Generative AI vs Classical AI in Human-AI Teams : An Analysis of Synergy Effects

This talk looks at how people work with two kinds of AI—Generative AI (like tools that chat and draft) and Classical AI (more fixed, rules- or model-based tools). We ask a simple question: when does a human-AI team do better than the single best person in the room? Beating an average human is common; beating the best human is rare. That rare win is what we call strong synergy.

We analyzed over a hundred studies with several hundred measured outcomes.
• "Team vs best person" checks for strong synergy.
• "Team vs human baseline" and "Team vs AI baseline" check lighter, everyday gains.
• "Team vs worst person" is a rough safety test: are we avoiding truly bad outcomes?
To compare apples to apples, we sort tasks by who currently has the edge (AI or humans), and by task type (decision making vs creative work). We also separate Generative from Classical systems.

What we found.
• When AI already does well on a task, pairing people with Generative AI brings a clear lift: teams often surpass AI alone, a real sign of strong synergy. Classical tools, by contrast, tend to fall short of that bar in the same setting.
• Across all tasks, teams using Generative AI show a small but reliable advantage over the single best individual, while Classical AI teams typically do not. In short, generative systems raise the ceiling of what a team can achieve.
• When skilled humans have the edge, Generative AI can backfire—it may distract, add confident-sounding but wrong ideas, or slow precise workflows. Classical aides behave like focused assistants and bring modest, steady improvements.
• Risk profile: Generative setups deliver bigger upside, but their safety margin is thinner. Results vary more, so guardrails matter.

How to use this.
Lead with Generative AI on tasks where AI already looks strong and the flow is "draft → edit" (ideation, writing, coding). Let people verify, edit, and handle exceptions. Prefer Classical AI where humans already excel; surface short, clear signals (highlights, alerts, confidence cues). Add guardrails everywhere: show uncertainty, ask "what would make this wrong?", trigger human checks when stakes are high, and monitor outcomes against best- and worst-case references to catch drift and negative synergy early. The big lesson is not "use generative AI everywhere," but match the tool to the task to capture upside while containing risk.

인간-AI 팀에서 생성형 AI vs 고전적 AI: 시너지 효과 분석

국문 요약
이 발표는 사람들이 생성형 AI(대화하며 초안을 써 주는 도구)와 고전적 AI(규칙이나 고정 모델 기반 도구)와 함께 일할 때, 팀이 최고 실력자보다 더 잘하는지를 분석하였습니다.

백여 편이 넘는 연구와 수백 건의 데이터를 분석했습니다.

• “팀 vs 최고 개인”은 강한 시너지가 있는지 보는 기준입니다.
• “팀 vs 인간 기준”, “팀 vs AI 기준”은 일상적인 개선이 있는지 확인합니다.
• “팀 vs 최약체”는 안전여유를 보는 거친 점검입니다. 아주 나쁜 결과를 피하고 있는지 보는 것이죠.
비교를 공정하게 하기 위해, 과업을 AI가 유리한지/사람이 유리한지로 나누고, 의사결정형 vs 창작형 같은 작업 유형도 구분했습니다. 또한 생성형과 고전적 시스템을 따로 보았습니다.

결과 요약
• AI가 이미 잘하는 과업에서는 생성형 AI와 팀을 이루는 방식이 명확한 상승 효과를 보였습니다. 팀이 AI 단독보다 더 잘하는 경우가 자주 나타납니다. 같은 환경에서 고전적 도구만 쓴 팀은 이 기준을 넘기 어려웠습니다.
• 전체적으로 보면, 생성형 AI를 쓰는 팀은 최고 개인을 약간이나마 꾸준히 앞서는 경향이 있습니다. 반면 고전적 AI 팀은 이 수준에 도달하지 못하는 경우가 많았습니다. 즉, 생성형 AI는 팀 성과의 상한선을 끌어올립니다.
• 사람이 더 잘하는 과업에서는 생성형 AI가 오히려 방해가 될 수 있습니다. 그럴듯하지만 틀린 방향으로 이끌거나, 정밀한 흐름을 깨뜨리기 쉽습니다. 반대로 고전적 보조도구는 짧고 분명한 신호를 주면서 소폭이지만 안정적인 개선을 냅니다.
• 위험 측면에서 생성형은 큰 기회가 있지만 안전결과의 변동폭이 큽니다. 그래서 안전장치가 중요합니다.

실무 적용
AI가 우수하고 ‘초안 → 편집’ 흐름인 작업(아이디어, 글, 코드)에서는 생성형 AI를 우선적으로 도입하고, 사람은 검증, 편집, 예외 처리에 집중해야 합니다. 반대로 사람이 이미 우수한 영역은 고전적 AI를 우선적으로 활용해야 합니다. 전반적으로는 불확실성 표시, 반증 질문(“이게 틀리려면?”), 고위험 시 사람 개입, 성과 모니터링을 통해 부정적인 시너지 효과를 초기에 예방하는 것이 핵심입니다. 결론은 “어디서나 생성형”이 아니라, 과업에 맞는 도구 선택입니다.