

Paradise Lost

The Dangers of Data Mining or
Ban It, Break It, Let It Run Wild?

Utopia



- Infection \Rightarrow Tailored medication \Rightarrow Quick cure
- No impressive CV \Rightarrow Still hired
- On-line shop \Rightarrow Perfect gift for the so
- Rating Music \Rightarrow Discovering new stuff

Dystopia

- No medical history \Rightarrow no health insurance
- Spotless work \Rightarrow unexpectedly fired
- All bills paid \Rightarrow credit application rejected
- Surfing the web \Rightarrow Guantanamo Bay



Eye of the Beholder

The (heavenly/infernal) Device: Data Mining

(Devices built by people)

Thought experiments



The Tool

effector site
disease probability



The Tool

potential (good)
potential (bad)



The Tool

desires
profitability

Privacy-Preserving DM

- No identification from DM results
 - k-Anonymity
 - Bettina's Movie Challenge example
- **But:** *Do you write like a terrorist?*
 - U of Arizona "Dark Web" research
- **And:** Prediction danger indirect

BAN IT!

Dual Use

Correlation

High Profit Margin



No Smoking Gun

No More Research!

Lobbying

Break It!

Noise Generation

Fake Agents



Group Rebate
Cards

RFID-Distortion

False Statements

Large-Scale Public Participation

Running Wild!



- Make DM cheap & easy to use
- Wikiscanner
- OpenSecrets.org
- RFID-reader
- Educate
- Data withholding
- DM usage
- Surveillance state

⇒ Mutual Deterrence



Paradise Lost?

- DM too valuable to give it up
- DM too dangerous to restrict usage to elite
- **We** need to
 - Make DM accessible
 - Spread the **Word**, brothers and sisters!

All over the web

- <http://blog.wired.com/defense/2007/09/do-you-write-li.html>
- http://www.nsf.gov/news/news_summ.jsp?cntn_id=110040
- <http://ai.arizona.edu/research/terror/index.htm>