

When AI Makes Each of Us Better & All of Us Worse (and what to do about it)¹

Gary King

Institute for Quantitative Social Science
Harvard University

Harvard AI Research Seminar, 10/6/2026

¹Papers, Slides, Data: [GaryKing.org](https://garyking.org). Based on joint work with (different subsets of) Steve Worthington, Queenie Luo, Mike Puett, and Mike Smith

Three Points

Three Points

1. Goals AI is optimized for: succeeding

Three Points

1. Goals AI is optimized for: succeeding
 - Makes each of us better off

Three Points

1. Goals AI is optimized for: succeeding
 - Makes each of us better off
2. Goals AI isn't optimized for: “failing” (obviously)

Three Points

1. Goals AI is optimized for: succeeding
 - Makes each of us better off
2. Goals AI isn't optimized for: “failing” (obviously)
 - Science/art/literature becoming conventional, homogenous

Three Points

1. Goals AI is optimized for: succeeding
 - Makes each of us better off
2. Goals AI isn't optimized for: “failing” (obviously)
 - Science/art/literature becoming conventional, homogenous
 - Makes all of us worse off;

Three Points

1. Goals AI is optimized for: succeeding
 - Makes each of us better off
2. Goals AI isn't optimized for: “failing” (obviously)
 - Science/art/literature becoming conventional, homogenous
 - Makes all of us worse off; plus big problems for you

Three Points

1. Goals AI is optimized for: succeeding
 - Makes each of us better off
2. Goals AI isn't optimized for: “failing” (obviously)
 - Science/art/literature becoming conventional, homogenous
 - Makes all of us worse off; plus big problems for you
3. How to fix the problem

Three Points

1. Goals AI is optimized for: succeeding
 - Makes each of us better off
2. Goals AI isn't optimized for: “failing” (obviously)
 - Science/art/literature becoming conventional, homogenous
 - Makes all of us worse off; plus big problems for you
3. How to fix the problem
 - Cool easy-to-use algorithm to come...

E.g.: AI Article Reviews >> Best Human Peer Reviews

E.g.: AI Article Reviews >> Best Human Peer Reviews

- **Humans:** IQSS RapidPeer (134 reviews; 2019–2026)

E.g.: AI Article Reviews >> Best Human Peer Reviews

- **Humans:** IQSS RapidPeer (134 reviews; 2019–2026)
 - **Highest quality:** double-blind, constructive, paid (well)

E.g.: AI Article Reviews >> Best Human Peer Reviews

- **Humans:** IQSS RapidPeer (134 reviews; 2019–2026)
 - **Highest quality:** double-blind, constructive, paid (well)
 - **Author evaluations:** very high

E.g.: AI Article Reviews >> Best Human Peer Reviews

- **Humans:** IQSS RapidPeer (134 reviews; 2019–2026)
 - **Highest quality:** double-blind, constructive, paid (well)
 - **Author evaluations:** very high
 - **(Human) journal editors:** “among the best I’ve seen”

E.g.: AI Article Reviews >> Best Human Peer Reviews

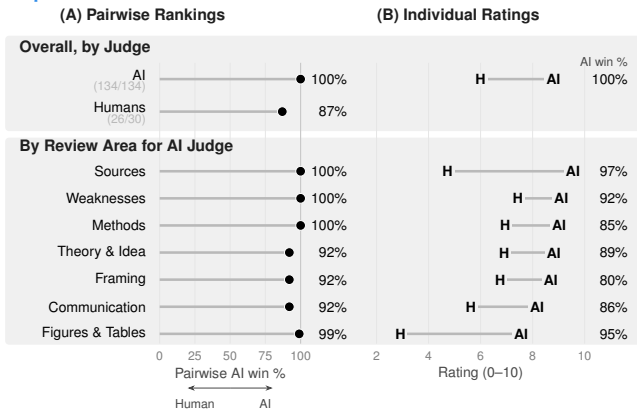
- **Humans:** IQSS RapidPeer (134 reviews; 2019–2026)
 - **Highest quality:** double-blind, constructive, paid (well)
 - **Author evaluations:** very high
 - **(Human) journal editors:** “among the best I’ve seen”
- **AI:** Opus 4.8

E.g.: AI Article Reviews >> Best Human Peer Reviews

- **Humans:** IQSS RapidPeer (134 reviews; 2019–2026)
 - **Highest quality:** double-blind, constructive, paid (well)
 - **Author evaluations:** very high
 - **(Human) journal editors:** “among the best I’ve seen”
- **AI:** Opus 4.8
- **Comparison:**

E.g.: AI Article Reviews >> Best Human Peer Reviews

- **Humans:** IQSS RapidPeer (134 reviews; 2019–2026)
 - **Highest quality:** double-blind, constructive, paid (well)
 - **Author evaluations:** very high
 - **(Human) journal editors:** “among the best I’ve seen”
- **AI:** Opus 4.8
- **Comparison:**



E.g.: AI Article Reviews >> Best Human Peer Reviews

- **Humans:** IQSS RapidPeer (134 reviews; 2019–2026)
 - **Highest quality:** double-blind, constructive, paid (well)
 - **Author evaluations:** very high
 - **(Human) journal editors:** “among the best I’ve seen”
- **AI:** Opus 4.8
- **Comparison:**

E.g.: AI Article Reviews >> Best Human Peer Reviews

- **Humans:** IQSS RapidPeer (134 reviews; 2019–2026)
 - **Highest quality:** double-blind, constructive, paid (well)
 - **Author evaluations:** very high
 - **(Human) journal editors:** “among the best I’ve seen”
- **AI:** Opus 4.8
- **Comparison:** AI wins

