北京大学 PEKING UNIVERSITY

高可信软件技术教育部重点实验室
Key Lab of High Confidence Software Technologies (Peking University), Ministry of Education



FeatX: Editing Software by Editing Features for Repository-Level Code Evolution

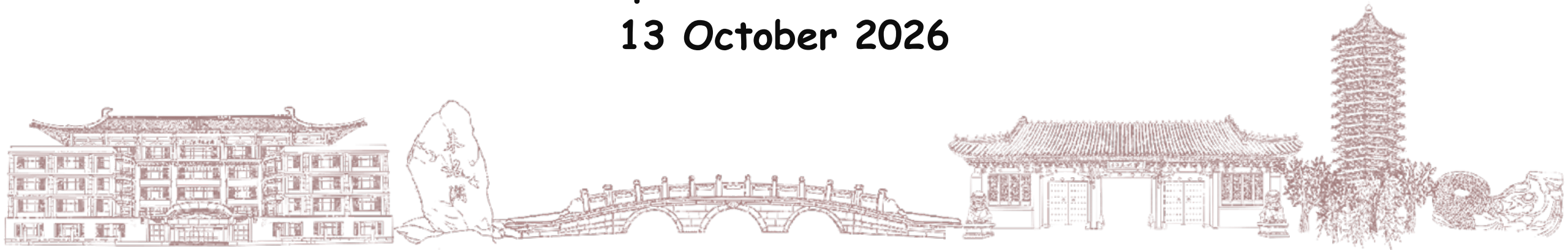
Xutian Li[1], Yifeng Zhu[1], Xianlin Zhao[1], Yanzhen Zou[1*], Lu Zhang[1], Bing Xie[1]

[1]Key Lab of HCST (PKU), MOE; SCS, Peking University, Beijing, China

{xtli25, yifengzhu25}@stu.pku.edu.cn, {zhaoxianlin, zouyz, zhanglucs, xiebing}@pku.edu.cn

