

LANGUAGE TECHNOLOGIES IN EUROPE: THE MARKET AND TRENDS

STUDY SUMMARY 2005-2006

1. OBJECTIVES

The 2005-2006 “Language Technologies in Europe: the Market and Trends” study, carried out on behalf of the *Direction de la Stratégie* of the *Ministère délégué à la Recherche et aux Nouvelles Technologies*, in the scope of the Technolangue programme, determines and assesses the European market as regards the language tools that process text and voice data, and identifies the axes of development.

The study consists of the update of the study carried out in 2003 involving the 13 countries of the European Union (Germany, Austria, Belgium, Denmark, Spain, Finland, France, Greece, Ireland, Italy, the Netherlands, the United Kingdom, Sweden), and its extension to the 12 other member countries, 10 new members having joined in 2004 (Cyprus, Estonia, Hungary, Latvia, Lithuania, Luxembourg, Malta, Poland, Portugal, the Czech Republic, Slovakia, Slovenia).

2. METHODOLOGY

The same methodology used for the previous study has been applied:

- research and exploitation of the available economic and financial data ;
- extraction of the market’s key indicators ;
- complementary interviews with the principal actors of language tools ;
- presentation of a group of experts for eventual refinements and enrichment.

The study was carried out, for the most part, during the year, 2006, in 3 phases:

- Phase 1: identification of the actors and analysis of the offer in Europe and in France.
- Phase 2: analysis of the language tools core market in France.
- Phase 3: determination of the trends and perspectives for 2010.

3. RESULTS

It can be confirmed that the language technologies market is a “niche” market represented by a multitude of actors, with some being both software publishers and service providers. The segmentation based on the uses and the applications deployed has been used again, in this study, to confirm the trends observed in 2003 and to better understand the evolutions.

This typology can be broken down into nine segments:

1. Content management	4. E-business	7. Voice and multimodal interfaces
2. Document management	5. Language resources	8. E-learning
3. Translation	6. Production and editing	9. Intelligent search engines

3.1. IDENTIFICATION AND EVALUATION OF THE SUPPLY

The language tools market is closely correlated to the variations in the different structural and economic conditions of the TIC: the evolutions in NTIC expenditures and in Research and Development; the position of the software market; the increase in data storage and distribution networks; the deployment of the NTIC sector in the global economy.

According to the OECD, the population count in 2005 for the 25 countries of the European Union was 460 million people, with 160 million of them Internet users, and 55% of them speaking at least a second language.

In 2005, the NTIC expenditures in Europe (Administration, Major companies and SME, private individuals) were evaluated at 625 billion euros. Forecasts for 2006 are approximately 6% of the GDP.

The portion of expenditures in Research and Development in Europe (private and public investments): 200 billion euros in 2004, that is to say 1,90% of the of the GDP. Germany, France, and the United Kingdom accounted for 120 billion euros, that is to say, 60% of the total.

The language-oriented R&D was evaluated at 800 million dollars in 2003. The R&D expenditures of the producers in the ITC sector's in France were 0,16% (against 0,5% in the United States).

The 2005 market position of the software publishers and software companies in France represented 31,5 billion euros. The French software market accounted for 7,2 billion euros in 2005 with 2500 companies totalling 60 000 persons in France and 8000 throughout the rest of the world.

Assessment of the supply-side of the language technologies market:

