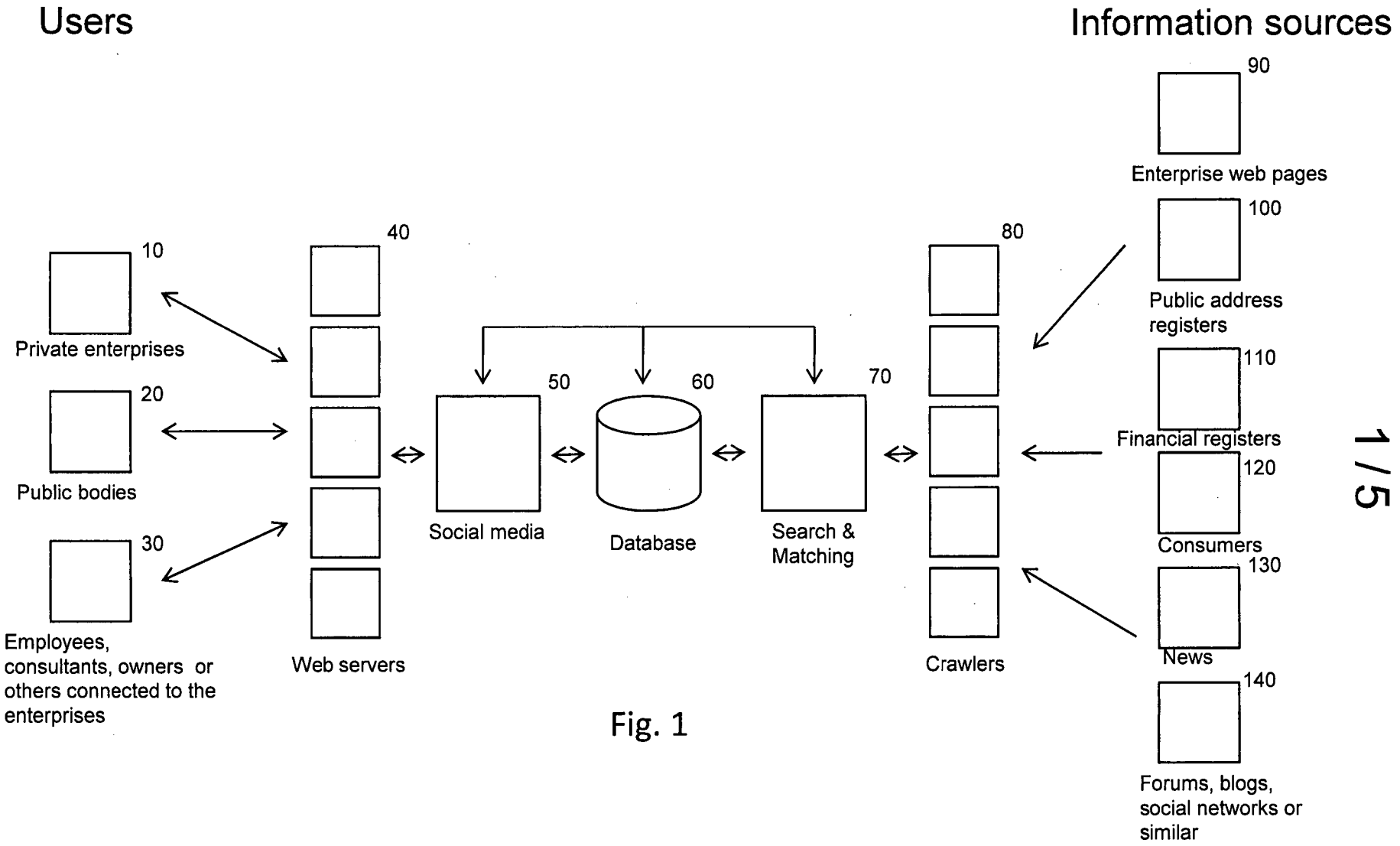

Users
Information sources
10
Private enterprises
20
Public bodies
30
Employees, consultants, owners or others connected to the enterprises
40
Web servers
50
Social media
60
Database
70
Search & Matching
80
Crawlers
90
Enterprise web pages
100
Public address registers
110
Financial registers
120
Consumers
130
News
140
Forums, blogs, social networks or similar