E.g.: AI Article Reviews >> Best Human Peer Reviews

- **Humans:** IQSS RapidPeer (134 reviews; 2019–2026)
 - **Highest quality:** double-blind, constructive, paid (well)
 - **Author evaluations:** very high
 - **(Human) journal editors:** “among the best I’ve seen”
- **AI:** Opus 4.8
- **Comparison:** AI wins
- **Yes, we thought of that:**

E.g.: AI Article Reviews >> Best Human Peer Reviews

- **Humans:** IQSS RapidPeer (134 reviews; 2019–2026)
 - **Highest quality:** double-blind, constructive, paid (well)
 - **Author evaluations:** very high
 - **(Human) journal editors:** “among the best I’ve seen”
- **AI:** Opus 4.8
- **Comparison:** AI wins
- **Yes, we thought of that:**
 - Control for length

E.g.: AI Article Reviews >> Best Human Peer Reviews

- **Humans:** IQSS RapidPeer (134 reviews; 2019–2026)
 - **Highest quality:** double-blind, constructive, paid (well)
 - **Author evaluations:** very high
 - **(Human) journal editors:** “among the best I’ve seen”
- **AI:** Opus 4.8
- **Comparison:** AI wins
- **Yes, we thought of that:**
 - Control for length
 - Compare to post hoc selection of top human reviewers

E.g.: AI Article Reviews >> Best Human Peer Reviews

- **Humans:** IQSS RapidPeer (134 reviews; 2019–2026)
 - **Highest quality:** double-blind, constructive, paid (well)
 - **Author evaluations:** very high
 - **(Human) journal editors:** “among the best I’ve seen”
- **AI:** Opus 4.8
- **Comparison:** AI wins
- **Yes, we thought of that:**
 - Control for length
 - Compare to post hoc selection of top human reviewers
 - Rewrite human reviews in AI style

E.g.: AI Article Reviews >> Best Human Peer Reviews

- **Humans:** IQSS RapidPeer (134 reviews; 2019–2026)
 - **Highest quality:** double-blind, constructive, paid (well)
 - **Author evaluations:** very high
 - **(Human) journal editors:** “among the best I’ve seen”
- **AI:** Opus 4.8
- **Comparison:** AI wins
- **Yes, we thought of that:**
 - Control for length
 - Compare to post hoc selection of top human reviewers
 - Rewrite human reviews in AI style
 - No change after ChatGPT

E.g.: AI Article Reviews >> Best Human Peer Reviews

- **Humans:** IQSS RapidPeer (134 reviews; 2019–2026)
 - **Highest quality:** double-blind, constructive, paid (well)
 - **Author evaluations:** very high
 - **(Human) journal editors:** “among the best I’ve seen”
- **AI:** Opus 4.8
- **Comparison:** AI wins
- **Yes, we thought of that:**
 - Control for length
 - Compare to post hoc selection of top human reviewers
 - Rewrite human reviews in AI style
 - No change after ChatGPT
 - AI: no hallucination; correct more often than humans

E.g.: AI Article Reviews >> Best Human Peer Reviews

- **Humans:** IQSS RapidPeer (134 reviews; 2019–2026)
 - **Highest quality:** double-blind, constructive, paid (well)
 - **Author evaluations:** very high
 - **(Human) journal editors:** “among the best I’ve seen”
- **AI:** Opus 4.8
- **Comparison:** AI wins
- **Yes, we thought of that:**
 - Control for length
 - Compare to post hoc selection of top human reviewers
 - Rewrite human reviews in AI style
 - No change after ChatGPT
 - AI: no hallucination; correct more often than humans
 - No change with repeat runs or different LLMs

E.g.: AI Article Reviews >> Best Human Peer Reviews

- **Humans:** IQSS RapidPeer (134 reviews; 2019–2026)
 - **Highest quality:** double-blind, constructive, paid (well)
 - **Author evaluations:** very high
 - **(Human) journal editors:** “among the best I’ve seen”
- **AI:** Opus 4.8
- **Comparison:** AI wins
- **Yes, we thought of that:**
 - Control for length
 - Compare to post hoc selection of top human reviewers
 - Rewrite human reviews in AI style
 - No change after ChatGPT
 - AI: no hallucination; correct more often than humans
 - No change with repeat runs or different LLMs
 - No difference by: “novelty”;

E.g.: AI Article Reviews >> Best Human Peer Reviews

- **Humans:** IQSS RapidPeer (134 reviews; 2019–2026)
 - **Highest quality:** double-blind, constructive, paid (well)
 - **Author evaluations:** very high
 - **(Human) journal editors:** “among the best I’ve seen”
- **AI:** Opus 4.8
- **Comparison:** AI wins
- **Yes, we thought of that:**
 - Control for length
 - Compare to post hoc selection of top human reviewers
 - Rewrite human reviews in AI style
 - No change after ChatGPT
 - AI: no hallucination; correct more often than humans
 - No change with repeat runs or different LLMs
 - No difference by: “novelty”; research style (stats, math, qualitative);

E.g.: AI Article Reviews >> Best Human Peer Reviews

- **Humans:** IQSS RapidPeer (134 reviews; 2019–2026)
 - **Highest quality:** double-blind, constructive, paid (well)
 - **Author evaluations:** very high
 - **(Human) journal editors:** “among the best I’ve seen”
- **AI:** Opus 4.8
- **Comparison:** AI wins
- **Yes, we thought of that:**
 - Control for length
 - Compare to post hoc selection of top human reviewers
 - Rewrite human reviews in AI style
 - No change after ChatGPT
 - AI: no hallucination; correct more often than humans
 - No change with repeat runs or different LLMs
 - No difference by: “novelty”; research style (stats, math, qualitative); author status, age;

