

網際網路上物件導向式資源查詢系統

Object-Oriented Resource Query System in Internet

吳宜鴻、陳彥寰、陳良弼

國立清華大學資訊科學系

Outline

- **Introduction**
- **System Architecture**
- **Methodology of Classification**
- **Query Processing**
- **Conclusion**

Introduction

- **Motivation**

- ☞ **internet resource discovery**

- ➔ WAIS, Archie, Gopher, Yahoo, YamWeb, GAIS

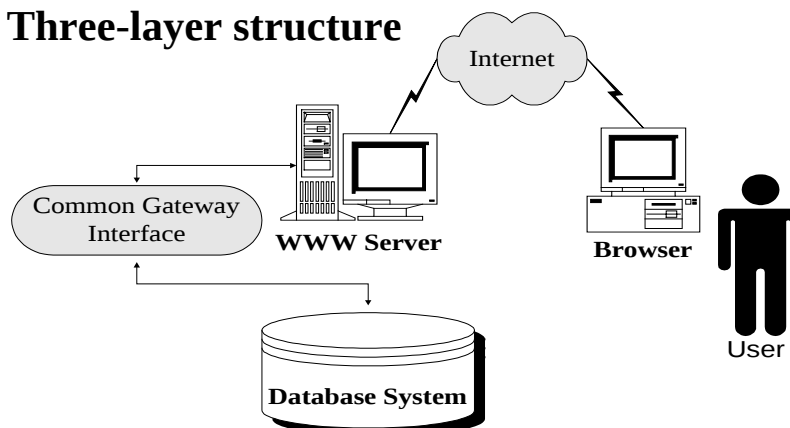
- **Goal**

- ☞ **provides a powerful facility to help users search their target resources**

- ➔ integrates resources with database querying

System Architecture

- **Three-layer structure**



System Architecture

- **Internet resource query system**
 - ☞ **resource collection**
 - ➔ **robot program**
 - ➔ **registration form**
 - ☞ **resource classification**
 - ☞ **resource query processing**