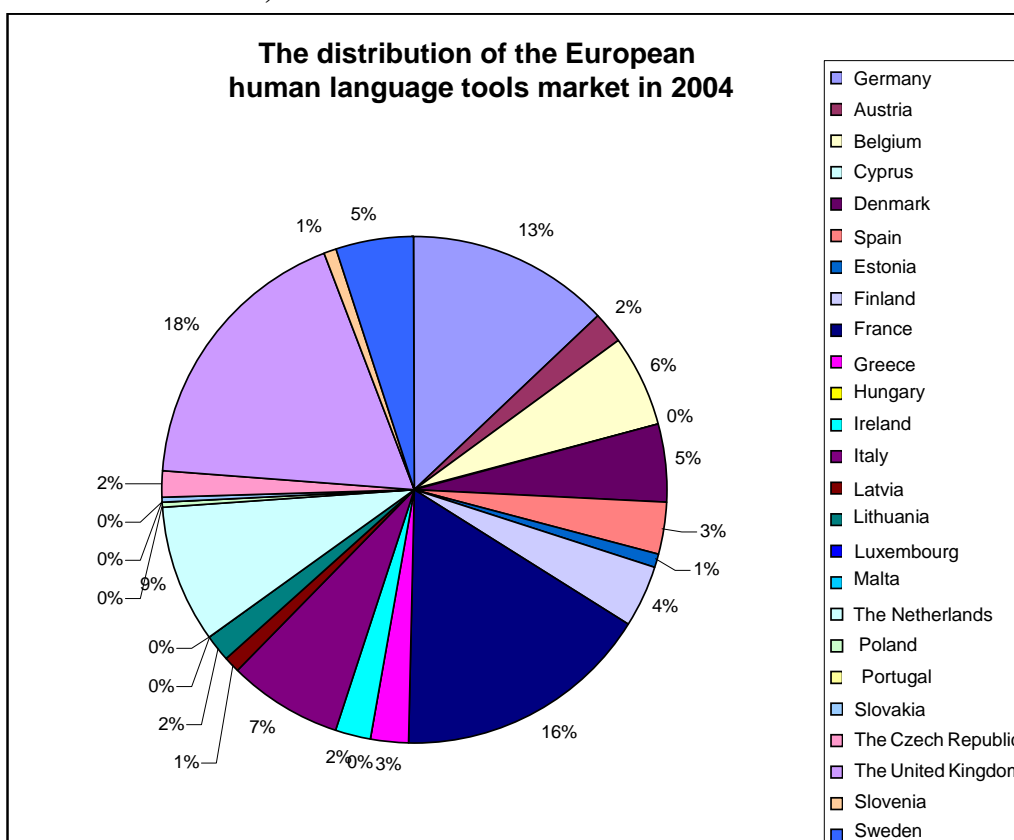
In Europe (25 countries), over the period 2000-2005, more than 50 European countries specialising in language processing disappeared (liquidations or buy-outs). There was a considerable increase in the number of companies created in the 1990s and in 2000.

In the 20-year span, the number of companies created has multiplied by more than tenfold, increasing from around 30 to more than 500 companies in Europe in 2006.

482 European suppliers are present in at least one of the 9 application segments:

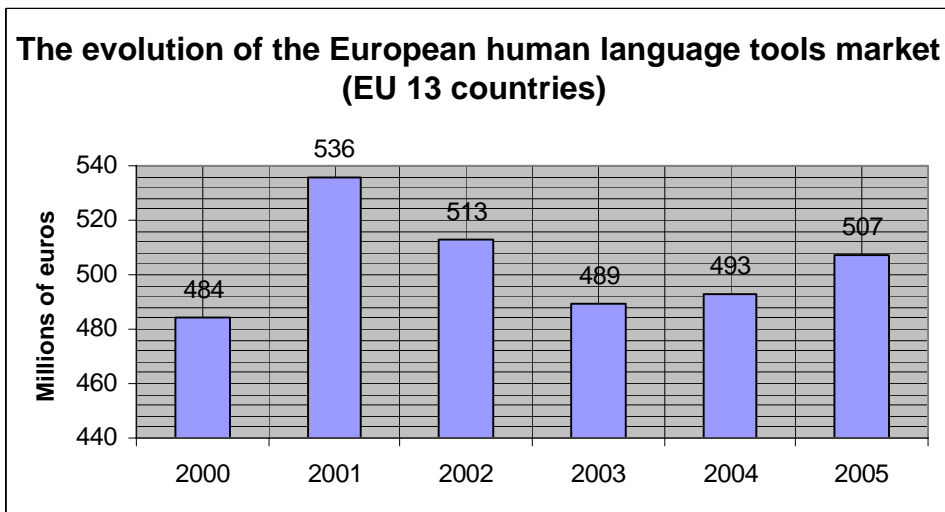
- the offer in Europe is distributed between text processing (70%) and voice processing (30%);
- the European market was evaluated at 610 million euros in 2004. The forecasts for 2005 are estimated at 328 million euros.