E.g.: AI Article Reviews >> Best Human Peer Reviews

- **Humans:** IQSS RapidPeer (134 reviews; 2019–2026)
 - **Highest quality:** double-blind, constructive, paid (well)
 - **Author evaluations:** very high
 - **(Human) journal editors:** “among the best I’ve seen”
- **AI:** Opus 4.8
- **Comparison:** AI wins
- **Yes, we thought of that:**
 - Control for length
 - Compare to post hoc selection of top human reviewers
 - Rewrite human reviews in AI style
 - No change after ChatGPT
 - AI: no hallucination; correct more often than humans
 - No change with repeat runs or different LLMs
 - No difference by: “novelty”; research style (stats, math, qualitative); author status, age; quality of journal where paper was eventually published

E.g.: AI Article Reviews >> Best Human Peer Reviews

- **Humans:** IQSS RapidPeer (134 reviews; 2019–2026)
 - **Highest quality:** double-blind, constructive, paid (well)
 - **Author evaluations:** very high
 - **(Human) journal editors:** “among the best I’ve seen”
- **AI:** Opus 4.8
- **Comparison:** AI wins

E.g.: AI Article Reviews >> Best Human Peer Reviews

- **Humans:** IQSS RapidPeer (134 reviews; 2019–2026)
 - **Highest quality:** double-blind, constructive, paid (well)
 - **Author evaluations:** very high
 - **(Human) journal editors:** “among the best I’ve seen”
- **AI:** Opus 4.8
- **Comparison:** AI wins
- **Individual Benefits**

E.g.: AI Article Reviews >> Best Human Peer Reviews

- **Humans:** IQSS RapidPeer (134 reviews; 2019–2026)
 - **Highest quality:** double-blind, constructive, paid (well)
 - **Author evaluations:** very high
 - **(Human) journal editors:** “among the best I’ve seen”
 - **AI:** Opus 4.8
 - **Comparison:** AI wins
 - **Individual Benefits**
-
- **Collective Costs**

E.g.: AI Article Reviews >> Best Human Peer Reviews

- **Humans:** IQSS RapidPeer (134 reviews; 2019–2026)
 - **Highest quality:** double-blind, constructive, paid (well)
 - **Author evaluations:** very high
 - **(Human) journal editors:** “among the best I’ve seen”
- **AI:** Opus 4.8
- **Comparison:** AI wins
- **Individual Benefits**
 - **Better papers** (self-supervised R&Rs): fewer errors in logic, math, references, argument, literature, data analyses, graphics, and journal style
- **Collective Costs**

E.g.: AI Article Reviews >> Best Human Peer Reviews

- **Humans:** IQSS RapidPeer (134 reviews; 2019–2026)
 - **Highest quality:** double-blind, constructive, paid (well)
 - **Author evaluations:** very high
 - **(Human) journal editors:** “among the best I’ve seen”
- **AI:** Opus 4.8
- **Comparison:** AI wins
- **Individual Benefits**
 - **Better papers** (self-supervised R&Rs): fewer errors in logic, math, references, argument, literature, data analyses, graphics, and journal style
 - **Better for reviewers:** only high level tasks
- **Collective Costs**

E.g.: AI Article Reviews >> Best Human Peer Reviews

- **Humans:** IQSS RapidPeer (134 reviews; 2019–2026)
 - **Highest quality:** double-blind, constructive, paid (well)
 - **Author evaluations:** very high
 - **(Human) journal editors:** “among the best I’ve seen”
- **AI:** Opus 4.8
- **Comparison:** AI wins
- **Individual Benefits**
 - **Better papers** (self-supervised R&Rs): fewer errors in logic, math, references, argument, literature, data analyses, graphics, and journal style
 - **Better for reviewers:** only high level tasks
 - **Better for journal editors:** no hunt for (lame) reviewers
- **Collective Costs**

E.g.: AI Article Reviews >> Best Human Peer Reviews

- **Humans:** IQSS RapidPeer (134 reviews; 2019–2026)
 - **Highest quality:** double-blind, constructive, paid (well)
 - **Author evaluations:** very high
 - **(Human) journal editors:** “among the best I’ve seen”
- **AI:** Opus 4.8
- **Comparison:** AI wins
- **Individual Benefits**
 - **Better papers** (self-supervised R&Rs): fewer errors in logic, math, references, argument, literature, data analyses, graphics, and journal style
 - **Better for reviewers:** only high level tasks
 - **Better for journal editors:** no hunt for (lame) reviewers
 - **Better for authors:** Programmatically controlled biases: demographic, status, political, field, sycophantic, etc.
- **Collective Costs**

E.g.: AI Article Reviews >> Best Human Peer Reviews

- **Humans:** IQSS RapidPeer (134 reviews; 2019–2026)
 - **Highest quality:** double-blind, constructive, paid (well)
 - **Author evaluations:** very high
 - **(Human) journal editors:** “among the best I’ve seen”
- **AI:** Opus 4.8
- **Comparison:** AI wins
- **Individual Benefits**
 - **Better papers** (self-supervised R&Rs): fewer errors in logic, math, references, argument, literature, data analyses, graphics, and journal style
 - **Better for reviewers:** only high level tasks
 - **Better for journal editors:** no hunt for (lame) reviewers
 - **Better for authors:** Programmatically controlled biases: demographic, status, political, field, sycophantic, etc.
- **Collective Costs**
 - AI for ideation $\rightsquigarrow$ homogenized results

The problem: AI's Not Optimized For All Goals

The problem: AI's Not Optimized For All Goals

- AI is good at: Factual questions

The problem: AI's Not Optimized For All Goals

- AI is good at: Factual questions
 - Optimized for the task; agreement on success; public leaderboards

The problem: AI's Not Optimized For All Goals

- AI is good at: Factual questions
 - Optimized for the task; agreement on success; public leaderboards
 - $\leadsto$ Improving fast

The problem: AI's Not Optimized For All Goals

- AI is good at: Factual questions
 - Optimized for the task; agreement on success; public leaderboards
 - $\leadsto$ Improving fast
- AI's bad at: “Search quests” (open-ended questions)