Resource Collection

```
graph TD; H1[Homepage] --> R[Robot]; R --> HC[/Homepage Cache/]; HC --> HP[Homepage Parser]; HP --> MS[/Metadata storage/]; MS --> C{Classification}; C --> DB[(Database)]; DB --> Reg[/Registration/]; Reg --> H2[Homepage]; R --> F1[ ]; F1 --> F2[ ]; F2 --> F3[ ]; F3 --> F4[ ]; F4 --> F5[ ]; F5 --> F6[ ]; F6 --> F7[ ]; F7 --> F8[ ]; F8 --> F9[ ]; F9 --> F10[ ]; F10 --> F11[ ]; F11 --> F12[ ]; F12 --> F13[ ]; F13 --> F14[ ]; F14 --> F15[ ]; F15 --> F16[ ]; F16 --> F17[ ]; F17 --> F18[ ]; F18 --> F19[ ]; F19 --> F20[ ]; F20 --> F21[ ]; F21 --> F22[ ]; F22 --> F23[ ]; F23 --> F24[ ]; F24 --> F25[ ]; F25 --> F26[ ]; F26 --> F27[ ]; F27 --> F28[ ]; F28 --> F29[ ]; F29 --> F30[ ]; F30 --> F31[ ]; F31 --> F32[ ]; F32 --> F33[ ]; F33 --> F34[ ]; F34 --> F35[ ]; F35 --> F36[ ]; F36 --> F37[ ]; F37 --> F38[ ]; F38 --> F39[ ]; F39 --> F40[ ]; F40 --> F41[ ]; F41 --> F42[ ]; F42 --> F43[ ]; F43 --> F44[ ]; F44 --> F45[ ]; F45 --> F46[ ]; F46 --> F47[ ]; F47 --> F48[ ]; F48 --> F49[ ]; F49 --> F50[ ]; F50 --> F51[ ]; F51 --> F52[ ]; F52 --> F53[ ]; F53 --> F54[ ]; F54 --> F55[ ]; F55 --> F56[ ]; F56 --> F57[ ]; F57 --> F58[ ]; F58 --> F59[ ]; F59 --> F60[ ]; F60 --> F61[ ]; F61 --> F62[ ]; F62 --> F63[ ]; F63 --> F64[ ]; F64 --> F65[ ]; F65 --> F66[ ]; F66 --> F67[ ]; F67 --> F68[ ]; F68 --> F69[ ]; F69 --> F70[ ]; F70 --> F71[ ]; F71 --> F72[ ]; F72 --> F73[ ]; F73 --> F74[ ]; F74 --> F75[ ]; F75 --> F76[ ]; F76 --> F77[ ]; F77 --> F78[ ]; F78 --> F79[ ]; F79 --> F80[ ]; F80 --> F81[ ]; F81 --> F82[ ]; F82 --> F83[ ]; F83 --> F84[ ]; F84 --> F85[ ]; F85 --> F86[ ]; F86 --> F87[ ]; F87 --> F88[ ]; F88 --> F89[ ]; F89 --> F90[ ]; F90 --> F91[ ]; F91 --> F92[ ]; F92 --> F93[ ]; F93 --> F94[ ]; F94 --> F95[ ]; F95 --> F96[ ]; F96 --> F97[ ]; F97 --> F98[ ]; F98 --> F99[ ]; F99 --> F100[ ]; F100 --> F101[ ]; F101 --> F102[ ]; F102 --> F103[ ]; F103 --> F104[ ]; F104 --> F105[ ]; F105 --> F106[ ]; F106 --> F107[ ]; F107 --> F108[ ]; F108 --> F109[ ]; F109 --> F110[ ]; F110 --> F111[ ]; F111 --> F112[ ]; F112 --> F113[ ]; F113 --> F114[ ]; F114 --> F115[ ]; F115 --> F116[ ]; F116 --> F117[ ]; F117 --> F118[ ]; F118 --> F119[ ]; F119 --> F120[ ]; F120 --> F121[ ]; F121 --> F122[ ]; F122 --> F123[ ]; F123 --> F124[ ]; F124 --> F125[ ]; F125 --> F126[ ]; F126 --> F127[ ]; F127 --> F128[ ]; F128 --> F129[ ]; F129 --> F130[ ]; F130 --> F131[ ]; F131 --> F132[ ]; F132 --> F133[ ]; F133 --> F134[ ]; F134 --> F135[ ]; F135 --> F136[ ]; F136 --> F137[ ]; F137 --> F138[ ]; F138 --> F139[ ]; F139 --> F140[ ]; F140 --> F141[ ]; F141 --> F142[ ]; F142 --> F143[ ]; F143 --> F144[ ]; F144 --> F145[ ]; F145 --> F146[ ]; F146 --> F147[ ]; F147 --> F148[ ]; F148 --> F149[ ]; F149 --> F150[ ]; F150 --> F151[ ]; F151 --> F152[ ]; F152 --> F153[ ]; F153 --> F154[ ]; F154 --> F155[ ]; F155 --> F156[ ]; F156 --> F157[ ]; F157 --> F158[ ]; F158 --> F159[ ]; F159 --> F160[ ]; F160 --> F161[ ]; F161 --> F162[ ]; F162 --> F163[ ]; F163 --> F164[ ]; F164 --> F165[ ]; F165 --> F166[ ]; F166 --> F167[ ]; F167 --> F168[ ]; F168 --> F169[ ]; F169 --> F170[ ]; F170 --> F171[ ]; F171 --> F172[ ]; F172 --> F173[ ]; F173 --> F174[ ]; F174 --> F175[ ]; F175 --> F176[ ]; F176 --> F177[ ]; F177 --> F178[ ]; F178 --> F179[ ]; F179 --> F180[ ]; F180 --> F181[ ]; F181 --> F182[ ]; F182 --> F183[ ]; F183 --> F184[ ]; F184 --> F185[ ]; F185 --> F186[ ]; F186 --> F187[ ]; F187 --> F188[ ]; F188 --> F189[ ]; F189 --> F190[ ]; F190 --> F191[ ]; F191 --> F192[ ]; F192 --> F193[ ]; F193 --> F194[ ]; F194 --> F195[ ]; F195 --> F196[ ]; F196 --> F197[ ]; F197 --> F198[ ]; F198 --> F199[ ]; F199 --> F200[ ]; F200 --> F201[ ]; F201 --> F202[ ]; F202 --> F203[ ]; F203 --> F204[ ]; F204 --> F205[ ]; F205 --> F206[ ]; F206 --> F207[ ]; F207 --> F208[ ]; F208 --> F209[ ]; F209 --> F210[ ]; F210 --> F211[ ]; F211 --> F212[ ]; F212 --> F213[ ]; F213 --> F214[ ]; F214 --> F215[ ]; F215 --> F216[ ]; F216 --> F217[ ]; F217 --> F218[ ]; F218 --> F219[ ]; F219 --> F220[ ]; F220 --> F221[ ]; F221 --> F222[ ]; F222 --> F223[ ]; F223 --> F224[ ]; F224 --> F225[ ]; F225 --> F226[ ]; F226 --> F227[ ]; F227 --> F228[ ]; F228 --> F229[ ]; F229 --> F230[ ]; F230 --> F231[ ]; F231 --> F232[ ]; F232 --> F233[ ]; F233 --> F234[ ]; F234 --> F235[ ]; F235 --> F236[ ]; F236 --> F237[ ]; F237 --> F238[ ]; F238 --> F239[ ]; F239 --> F240[ ]; F240 --> F241[ ]; F241 --> F242[ ]; F242 --> F243[ ]; F243 --> F244[ ]; F244 --> F245[ ]; F245 --> F246[ ]; F246 --> F247[ ]; F247 --> F248[ ]; F248 --> F249[ ]; F249 --> F250[ ]; F250 --> F251[ ]; F251 --> F252[ ]; F252 --> F253[ ]; F253 --> F254[ ]; F254 --> F255[ ]; F255 --> F256[ ]; F256 --> F257[ ]; F257 --> F258[ ]; F258 --> F259[ ]; F259 --> F260[ ]; F260 --> F261[ ]; F261 --> F262[ ]; F262 --> F263[ ]; F263 --> F264[ ]; F264 --> F265[ ]; F265 --> F266[ ]; F266 --> F267[ ]; F267 --> F268[ ]; F268 --> F269[ ]; F269 --> F270[ ]; F270 --> F271[ ]; F271 --> F272[ ]; F272 --> F273[ ]; F273 --> F274[ ]; F274 --> F275[ ]; F275 --> F276[ ]; F276 --> F277[ ]; F277 --> F278[ ]; F278 --> F279[ ]; F279 --> F280[ ]; F280 --> F281[ ]; F281 --> F282[ ]; F282 --> F283[ ]; F283 --> F284[ ]; F284 --> F285[ ]; F285 --> F286[ ]; F286 --> F287[ ]; F287 --> F288[ ]; F288 --> F289[ ]; F289 --> F290[ ]; F290 --> F291[ ]; F291 --> F292[ ]; F292 --> F293[ ]; F293 --> F294[ ]; F294 --> F295[ ]; F295 --> F296[ ]; F296 --> F297[ ]; F297 --> F298[ ]; F298 --> F299[ ]; F299 --> F300[ ]; F300 -->
```