Speaker: Xutian Li

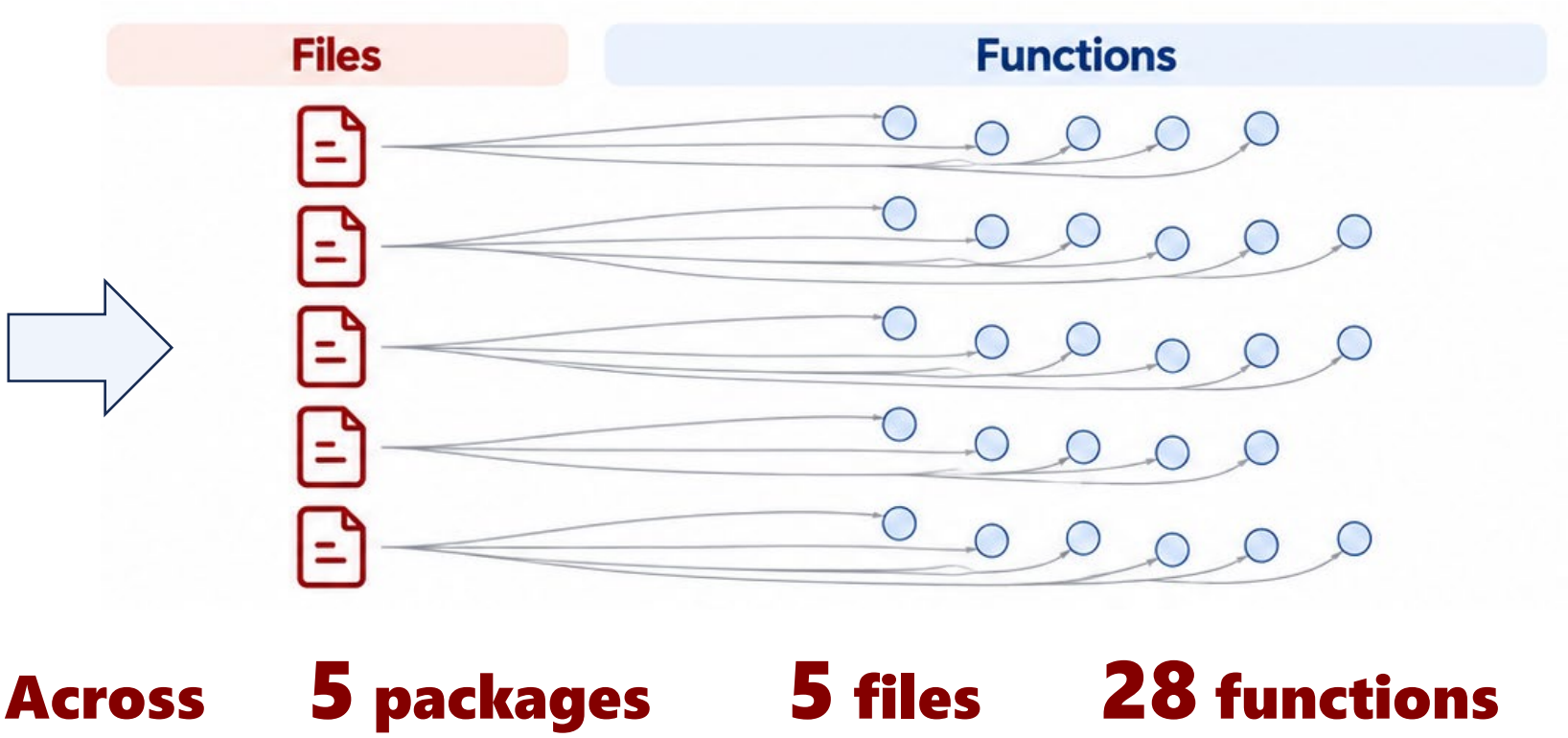
13 October 2026



Motivation Example

- Software evolution requires translating repository-level feature intent into coordinated code edits across multiple files.

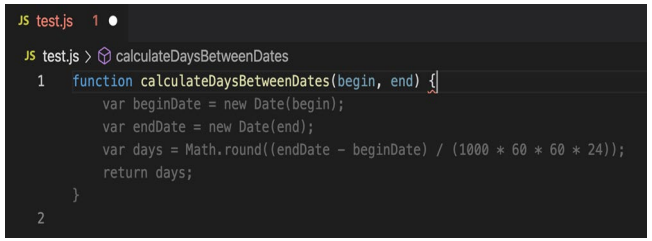




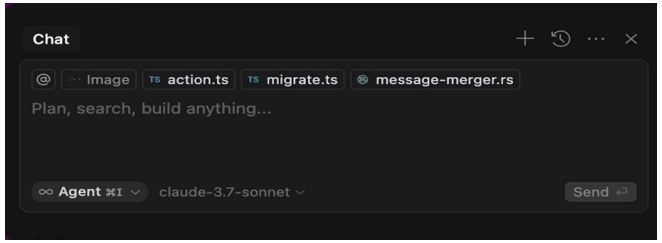
[1] https://github.com/Naccl/NBlog/commit/d47882e69d318b82827ada83b197742bec4d9669