The problem: AI's Not Optimized For All Goals

- AI is good at: Factual questions
 - Optimized for the task; agreement on success; public leaderboards
 - $\leadsto$ Improving fast
- AI's bad at: “Search quests” (open-ended questions)
 - Examples:

The problem: AI's Not Optimized For All Goals

- AI is good at: Factual questions
 - Optimized for the task; agreement on success; public leaderboards
 - $\leadsto$ Improving fast
- AI's bad at: “Search quests” (open-ended questions)
 - **Examples:** Paper topics,

The problem: AI's Not Optimized For All Goals

- AI is good at: Factual questions
 - Optimized for the task; agreement on success; public leaderboards
 - $\leadsto$ Improving fast
- AI's bad at: “Search quests” (open-ended questions)
 - **Examples:** Paper topics, headphones,

The problem: AI's Not Optimized For All Goals

- AI is good at: Factual questions
 - Optimized for the task; agreement on success; public leaderboards
 - $\leadsto$ Improving fast
- AI's bad at: “Search quests” (open-ended questions)
 - **Examples:** Paper topics, headphones, perfect wedding dress,

The problem: AI's Not Optimized For All Goals

- AI is good at: Factual questions
 - Optimized for the task; agreement on success; public leaderboards
 - $\leadsto$ Improving fast
- AI's bad at: “Search quests” (open-ended questions)
 - **Examples:** Paper topics, headphones, perfect wedding dress, unintended policy consequences,

The problem: AI's Not Optimized For All Goals

- AI is good at: Factual questions
 - Optimized for the task; agreement on success; public leaderboards
 - $\leadsto$ Improving fast
- AI's bad at: “Search quests” (open-ended questions)
 - **Examples:** Paper topics, headphones, perfect wedding dress, unintended policy consequences, killer startup idea

The problem: AI's Not Optimized For All Goals

- **AI is good at:** Factual questions
 - Optimized for the task; agreement on success; public leaderboards
 - $\leadsto$ Improving fast
- **AI's bad at:** “Search quests” (open-ended questions)
 - **Examples:** Paper topics, headphones, perfect wedding dress, unintended policy consequences, killer startup idea
 - **Definition:** (1) ill-defined goal; (2) personally important; (3) both destination & journey (understand search space) are essential

The problem: AI's Not Optimized For All Goals

- **AI is good at:** Factual questions
 - Optimized for the task; agreement on success; public leaderboards
 - $\leadsto$ Improving fast
- **AI's bad at:** “Search quests” (open-ended questions)
 - **Examples:** Paper topics, headphones, perfect wedding dress, unintended policy consequences, killer startup idea
 - **Definition:** (1) ill-defined goal; (2) personally important; (3) both destination & journey (understand search space) are essential
 - **Outputs must be:** relevant, diverse, creative, sustained

The problem: AI's Not Optimized For All Goals

- **AI is good at:** Factual questions
 - Optimized for the task; agreement on success; public leaderboards
 - $\leadsto$ Improving fast
- **AI's bad at:** “Search quests” (open-ended questions)
 - **Examples:** Paper topics, headphones, perfect wedding dress, unintended policy consequences, killer startup idea
 - **Definition:** (1) ill-defined goal; (2) personally important; (3) both destination & journey (understand search space) are essential
 - **Outputs must be:** relevant, diverse, creative, sustained
 - **Consequences:**

The problem: AI's Not Optimized For All Goals

- **AI is good at:** Factual questions
 - Optimized for the task; agreement on success; public leaderboards
 - $\leadsto$ Improving fast
- **AI's bad at:** “Search quests” (open-ended questions)
 - **Examples:** Paper topics, headphones, perfect wedding dress, unintended policy consequences, killer startup idea
 - **Definition:** (1) ill-defined goal; (2) personally important; (3) both destination & journey (understand search space) are essential
 - **Outputs must be:** relevant, diverse, creative, sustained
 - **Consequences:**
 - **Everyone:** homogenized, orthodox, conventional, majoritarian

The problem: AI's Not Optimized For All Goals

- **AI is good at:** Factual questions
 - Optimized for the task; agreement on success; public leaderboards
 - $\leadsto$ Improving fast
- **AI's bad at:** “Search quests” (open-ended questions)
 - **Examples:** Paper topics, headphones, perfect wedding dress, unintended policy consequences, killer startup idea
 - **Definition:** (1) ill-defined goal; (2) personally important; (3) both destination & journey (understand search space) are essential
 - **Outputs must be:** relevant, diverse, creative, sustained
 - **Consequences:**
 - **Everyone:** homogenized, orthodox, conventional, majoritarian
 - **You:**

The problem: AI's Not Optimized For All Goals

- **AI is good at:** Factual questions
 - Optimized for the task; agreement on success; public leaderboards
 - $\leadsto$ Improving fast
- **AI's bad at:** “Search quests” (open-ended questions)
 - **Examples:** Paper topics, headphones, perfect wedding dress, unintended policy consequences, killer startup idea
 - **Definition:** (1) ill-defined goal; (2) personally important; (3) both destination & journey (understand search space) are essential
 - **Outputs must be:** relevant, diverse, creative, sustained
 - **Consequences:**
 - **Everyone:** homogenized, orthodox, conventional, majoritarian
 - **You:** Better with AI

The problem: AI's Not Optimized For All Goals

- AI is good at: Factual questions
 - Optimized for the task; agreement on success; public leaderboards
 - $\leadsto$ Improving fast
- AI's bad at: “Search quests” (open-ended questions)
 - **Examples**: Paper topics, headphones, perfect wedding dress, unintended policy consequences, killer startup idea
 - **Definition**: (1) ill-defined goal; (2) personally important; (3) both destination & journey (understand search space) are essential
 - **Outputs must be**: relevant, diverse, creative, sustained
 - **Consequences**:
 - **Everyone**: homogenized, orthodox, conventional, majoritarian
 - **You**: Better with AI — Same as everyone else