According to the turnover figures made available, the European language tools market in 2004 was estimated at 610 million euros, broken down as follows:

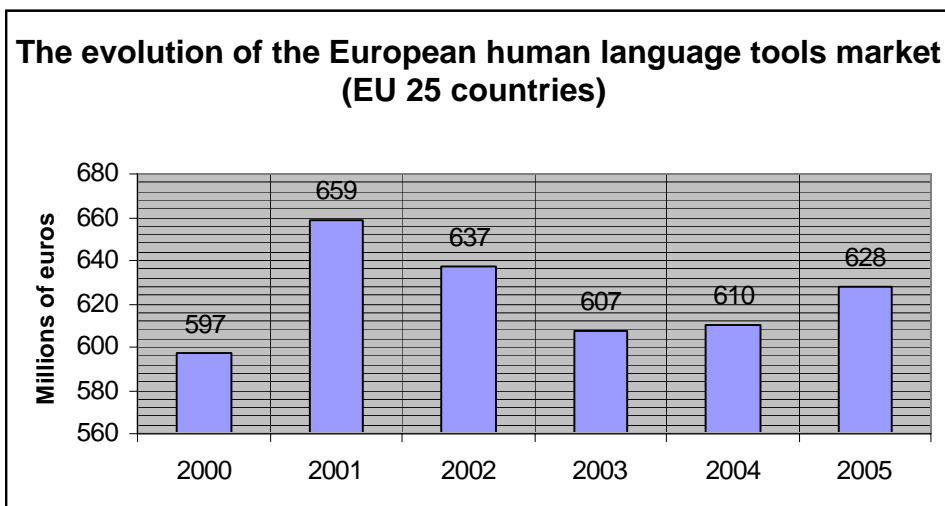


Thus, the leading countries on the language tools market are the United Kingdom, France and Germany, the three of them accounting for 47% of the European market. The United Kingdom is the leader on the language tools market. This position can be explained by the predominance of voice processing in this country.

In the 13-countries European zone, 394 European suppliers are present in at least one of the 9 application segments. According to our updated estimations, the European language tools market (the 13-countries Europe) confirms that there is again an increase which had begun in 2004 after the bursting of the Internet bubble.

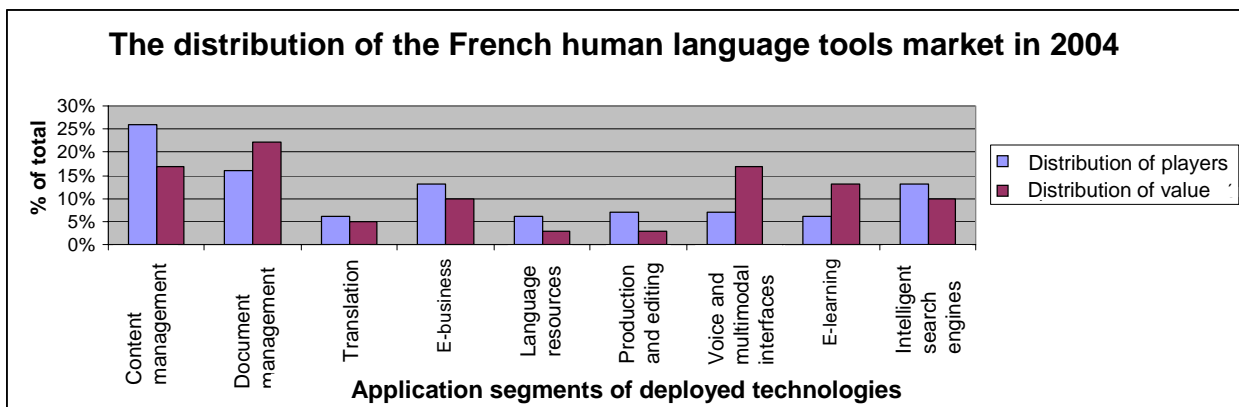


In the 25-countries European zone, 482 companies have been identified, and only 252 have communicated their turnover figures. The market's evolution (25-countries zone) is following the same trend as that observed in the 13-countries European Union zone.



In France, 109 companies have been able to be identified as falling within the perimeter of language engineering: the generated turnover represents approximating 78,8 million euros, that is to say, 16% of the European market and France is the second leading country after the United Kingdom.

Based on the leaders of the French market, we can determine the predominance of text processing on the French market: 77% of the leading companies on the French market are positioned on text which is, still today, the driving force behind the value of the language tools market in France, although the income generated by voice processing is undergoing a significant increase.



In 2005, approximately 36% (40% in 2003) of the actors identified were positioned on only one application segment.