Gaps

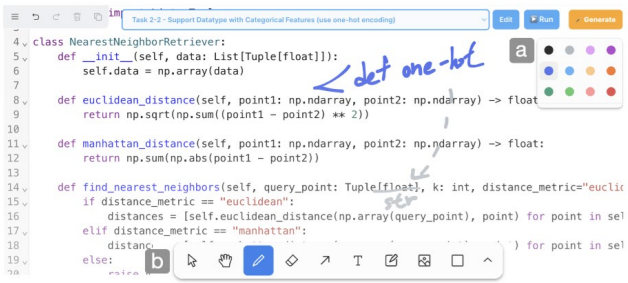
Existing Approach



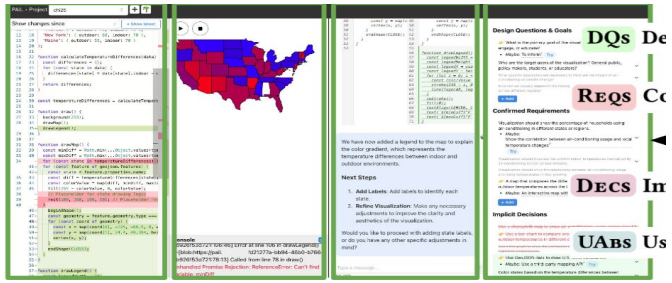
Autocompletion



Chat-based

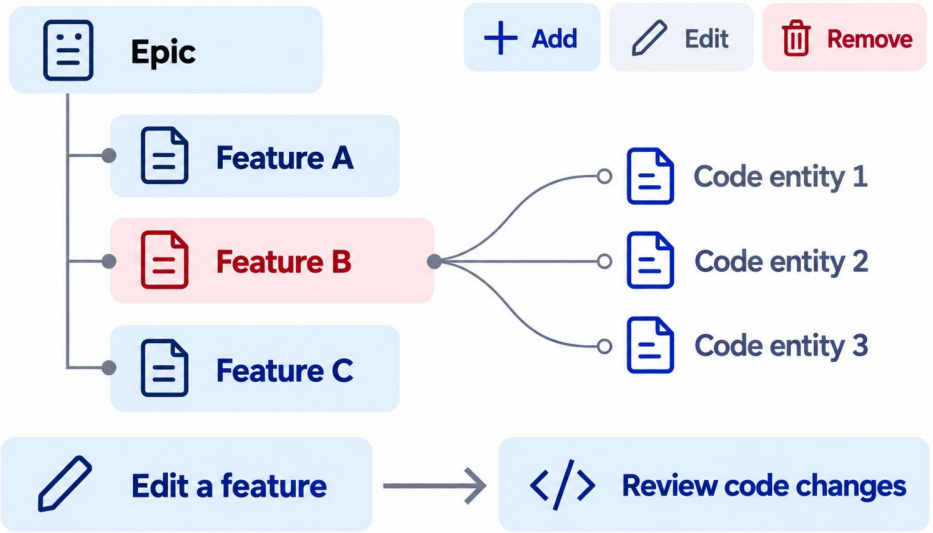


Annotation-based



Design-oriented

With FeatX



Gaps

No explicit feature list → Inconsistent!

Lack of a feature-to-code mapping → Less accurate localization!

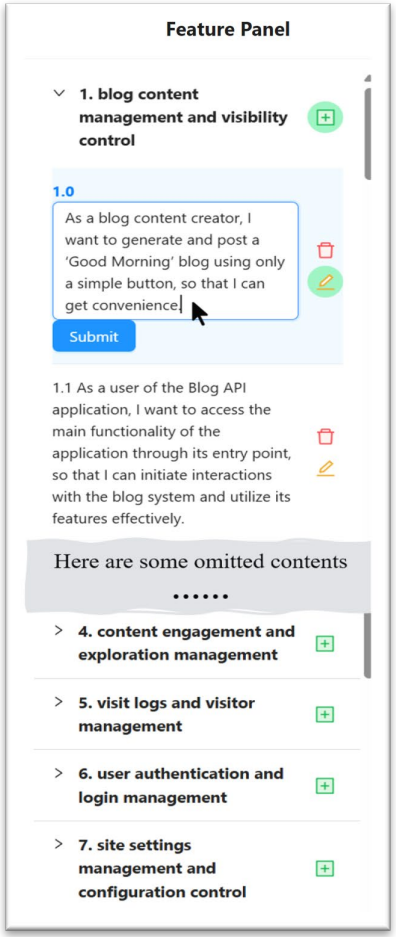
Usage Workflow

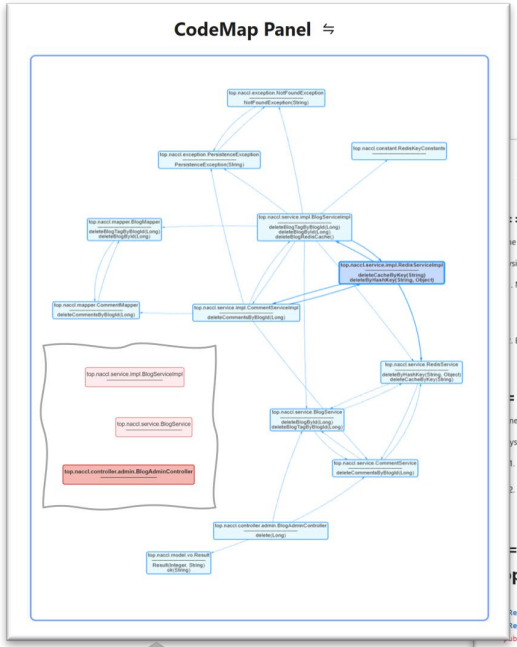
A Feature Panel
Lists hierarchical Epics/Features with user-controlled CRUD

B CodeMap Panel
Shows feature-related code entities, and highlights changed entity

C Agent Panel
Displays the reasoning process of the Revolution Agent in three stages