The problem: AI's Not Optimized For All Goals

- **AI is good at:** Factual questions
 - Optimized for the task; agreement on success; public leaderboards
 - $\leadsto$ Improving fast
- **AI's bad at:** “Search quests” (open-ended questions)
 - **Examples:** Paper topics, headphones, perfect wedding dress, unintended policy consequences, killer startup idea
 - **Definition:** (1) ill-defined goal; (2) personally important; (3) both destination & journey (understand search space) are essential
 - **Outputs must be:** relevant, diverse, creative, sustained
 - **Consequences:**
 - **Everyone:** homogenized, orthodox, conventional, majoritarian
 - **You:** Better with AI — Same as everyone else
 - **When AI's better at facts:** worse at search quests

The problem: AI's Not Optimized For All Goals

- **AI is good at:** Factual questions
 - Optimized for the task; agreement on success; public leaderboards
 - $\leadsto$ Improving fast
- **AI's bad at:** “Search quests” (open-ended questions)
 - **Examples:** Paper topics, headphones, perfect wedding dress, unintended policy consequences, killer startup idea
 - **Definition:** (1) ill-defined goal; (2) personally important; (3) both destination & journey (understand search space) are essential
 - **Outputs must be:** relevant, diverse, creative, sustained
 - **Consequences:**
 - **Everyone:** homogenized, orthodox, conventional, majoritarian
 - **You:** Better with AI — Same as everyone else
 - **When AI's better at facts:** worse at search quests
 - **Different LLMs:** essentially the same answers

The problem: AI's Not Optimized For All Goals

- **AI is good at:** Factual questions
 - Optimized for the task; agreement on success; public leaderboards
 - $\leadsto$ Improving fast
- **AI's bad at:** “Search quests” (open-ended questions)
 - **Examples:** Paper topics, headphones, perfect wedding dress, unintended policy consequences, killer startup idea
 - **Definition:** (1) ill-defined goal; (2) personally important; (3) both destination & journey (understand search space) are essential
 - **Outputs must be:** relevant, diverse, creative, sustained
 - **Consequences:**
 - **Everyone:** homogenized, orthodox, conventional, majoritarian
 - **You:** Better with AI — Same as everyone else
 - **When AI's better at facts:** worse at search quests
 - **Different LLMs:** essentially the same answers
 - **Examples:**

The problem: AI's Not Optimized For All Goals

- **AI is good at:** Factual questions
 - Optimized for the task; agreement on success; public leaderboards
 - $\leadsto$ Improving fast
- **AI's bad at:** “Search quests” (open-ended questions)
 - **Examples:** Paper topics, headphones, perfect wedding dress, unintended policy consequences, killer startup idea
 - **Definition:** (1) ill-defined goal; (2) personally important; (3) both destination & journey (understand search space) are essential
 - **Outputs must be:** relevant, diverse, creative, sustained
 - **Consequences:**
 - **Everyone:** homogenized, orthodox, conventional, majoritarian
 - **You:** Better with AI — Same as everyone else
 - **When AI's better at facts:** worse at search quests
 - **Different LLMs:** essentially the same answers
 - **Examples:**
 - **Science, art, writing:** Individually more creative; collectively less

The problem: AI's Not Optimized For All Goals

- **AI is good at:** Factual questions
 - Optimized for the task; agreement on success; public leaderboards
 - $\leadsto$ Improving fast
- **AI's bad at:** “Search quests” (open-ended questions)
 - **Examples:** Paper topics, headphones, perfect wedding dress, unintended policy consequences, killer startup idea
 - **Definition:** (1) ill-defined goal; (2) personally important; (3) both destination & journey (understand search space) are essential
 - **Outputs must be:** relevant, diverse, creative, sustained
 - **Consequences:**
 - **Everyone:** homogenized, orthodox, conventional, majoritarian
 - **You:** Better with AI — Same as everyone else
 - **When AI's better at facts:** worse at search quests
 - **Different LLMs:** essentially the same answers
 - **Examples:**
 - **Science, art, writing:** Individually more creative; collectively less
 - **Large Class:** Papers with identical arguments

The problem: AI's Not Optimized For All Goals

- **AI is good at:** Factual questions
 - Optimized for the task; agreement on success; public leaderboards
 - $\leadsto$ Improving fast
- **AI's bad at:** “Search quests” (open-ended questions)
 - **Examples:** Paper topics, headphones, perfect wedding dress, unintended policy consequences, killer startup idea
 - **Definition:** (1) ill-defined goal; (2) personally important; (3) both destination & journey (understand search space) are essential
 - **Outputs must be:** relevant, diverse, creative, sustained
 - **Consequences:**
 - **Everyone:** homogenized, orthodox, conventional, majoritarian
 - **You:** Better with AI — Same as everyone else
 - **When AI's better at facts:** worse at search quests
 - **Different LLMs:** essentially the same answers
 - **Examples:**
 - **Science, art, writing:** Individually more creative; collectively less
 - **Large Class:** Papers with identical arguments
 - **US Govt:** How to get the uranium out of Iran?