Fig. 1

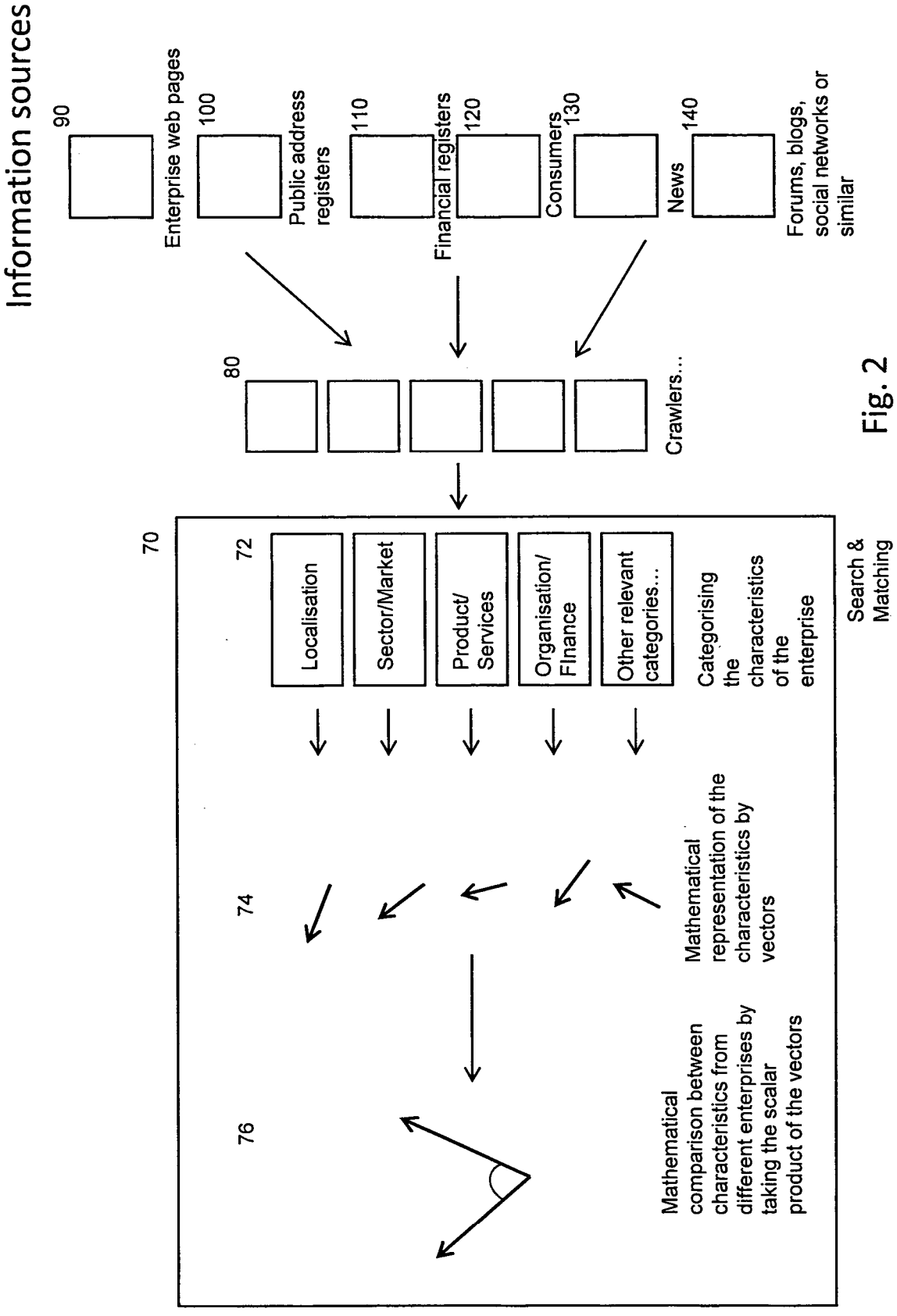
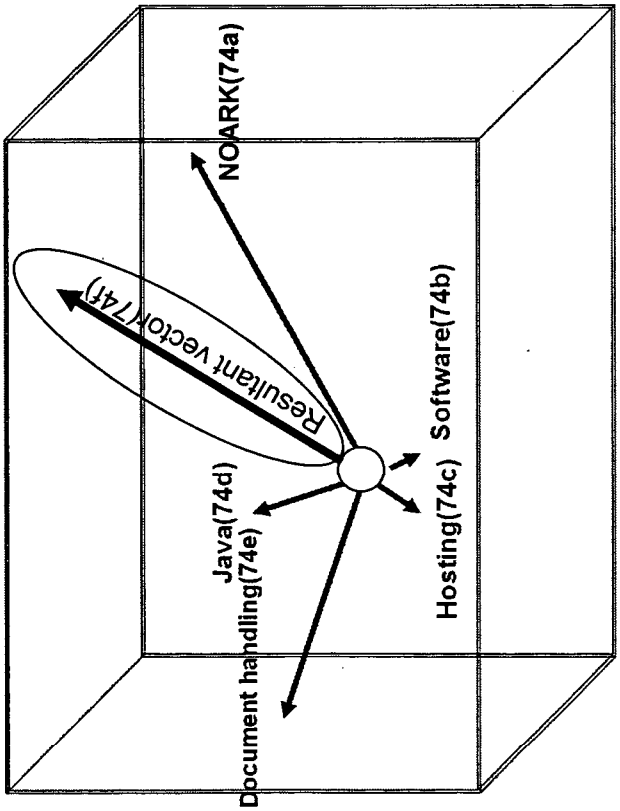