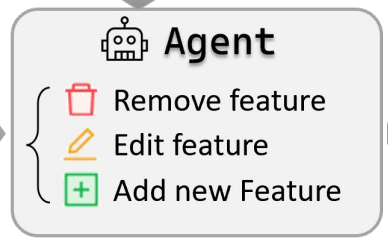
D Diff Panel
Presents modifications in line-level code diff of each class for confirmation

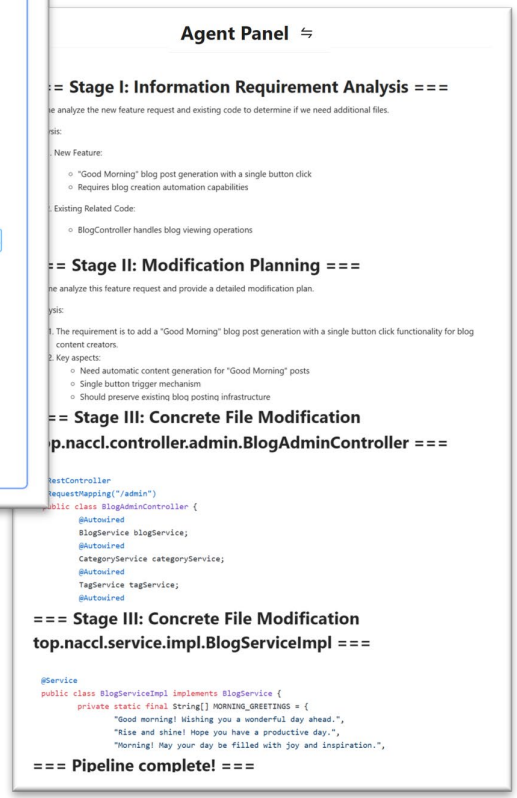




Switch automatically with task or manually

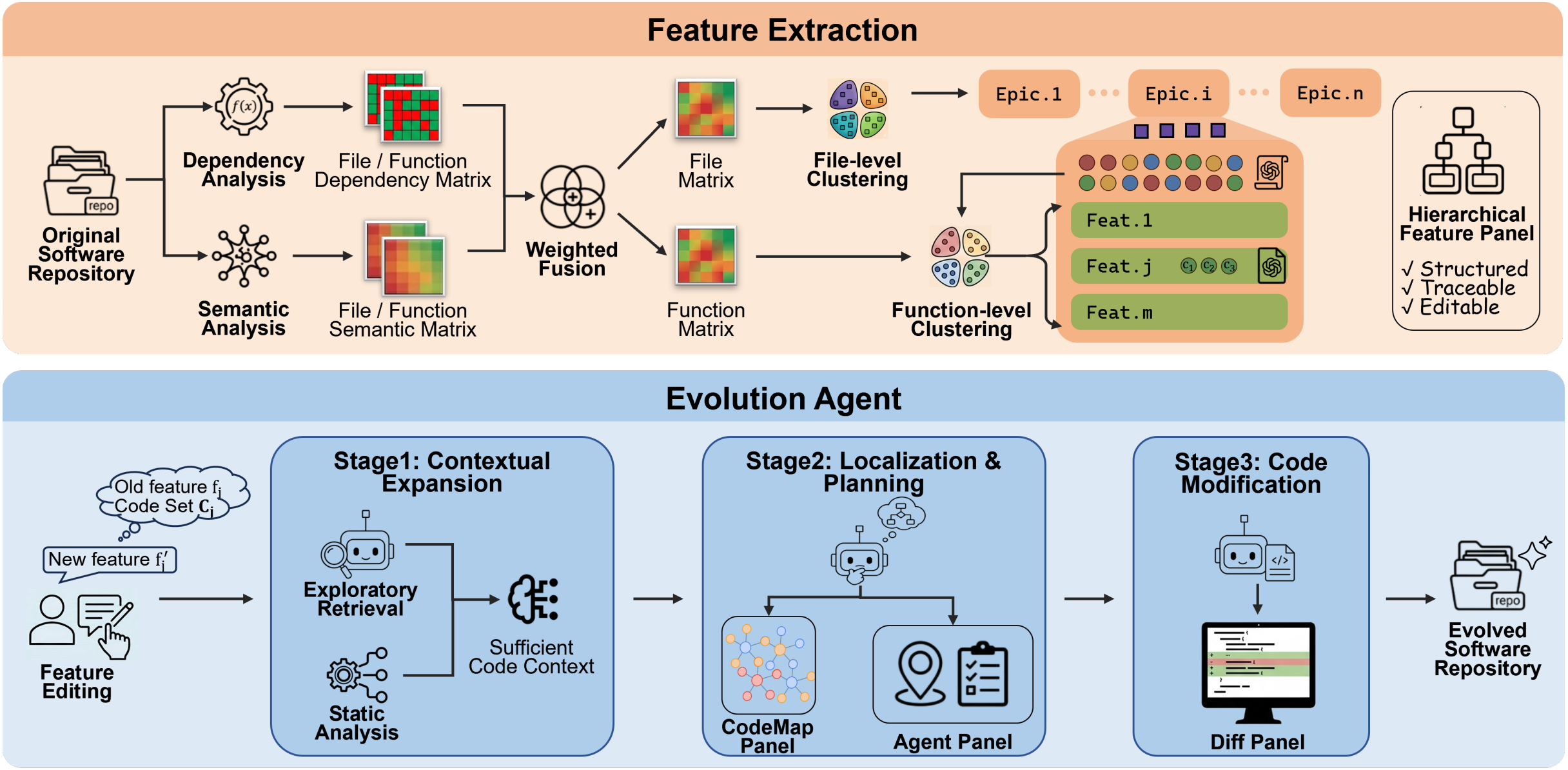
Context







