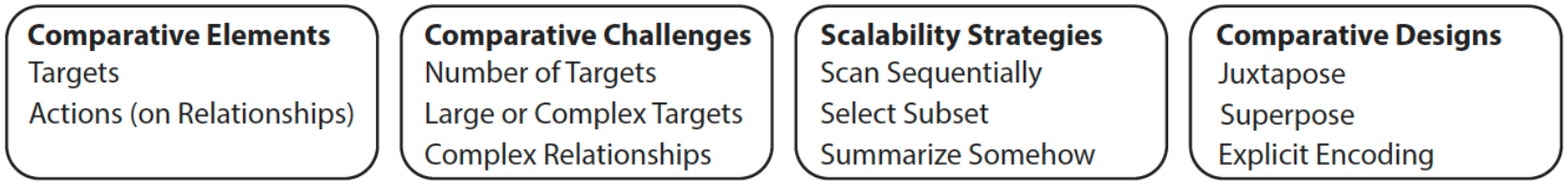


Fig. 1. Overview of the four considerations of comparison and the abstract categories they impose.

Considerations for Visualizing Comparison, Gleicher, TVCG, 2018

Cost Modeling

Cost of manual paradigm (👤): $C_{manual} = c_{generate}^{user}$

Validatory paradigm (👤 + 🤖)

1. 👤**Steering:** User requests AI to generate/edit content. (cost: c_{steer}^{user})
2. 🤖**Generation:** AI generates content (cost: $c_{generate}^{machine}$)
3. 👤**Validation:** User validates AI-generated content. (cost: $c_{validate}^{user}$)
   - With $p_{complete}$, the content is acceptable. Task **completed**.
   - With p_{refine}, the user manually refines the content. Task **completed**. (cost: c_{refine}^{user})
   - With $1 - p_{complete} - p_{refine}$, the user **goes back to the steering step**.

Expected task completion cost:

$$C = \sum_{n=1}^{\infty} (1-p)^{n-1} \left(pnc_{loop} + p_{refine} c_{refine}^{user} \right) = \boxed{\frac{1}{p} c_{loop} + \frac{1}{p} p_{refine} c_{refine}^{user}}$$

where:

$c_{loop} = c_{steer}^{user} + c_{generate}^{machine} + c_{validate}^{user}$ and $p = p_{complete} + p_{refine}$

Cost Modeling – Interpretation

Benefit of validatory paradigm: $r = 1 - C/C_{manual} = 1 - \frac{c_{loop} + p_{refine} c^{user}_{refine}}{p c^{user}_{generate}}$

⚖️ *When is the validatory paradigm beneficial (i.e., $r > 0$)?*

$c_{loop} + p_{refine} c^{user}_{refine} < p c^{user}_{generate}$ when:

- (🤖) AI is skilled (large p, small p_{refine})
- (👤) Manual content generation is costly (large $c^{user}_{generate}$)
- (👤 💻 🤖) Swift steer-generate-validate loop (small c_{loop} and c^{user}_{refine})

Cost Modeling – Interpretation

Benefit of validatory paradigm: $r = 1 - \frac{c_{loop} + p_{refine} c_{refine}^{user}}{pc_{generate}^{user}}$

⚖️ *Compare $c_{validate}^{user}$ and $c_{generate}^{user}$ (informally)*

- Evidence 1: P ≠ NP belief implies $c_{validate}^{user} < c_{generate}^{user}$ (in certain complexity classes).
- Evidence 2: RLHF assumes rating is simpler than generating ($c_{validate}^{user} < c_{generate}^{user}$).
- Interpretation: **The validatory paradigm converts a complex task to a simpler task.**
  - 🧐 Note: AI's competence determines how many times the simpler task needs to be executed.

Issues of the Validatory Paradigm

💭 *Assumptions*

- Output-centric
- Cost-sensitive

🧐 *Criticisms*

- Automation bias
- (Potential) loss of context, serendipity, creativity, and frictions

Conclusion

- Validatory visualization deserves more attention
- **VIS opportunities:** productivity tools, mixed-initiative data analysis

👀 **Paper URL:**

https://zhangyu94.github.io/projects/2025/Zhang2025Validatory/paper.pdf

Paper QR Code:

Email: yuzhang94@outlook.com
Website: zhangyu94.github.io

VIS2025

📢 **Related Talks:**

ProactiveVA, *LightVA*
by **Yuheng Zhao**

@Session: Agentic Visualization and Intelligent Systems
3:45pm - 5:00pm
Wednesday, Nov 5