The problem: AI's Not Optimized For All Goals

- **AI is good at:** Factual questions
 - Optimized for the task; agreement on success; public leaderboards
 - $\leadsto$ Improving fast
- **AI's bad at:** “Search quests” (open-ended questions)
 - **Examples:** Paper topics, headphones, perfect wedding dress, unintended policy consequences, killer startup idea
 - **Definition:** (1) ill-defined goal; (2) personally important; (3) both destination & journey (understand search space) are essential
 - **Outputs must be:** relevant, diverse, creative, sustained
 - **Consequences:**
 - **Everyone:** homogenized, orthodox, conventional, majoritarian
 - **You:** Better with AI — Same as everyone else
 - **When AI's better at facts:** worse at search quests
 - **Different LLMs:** essentially the same answers
 - **Examples:**
 - **Science, art, writing:** Individually more creative; collectively less
 - **Large Class:** Papers with identical arguments
 - **US Govt:** How to get the uranium out of Iran?
 - **Research:** AI for creative ideas; same ones as your competitors

Inducing Sustained Creativity and Diversity in LLMs

Inducing Sustained Creativity and Diversity in LLMs

- LLMs: encode all knowledge; decode to write your answer

Inducing Sustained Creativity and Diversity in LLMs

- LLMs: **encode** all knowledge; **decode** to write your answer
- What doesn't work:

Inducing Sustained Creativity and Diversity in LLMs

- LLMs: **encode** all knowledge; **decode** to write your answer
- What doesn't work:
 - Ordinary **decoding** (worse due to RLHF+majority vote, competition, training on LLM outputs, ...)

Inducing Sustained Creativity and Diversity in LLMs

- LLMs: **encode** all knowledge; **decode** to write your answer
- What doesn't work:
 - **Ordinary decoding** (worse due to RLHF+majority vote, competition, training on LLM outputs, ...)
 - **Prior Attempts** (changing prompts, higher “temperature”, small-batch diversity algorithms, post-training, fine-tuning, harnesses, ...)

Inducing Sustained Creativity and Diversity in LLMs

- LLMs: **encode** all knowledge; **decode** to write your answer
- What doesn't work:
 - **Ordinary decoding** (worse due to RLHF+majority vote, competition, training on LLM outputs, ...)
 - **Prior Attempts** (changing prompts, higher “temperature”, small-batch diversity algorithms, post-training, fine-tuning, harnesses, ...)
- What works:

Inducing Sustained Creativity and Diversity in LLMs

- LLMs: **encode** all knowledge; **decode** to write your answer
- What doesn't work:
 - **Ordinary decoding** (worse due to RLHF+majority vote, competition, training on LLM outputs, ...)
 - **Prior Attempts** (changing prompts, higher “temperature”, small-batch diversity algorithms, post-training, fine-tuning, harnesses, ...)
- What works:
 - **Recoding-Decoding** algorithm

Inducing Sustained Creativity and Diversity in LLMs

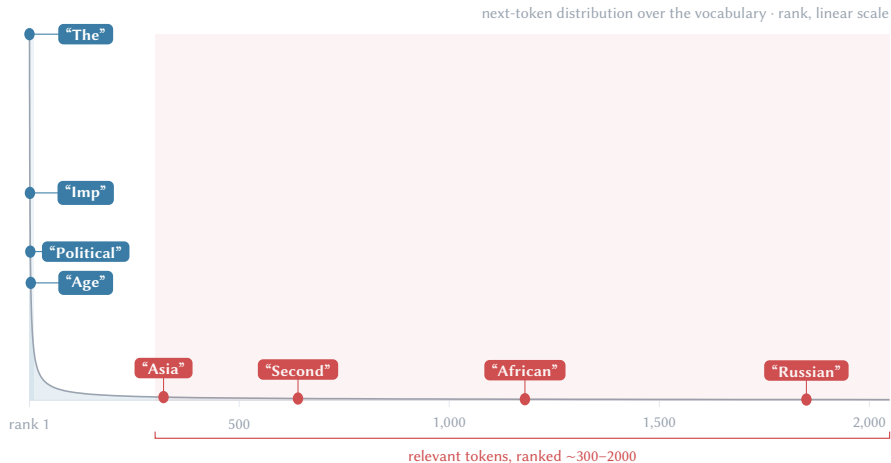
- LLMs: **encode** all knowledge; **decode** to write your answer
- What doesn't work:
 - **Ordinary decoding** (worse due to RLHF+majority vote, competition, training on LLM outputs, ...)
 - **Prior Attempts** (changing prompts, higher “temperature”, small-batch diversity algorithms, post-training, fine-tuning, harnesses, ...)
- What works:
 - **Recoding-Decoding** algorithm
 - **How it works:** Randomized diversion, LLM completion

Inducing Sustained Creativity and Diversity in LLMs

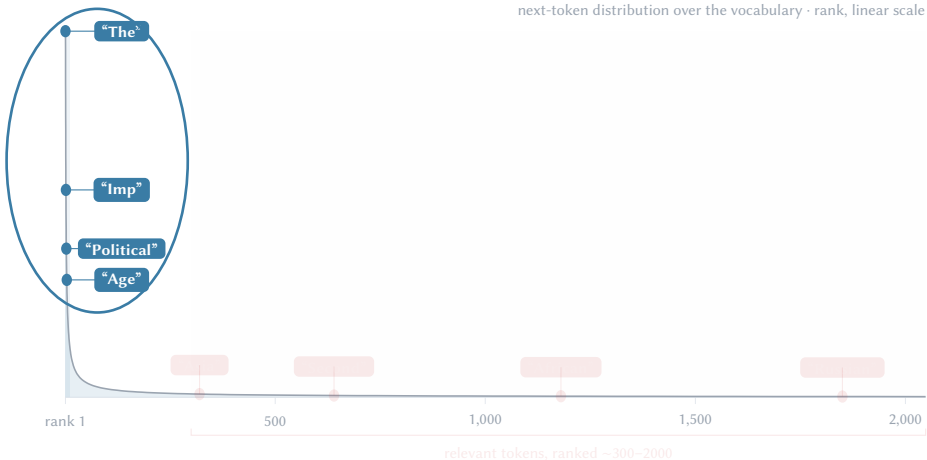
- **LLMs:** **encode** all knowledge; **decode** to write your answer
- **What doesn't work:**
 - **Ordinary decoding** (worse due to RLHF+majority vote, competition, training on LLM outputs, ...)
 - **Prior Attempts** (changing prompts, higher “temperature”, small-batch diversity algorithms, post-training, fine-tuning, harnesses, ...)
- **What works:**
 - **Recoding-Decoding** algorithm
 - **How it works:** Randomized diversion, LLM completion
 - **Easy to implement:** Built on any LLM (you can do it)

