

Contents

EXITON式 AI開発システム v1.3 (Claudeなし最適化版)	1
0. はじめに – v1.3 の目的	1
1. EXITON v1.3 の AI 役割構成図 (Claude完全排除版)	2
2. 重要アップデート (v1.3)	2
2.1 Googleモデル「エコーチェンバー」破壊機構を導入 効果	2
2.2 ChatGPT 破壊QA人格を“サイコモード”に公式化 破壊QAモデルを ChatGPT o1-preview に固定	3
破壊人格の追加ルール	3
2.3 PM (ChatGPTブラウザ) のボトルネック問題 → 自動承認ロードマップ 将来構成 (v2.0)	3
3. AI 役割の詳細 (v1.3 正式版)	3
3.1 ChatGPT (ブラウザPM)	3
3.2 Antigravity (Dev)	4
3.3 CODEX (VSCode)	4
3.4 Gemini CLI (外部批判監査員)	4
3.5 ChatGPT o1-preview (破壊QA / 冷酷担当)	4
4. PRワークフロー (v1.3)	5
5. 重要ルールファイルの改訂ポイント	5
5.1 QA_RULES_Gemini.md (追加文)	5
5.2 QA_RULES_ChatGPT_Destroyer.md (追加文)	5
6. EXITONの最終哲学 (v1.3)	6
END OF DOCUMENT	6

EXITON式 AI開発システム v1.3 (Claudeなし最適化版)

Updated AI Multi-Agent Constitution with Bias-Control & Enhanced QA Modes

0. はじめに – v1.3 の目的

EXITON v1.3 では、Gemini が指摘した以下の構造的リスクを正式に解消し、より堅牢で自律的な AI開発チームアーキテクチャ に進化させた。

1. Googleモデルのエコーチェンバー問題 (Dev=Antigravity ⇔ QA=Gemini)
2. ChatGPT破壊QAの甘さ・RLHF問題 (Claudeの冷徹さの代替)
3. PM (ChatGPTブラウザ) の手動承認ボトルネック

これらを改善し、EXITON チームを完全な“自律型AI開発OS”に近づける。

1. EXITON v1.3 の AI 役割構成図 (Claude完全排除版)

PM / Architecture (ChatGPT Browser)

docs/specs/*.md

Antigravity (Dev) CODEX (Test AI) Gemini CLI (External Reviewer)

PR

ChatGPT o1-preview QA

PM

Google系 (Antigravity/Gemini) と OpenAI系 (ChatGPT/Codex) でサンドイッチ構造を形成。
各AIのバイアスが互いを補正するように設計している。

2. 重要アップデート (v1.3)

2.1 Googleモデル「エコーチェンバー」破壊機構を導入

Gemini 用 QA ルールに以下を追加し、
Antigravity と似た推論方向をわざと破壊する人格を付与する。

Google
Antigravity AI

Antigravity
Google

効果

- Google系モデル同士の“お互い褒め合う問題”を物理的に排除
 - Gemini が「もう一人のAntigravity」にならない
 - OpenAI系 QA (o1) が最後にバイアスを潰す構造を補強
-

2.2 ChatGPT 破壊QA人格を“サイコモード”に公式化

Claude の代替として、以下を採用：

破壊QAモデルを ChatGPT o1-preview に固定

- 空気を読まない
- 人間への配慮が極端に少ない
- 純粋な論理推論
- Claude 並みに容赦がない

破壊人格の追加ルール

これにより、Claude の冷徹さを超える“論理的サイコパスQA”を獲得。

2.3 PM（ChatGPTブラウザ）のボトルネック問題 → 自動承認ロードマップ

今はブラウザPMが最終判断をしているが、
v1.3 からは自動化ロードマップを正式に規定する。

将来構成（v2.0）

GitHub Actions → → OpenAI PM API → or

つまり、
人間PM（こうすけ）は **思想のアップデートのみ担当** となり、
個別PRの承認は **AI PM（API 経由）** が行う。

3. AI 役割の詳細（v1.3 正式版）

3.1 ChatGPT（ブラウザPM）

思想・価値基準・非機能要件を維持する唯一の存在。

- FUNC_SPEC / NFR_SPEC の更新
 - AI間の判断軸の統一
 - 開発憲法の改訂
 - 最終的な「EXITONらしさ」の監査
-

3.2 Antigravity (Dev)

Google系実働エンジニア。

- 仕様通り実装
 - 自己pytest
 - PR作成
 - WHY・REQ-ID の明記
 - ログ & エラー処理の徹底
-

3.3 CODEX (VSCode)

「仕様→テスト」の自動化装置。

- pytest自動生成
 - カバレッジ向上
 - 境界値・異常系テスト
 - 仕様との整合チェック
-

3.4 Gemini CLI (外部批判監査員)

Google系モデルの偏りを逆に利用し、
あえて“**Googleの自分を疑わせる人格**”を与えて使う。

- 設計矛盾
- 非機能要件違反
- Antigravity の過度な楽観ロジック
- 誤った仮定

を指摘する。

3.5 ChatGPT o1-preview (破壊QA / 冷酷担当)

Claude の冷徹さ + 数学的破壊思考。

- 例外処理の抜け
- 競合条件
- 無限ループ
- スレッド安全性

- ルックアヘッド検知
 - データ品質の破壊チェック
-

4. PRワークフロー (v1.3)

1. Antigravity PR
 - WHY
 - REQ-ID
 - pytest
 2. CODEX
 -
 - pytest
 3. Gemini
 -
 -
 4. ChatGPT o1-preview QA
 -
 -
 5. PM ChatGPT
 - EXITON
 -
-

5. 重要ルールファイルの改訂ポイント

5.1 QA_RULES_Gemini.md (追加文)

Google
Antigravity

5.2 QA_RULES_ChatGPT_Destroyer.md (追加文)

6. EXITONの最終哲学 (v1.3)

AI

AI Pull Request

AI PM

Google OpenAI

OpenAI Google

EXITON AI

END OF DOCUMENT