Methodology of Classification

- Resource classification

- ☞ classification hierarchy

- ☞ classify method

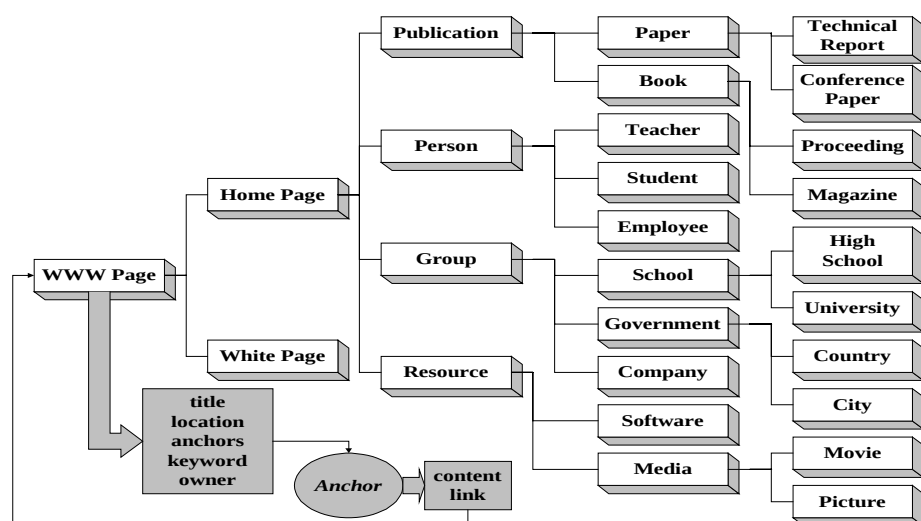
- ➔ semantic analysis

- ⇒ learning stage

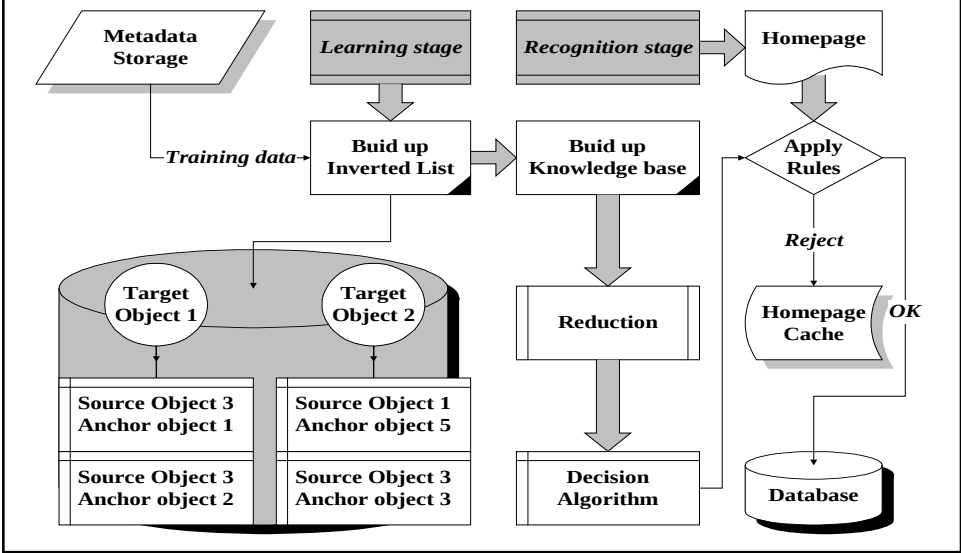
- ⇒ recognition stage

- ➔ syntactic analysis

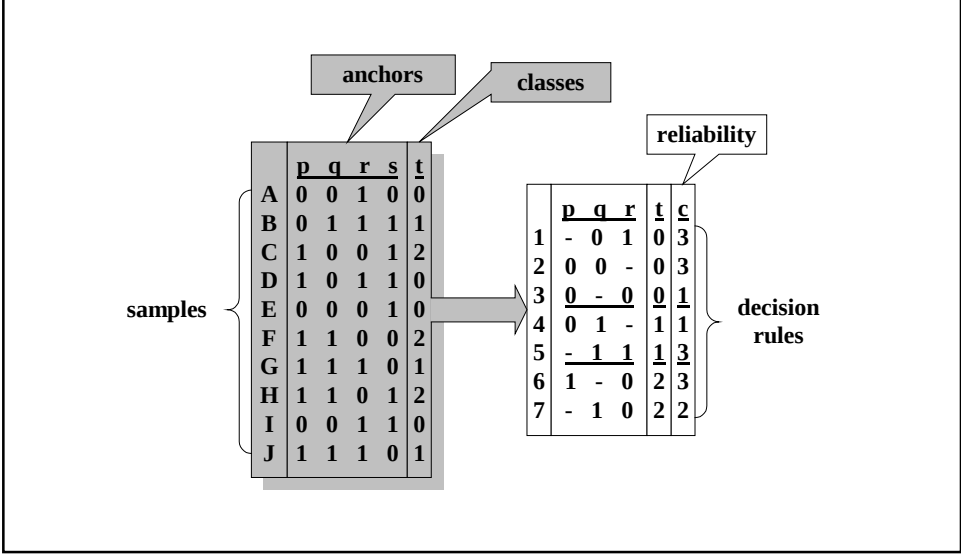
Classification Hierarchy



Semantic Analysis



Rule Reduction



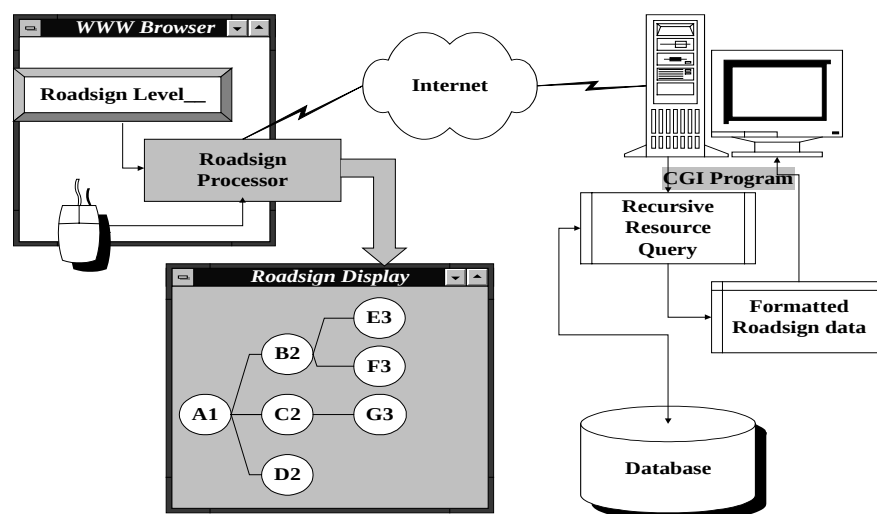
Query processing

- Query translation

☞ **SELECT** y **FROM** root* x, root* y **WHERE** x.location='http://www.cs.nthu.edu.tw' **and** x[level 3:5.anchors.link] **has** y **and** y.keyword **has** ['WWW', 'Database']

☞ **SELECT** y **FROM** root* x, root* y **WHERE** x.location='http://www.cs.nthu.edu.tw' **and** *is_recv_linked*(x,y,3,5)<>0 **and** y.keyword **has** ['WWW', 'Database']

Roadsign



Conclusion

- **Contribution**

- ☞ provide an effective classification scheme to represent the resources in Internet
- ☞ provide a query interface and roadsign service

- **Future work**

- ☞ rule refinement for real data