Brainstorm Book Topics on 18th Century World History

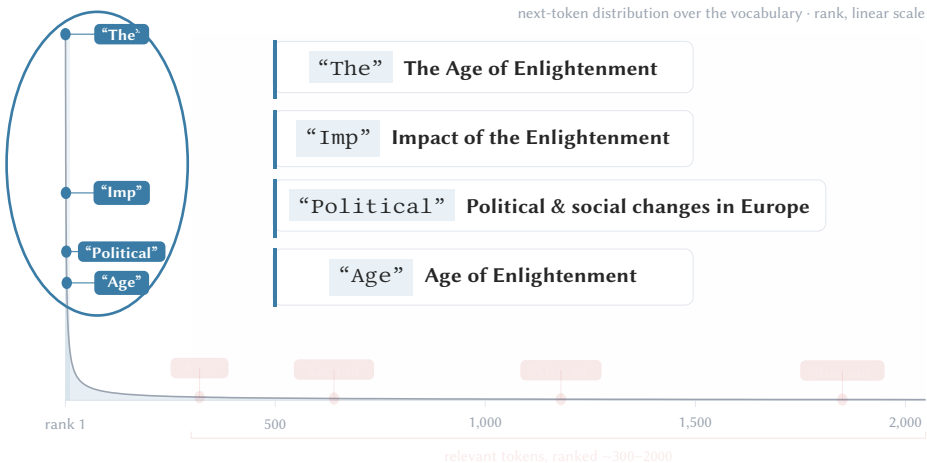
Brainstorm Book Topics on 18th Century World History



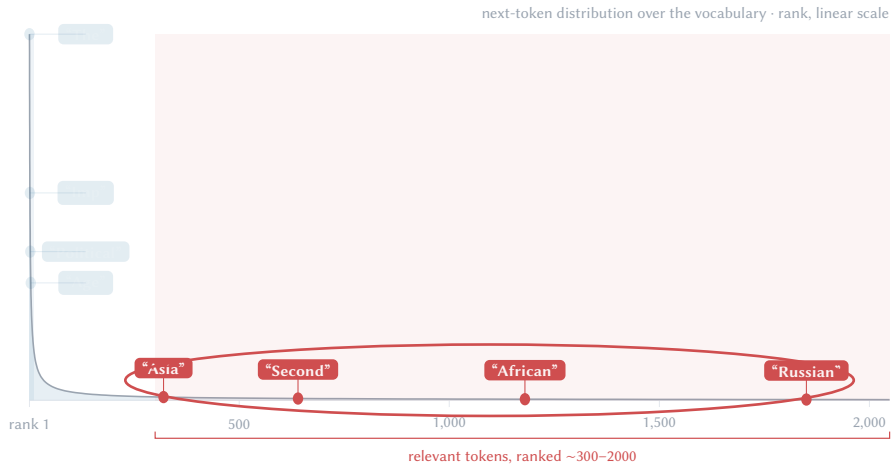
Brainstorm Book Topics on 18th Century World History



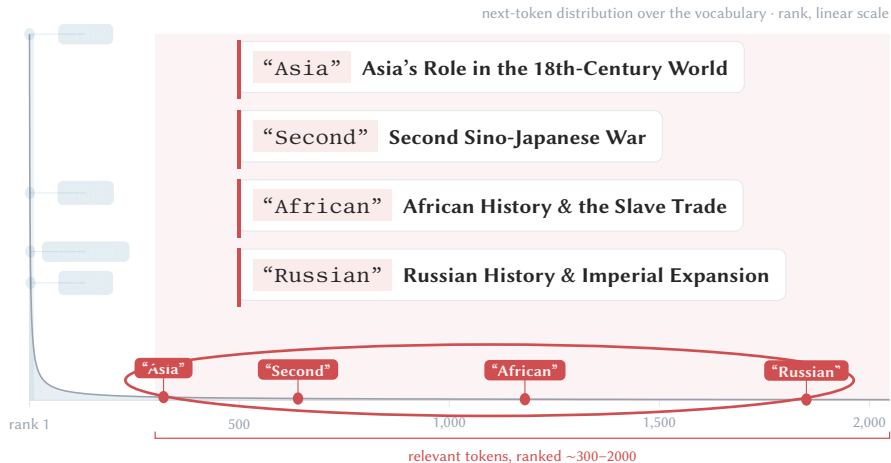
Brainstorm Book Topics on 18th Century World History



Brainstorm Book Topics on 18th Century World History



Brainstorm Book Topics on 18th Century World History



Recoding-Decoding, Step by Step

Recoding-Decoding, Step by Step

Brainstorm 2 world history topics.



Recoding-Decoding, Step by Step

Sampled Priming Phrase

Sampled Diverting Token

**Related to FOOD: ** Brainstorm 2 world history topics. Pas



Recoding-Decoding, Step by Step

Sampled Priming Phrase

Sampled Diverting Token

**Related to FOOD: ** Brainstorm 2 world history topics. Pas



Recoding-Decoding, Step by Step

Sampled Priming Phrase

Sampled Diverting Token

**Related to FOOD: ** Brainstorm 2 world history topics. Pas



t

Recoding-Decoding, Step by Step

Sampled Priming Phrase

Sampled Diverting Token

**Related to FOOD: ** Brainstorm 2 world history topics. Pas



ta and the

Recoding-Decoding, Step by Step

Sampled Priming Phrase

Sampled Diverting Token

**Related to FOOD: ** Brainstorm 2 world history topics. Pas



ta and the silk road.