Information sources
90
Enterprise web pages
100
Public address registers
110
Financial registers
120
Consumers
130
News
140
Forums, blogs, social networks or similar
80
Crawlers...
70
72
Localisation
Sector/Market
Product/ Services
Organisation/ FInance
Other relevant categories...
Categorising the characteristics of the enterprise
Search & Matching
74
Mathematical representation of the characteristics by vectors
76
Mathematical comparison between characteristics from different enterprises by taking the scalar product of the vectors

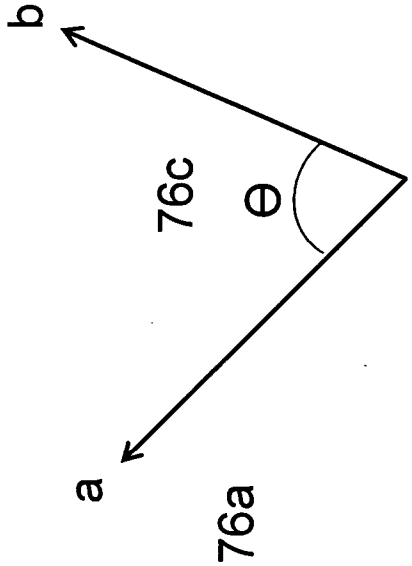
Fig. 2



Adaptive word list(74g)

Word combination	Word appearance	Word uniqueness
NOARK	100	10
Software	2000	2
Hosting	1500	3
Document handling	500	7
JAVA	1000	4

Fig 3



76d

$$\cos\theta = \frac{a_1b_1+a_2b_2+a_3b_3}{|\vec{a}|\,|\vec{b}|} = \frac{\vec{a}\cdot\vec{b}}{|\vec{a}|\,|\vec{b}|}$$

Fig. 4

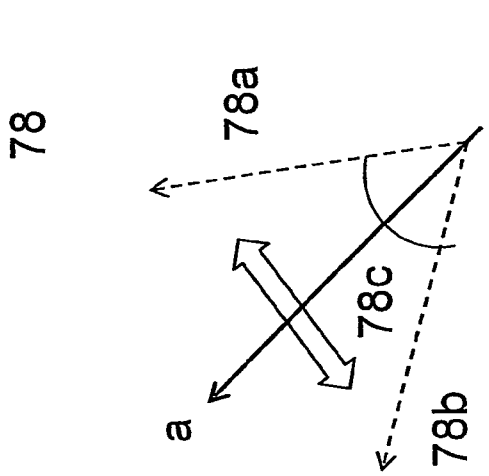


Fig. 5