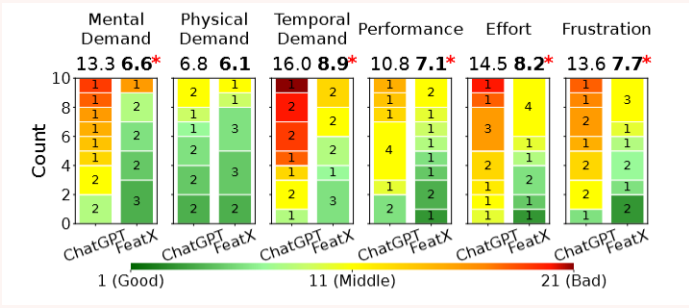
System Design

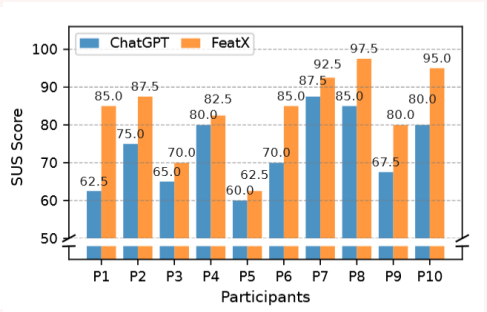


Evaluation Results

User Study

10 participants, Compared with vanilla ChatGPT





Mean NASA-TLX: 12.5→7.4

Mean SUS: 73→84

Cognitive Load (41% ↘)

System Usability (15% ↗)