Recoding-Decoding, Step by Step

Sampled Priming Phrase



**Related to FOOD: ** Brainstorm 2 world history topics. Pas

Sampled Diverting Token



ta and the silk road.

Recoding-Decoding, Step by Step

Sampled Priming Phrase

Sampled Diverting Token

**Related to FOOD: ** Brainstorm 2 world history topics. Pasta and the silk road.



Recoding-Decoding, Step by Step

Sampled Priming Phrase

Sampled Diverging Topic

Brainstorm 2 world history topics. Pasta and the silk road.



Recoding-Decoding, Step by Step

Replace with a New Priming Phrase

Sample a New Diverting Token

**Related to SKY: ** Brainstorm 2 world history topics. Pasta and the silk road. Tib



Recoding-Decoding, Step by Step

Replace with a New Priming Phrase

Sample a New Diverting Token

**Related to SKY: ** Brainstorm 2 world history topics. Pasta and the silk road. Tib



Recoding-Decoding, Step by Step

Replace with a New Priming Phrase

Sample a New Diverting Token

**Related to SKY: ** Brainstorm 2 world history topics. Pasta and the silk road. Tib



etan

Recoding-Decoding, Step by Step

Replace with a New Priming Phrase

Sample a New Diverting Token

**Related to SKY: ** Brainstorm 2 world history topics. Pasta and the silk road. Tib



etan sky burial

Recoding-Decoding, Step by Step

Replace with a New Priming Phrase

Sample a New Diverting Token

**Related to SKY: ** Brainstorm 2 world history topics. Pasta and the silk road. Tib



etan sky burial.

Recoding-Decoding, Step by Step

Replace with a New Priming Phrase

Sample a New Diverting Token

**Related to SKY: ** Brainstorm 2 world history topics. Pasta and the silk road. Tibetan sky burial.



Recoding-Decoding, Step by Step

Replace with a New Priming Phrase

Sample a New Diverting Token

**Related to SKY: ** Brainstorm 2 world history topics.

Pasta and the silk road. Tibetan sky burial.



Recoding-Decoding, Step by Step

Pasta and the silk road. Tibetan sky burial.

List interesting battlefields in world history

List interesting battlefields in world history

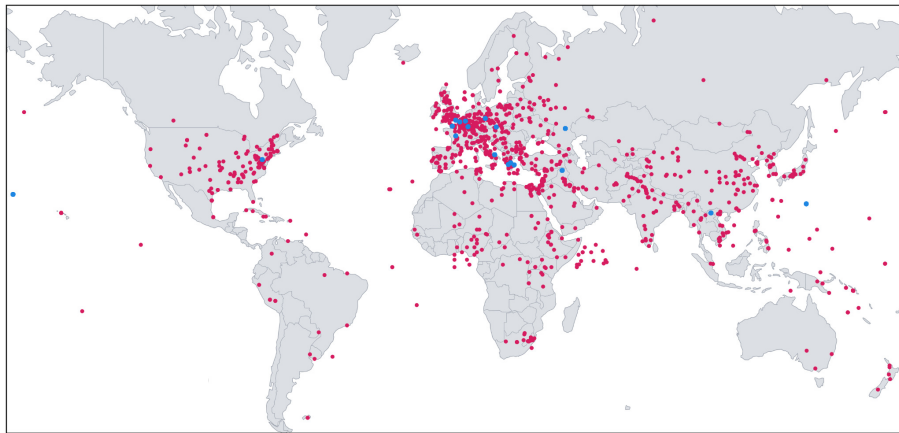


List interesting battlefields in world history



OD: Only Famous Battles — Gettysburg, Waterloo, Stalingrad, Marathon, etc.

List interesting battlefields in world history



OD: Only Famous Battles — Gettysburg, Waterloo, Stalingrad, Marathon, etc.

RD: (Real) Battles everywhere

Bridal Dress Design Ideas

Bridal Dress Design Ideas

OD



Bridal Dress Design Ideas

OD



RD



Bridal Dress Design Ideas

OD



RD



gender-neutral wedding outfit

Bridal Dress Design Ideas

OD



RD



Mongolian-style wedding dress

Bridal Dress Design Ideas

OD



RD



with music-note embroidery for a musician bride

Bridal Dress Design Ideas

OD



RD



wedding dress with optical fiber veil

Let's Remember

Let's Remember

- Your search quests:

Let's Remember

- Your search quests: Better with AI

Let's Remember

- Your search quests: Better with AI — Same as everyone else

Let's Remember

- Your search quests: Better with AI — Same as everyone else
- You can fix this!

Let's Remember

- Your search quests: Better with AI — Same as everyone else
- You can fix this!

Tell Claude: **Make a /quest skill based on GaryKing.org/quest**

Let's Remember

- Your search quests: Better with AI — Same as everyone else
- You can fix this!

Tell Claude: **Make a /quest skill based on GaryKing.org/quest**

Papers, Slides, Data: **GaryKing.org/quest**

Bouquet Design Ideas

Bouquet Design Ideas

OD

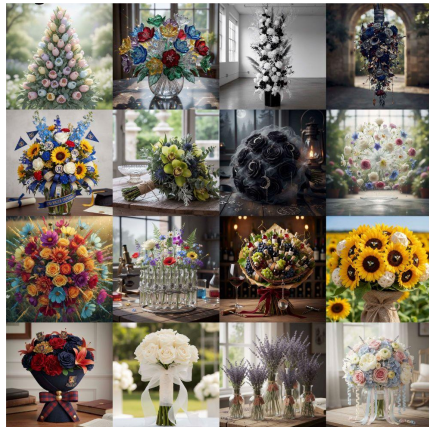


Bouquet Design Ideas

OD



RD



Halloween Party Themes

Halloween Party Themes

OD



Halloween Party Themes

OD



RD

