

Denis Türker

**The Optimal Design of a Search Engine
from an Agency Theory Perspective***

List of Figures	IV
Abbreviations	V
1. Introduction	1
2. Web Information Retrieval.....	3
2.1. Overview of IR-Models	3
2.1.1. Retrieval: Ad Hoc and Filtering	3
2.1.2. Models for Browsing	6
2.2. Web Searching.....	6
2.2.1. Directories	9
2.2.2. Traditional Search Engines	9
2.2.3. Niche Search Engines	9
2.2.4. Metasearch Engines.....	10
2.2.5. Comparison Shopping Engines	10
2.2.6. Web Question Answering Systems	11
3. Search Engines	13
3.1. Challenges	13
3.1.1. Spam	13
3.1.2. Content Quality.....	13
3.1.3. Quality Evaluation.....	14
3.1.4. Invisible Web	14

* Slightly modified version of a thesis that was accepted by the economic faculty of the University of Cologne in spring term 2004.



3.2. Search Engine Architecture	14
3.2.1. Crawler Module.....	15
3.2.2. Indexer.....	16
3.2.3. Document Preprocessing.....	16
3.2.4. Query Module	17
3.2.5. Ranking Module.....	18
3.3. Ranking Techniques	18
3.3.1. Internal Content	18
3.3.2. Usage Information.....	19
3.3.3. Link-Based.....	19
3.4. Trends.....	21
3.4.1. Usability, Visualization, and User Interface.....	21
3.4.2. Semantic Web.....	22
3.4.3. Comparison Shopping Enhancements.....	22
3.4.4. Query Reformulation and Cross-Language Retrieval	23
3.4.5. Hybrid Search	24
3.4.6. Natural Language Processing and Web Question Answering ..	24
3.4.7. Personalized Search.....	24
3.4.8. Local Search.....	25
3.4.9. Social Networks	26
3.5. Revenue Sources	27
3.5.1. Direct Revenues from Users	28
3.5.2. Indirect Revenues	28
4. The Information Market.....	31
4.1. The Information Economy	31
4.2. Transaction Costs.....	31
4.3. Economics of Attention	32
4.4. Mass Media	33
4.5. Diversity of Opinions.....	35

5. Search Engines from an Agency Theory Perspective	37
5.1. The Basic Concept of Agency Theory	37
5.1.1. Main Concept and Assumptions	37
5.1.2. Agency Problems	38
5.1.3. Solutions to Agency Problems	38
5.2. The Optimal Delegation of Power	39
5.3. Power Roles	40
5.4. Design Patterns	43
5.4.1. A Trilateral Relationship Pattern	43
5.4.2. Direct Relationship Pattern	47
5.5. Design Strategies	48
5.5.1. Concentration Strategy	48
5.5.2. Niche Strategy	52
5.5.3. Bundling Strategy	54
5.5.4. Premium Search	55
5.6. Choice of Strategy	56
5.6.1. Goals and Compatible Strategies	56
5.6.2. Circumstances Restricting Choice	58
6. Conclusions	61
Bibliography	63



List of Figures

No. Contents	Page
1 Information Retrieval Models	4
2 Major Search Engines: Who Powers Whom?	7
3 Share of Search Referrals (March 2004)	8
4 Referral Trends of Major Search Engines.....	8
5 Search Engine Architecture	15
6 Revenue Sources	27
7 Trilateral Relationship in Traditional Advertising-Financed Media.....	34
8 Trilateral Relationship of Advertising-Financed Search Engines.....	35
9 Agency Problems.....	38
10 Optimal Level of Delegation.....	40
11 Trilateral Relationship	43
12 Direct Relationship.....	47
13 Google and a Concentration Strategy.....	49
14 Yahoo and a Concentration Strategy	51
15 A9 and a Niche Strategy	53
16 Premium Search Strategy	55
17 Choice of Strategy	56
18 Goals and Strategies	57
19 Circumstances Restricting Choice	59