How to Fail: a Subjective Disaster

Gerhard Weikum
http://www.mpi-inf.mpg.de/~weikum/
How to Fail: 5 Golden Rules

Rule 1: Start with list of 10 topics and hire 10+ students

Rule 2: If you don’t make progress hire more students

Rule 3: If you run out of money write more grant proposals

Rule 4: Be visible everywhere: 5+ papers in DB, 5+ in DM&ML, …

Rule 5: Don’t miss any opportunities: PCs, special issues, panels, …

Overcommitment is our professional curse! Learn how to say No!
Subjective Advice #1

Be passionate and committed!

„There are only two mistakes one can make along the road to truth: not going all the way, and not starting.”

(Buddha)
Subjective Advice #2

Be bold, take risks!
No risk, no gain!

„Avoiding failure is not the path to success“

(Simon Pegg in „Hector and the Search for Happiness“)
Subjective Advice #3

Seek advice, but think for yourself!

„Don‘t follow the advice of senior people only because they have grey hair“

„QA: Question Authority“

(Yours Truly)

(Jim Gray)
What are Good Research Directions: Compass and Map for Research Taste
Key Choices and Trade-Offs

<table>
<thead>
<tr>
<th>Difficulty</th>
<th>Popularity</th>
</tr>
</thead>
<tbody>
<tr>
<td>high</td>
<td>low</td>
</tr>
<tr>
<td>low</td>
<td>high</td>
</tr>
</tbody>
</table>

- **off the beaten path**
  - exciting: trendsetting opportunity
  - underexplored but perhaps boring

- **main-stream**
  - still interesting: be early & fast to beat the crowds (or be better)
  - too easy
My Personal Choices

Difficulty

off the beaten path

low

high

main-stream

Popularity

DB Auto Tuning

DB&IR

Knowledge Harvesting

Transactions

Multimedia QoS

Workflow management
Criteria for Good Topics

• **Significant benefit:** people care about solution
• **Difficult:** no obvious solution
• **Reachable:** hopefully solvable
• **Methodologically open:**
  variety of approaches conceivable

• **Simple to state:**
  can be explained in few minutes to any scientist
• **Incrementally solvable:**
  can pursue intermediate milestones
• **Testable:**
  can measure progress and success
# Meeting Criteria for Good Topics

## DB Auto Tuning:

- **Beneficial:** overcome DBA bottleneck, reduce $$$ cost
- **Difficult:** automate human expertise, considered daunting
- **Reachable:** automate at least special tasks
- **Open:** heuristics, combinat opt, stochastic model, control theory, …
- **Explainable:** by analogies (car) & specific cases (indexes, load control)
- **Incremental:** approximate DBA, focus on special issues
- **Testable:** competitiveness, workload traces, dynamic load generators

## Making Sense of Web Tables:

- **Beneficial:** a lot of interesting contents
- **Difficult:** ultimate form of heterogeneous data integration
- **Reachable:** partial success already big progress
- **Open:** declarative programs, machine learning, crowdsourcing, …
- **Explainable:** put structured Web in an integrated DB
- **Incremental:** good enough for search, domain-specific tables, …
- **Testable:** test corpora & tasks, comparison to human
Not Meeting (All) Criteria for Good Topics

Semantic Desktop:
- **Beneficial**: perhaps, but need to consider user tasks, not just the data
- **Explainable**: “envision a semantic web just for yourself “ … ?
- **Incremental**: semi-semantic desktop?
- **Testable**: user studies difficult, usage logs hardly available, …

Web-Scale Graph Algorithms with Map-Reduce:
- **Difficult**: $10^9 \times 10^3 \times 10 \times 0.1$ ☑ 1 TB memory ☑ 10K$
- **Reachable**: easy on paths, daunting on Steiner trees
- **Open**: already sold on one method
- **Explainable**: already lost on explaining MR
Take-Home Message

Success requires: skills, creativity, fast learning, insight, stamina, luck, ... good research taste

“It was much nicer before people started storing all their personal data in the cloud.“

“I don’t know if this is such a wise thing to do, Julia.“

- go mainstream
- must be fast (or better)

- go off the beaten path
  - more risk, more fun,
  - more long-term gain