Commit Replay

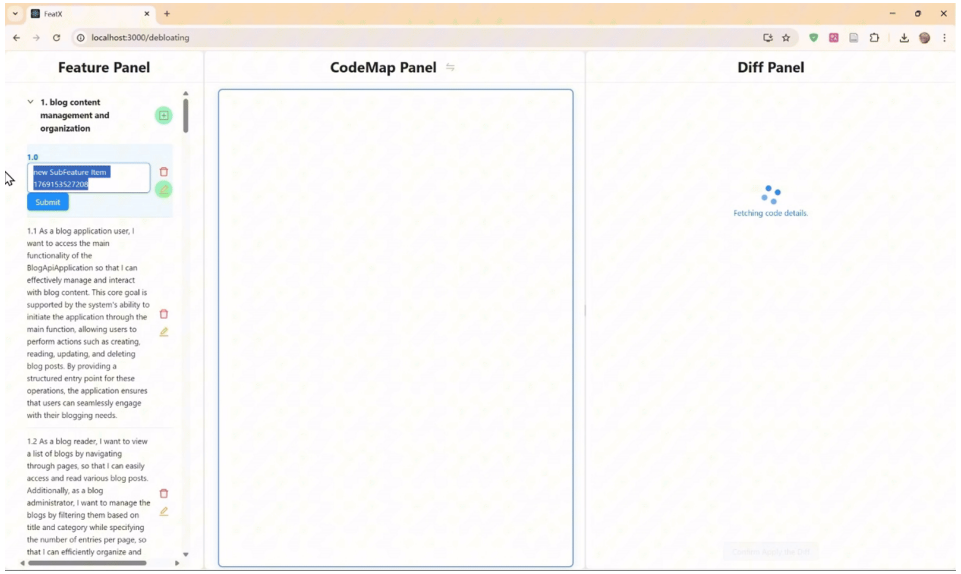
38 commits in 5 Java repositories

Approach	P (%)	R (%)	F1	Cost ($)
DeepSeek-v3	36.0	9.0	0.143	–
DeepSeek-v3.2-think	29.7	5.5	0.092	–
GPT-4o-mini	37.0	9.7	0.154	1.28
GPT-5.2	30.0	6.0	0.100	17.81
Claude-opus-4.5	**50.7**	*18.4*	*0.270*	45.05
Cursor (Agent)	15.9	13.0	0.143	–
FeatX (Ours)	*41.6*	**35.8**	**0.385**	**0.07**

Localization F1 (42.6%↗)

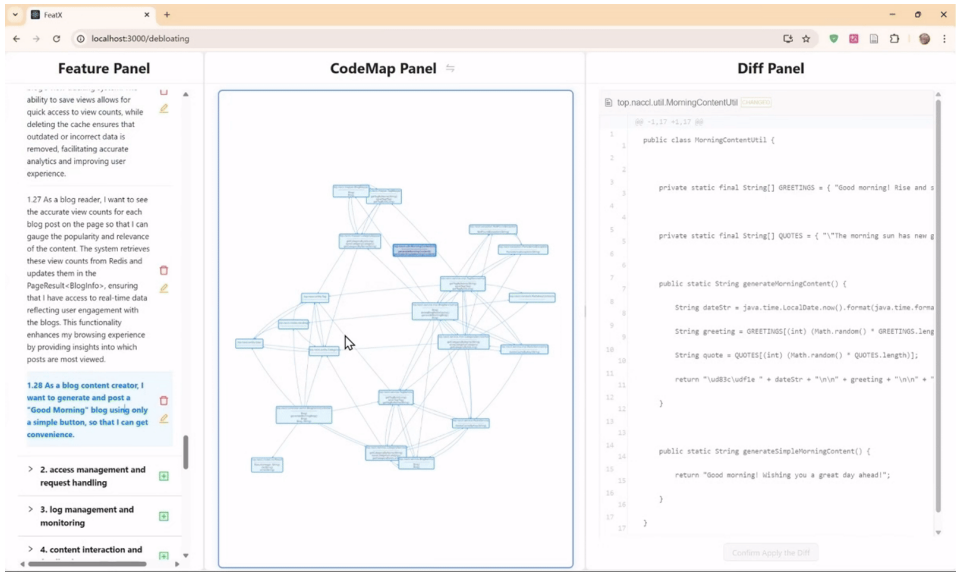
Monetary Cost (only $0.07)

