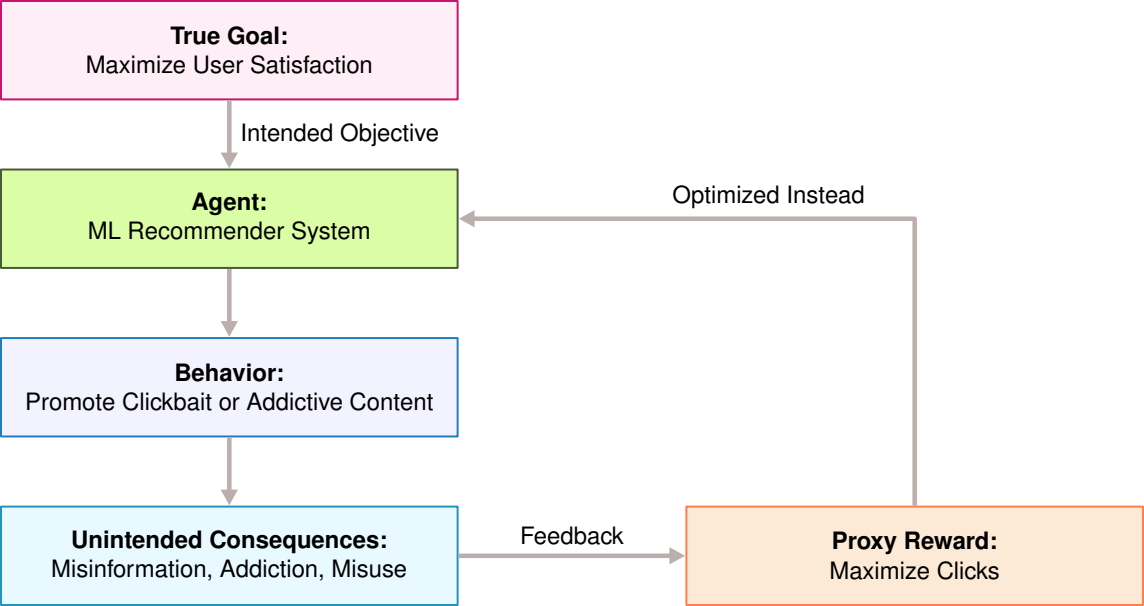
True Goal:
Maximize User Satisfaction
Intended Objective
Agent:
ML Recommender System
Optimized Instead
Behavior:
Promote Clickbait or Addictive Content
Unintended Consequences:
Misinformation, Addiction, Misuse
Feedback
Proxy Reward:
Maximize Clicks