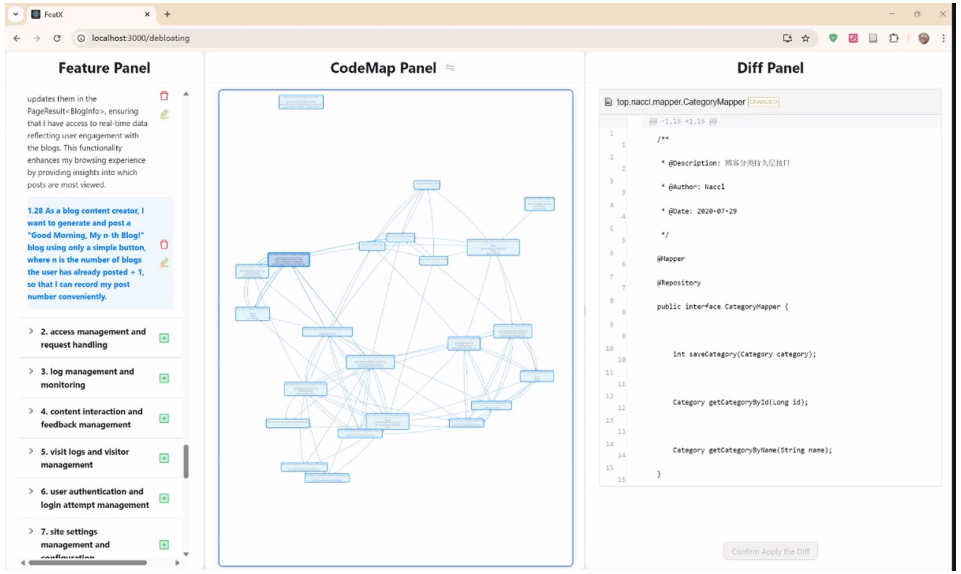
Case Demo



Add Feature

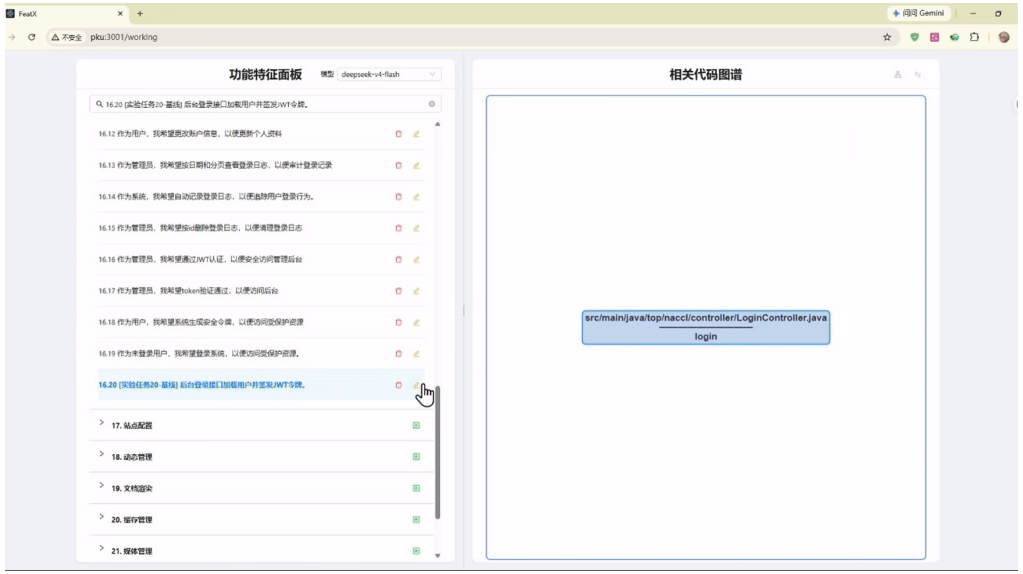
Edit Feature





Remove Feature

Edit Feature



Beyond FeatX: FeatLens

FeatX →

In FeatX, we have determined **where to edit**.



What dependency does LLM need to know before generating the code edits?

FeatLens[1]

① Feature-Oriented Repository Indexing

② Feature Retrieval & Dependency Seeding

③ Semantic-Structural Dependency Context Refinement

Better dependency retrieval

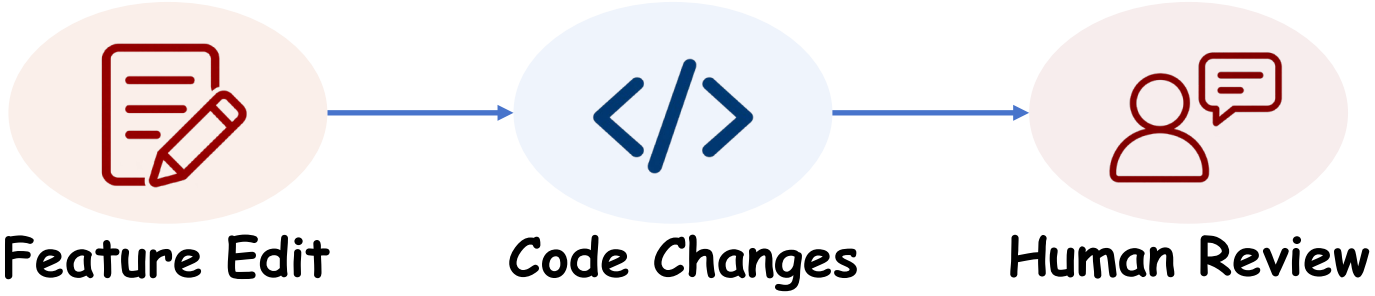
↓

Better code generation

[1] https://arxiv.org/abs/2609.26480

Q&A

Thanks!



- Contact Me (Xutian Li):
  - xtli25@stu.pku.edu.cn
  - shooter.lee492@gmail.com
  - https://github.com/a496263365

WeChat Telegram

GitHub Demo Video