The companies that are the most present on the language tools market are those that deal with document content (26%), document management (16%), intelligent search engines (13%), then specialists in e-businesses (13%), those specialising in voice and multimodal interfaces (7%), and those specialising in production and editing (7%).

The companies that are the most present on the French language tools market in terms of turnover potential are the EDM companies (22%), content management (17%), voice recognition and synthesis (17%), e-learning (13%), e-business (11%), and the intelligent search engines (10%), automatic translation (4%), production and edition (4%) and language resources (3%). In terms of turnover potential, EDM companies (27%) and document content (21%) represented nearly 50% of the market in 2002.

The French market remains concentrated. The first 10 companies of those chosen in the sampling of the French language tools market represent 70% of the turnover while the next 10 represent only 20% and the last 20 account for less than 5% of the turnover. The leading actor, mainly represented in segments 7 and 8, has a turnover far greater than those of the other countries chosen in the sampling. The majority of the sector's significant turnovers often come from actors deploying activities related to language engineering.

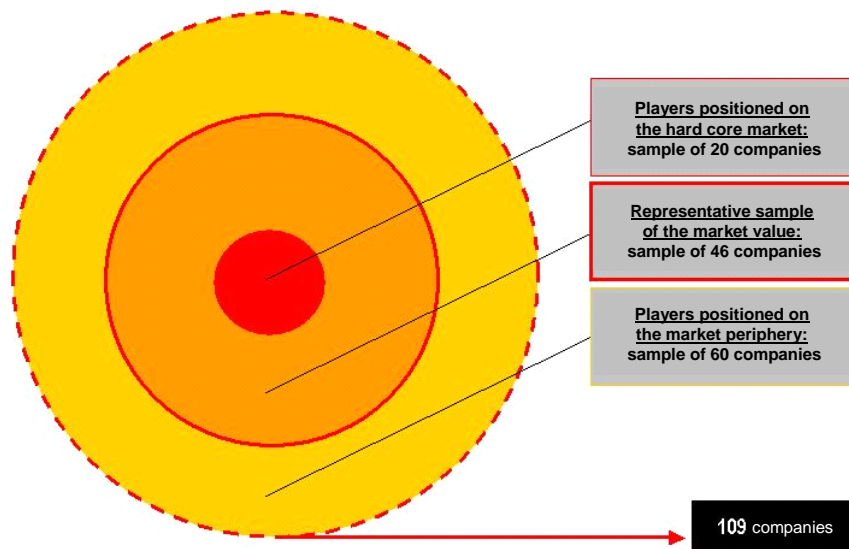
Note:

On the language tools market, the average turnover of the smallest companies is 300 000 euros. For the most important companies, principally positioned in the language technologies sector, the average turnover is 5 billion euros.

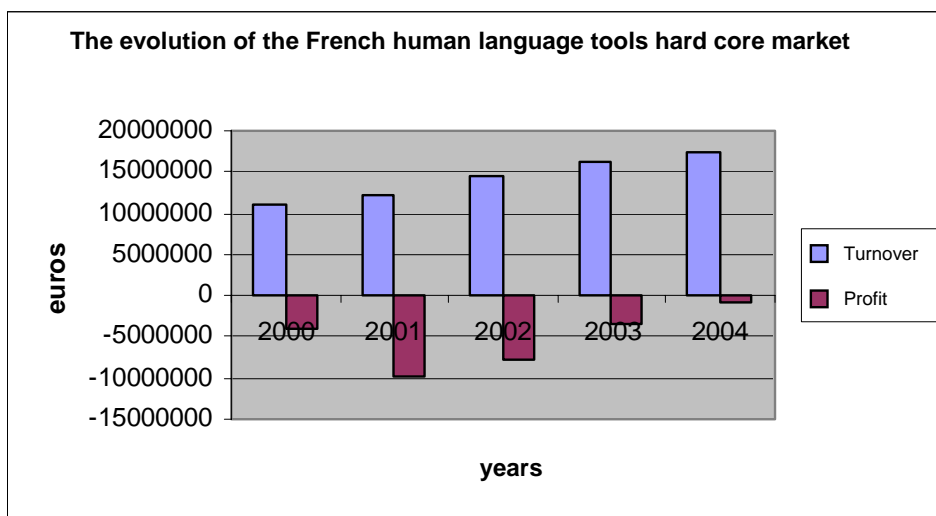
3.2.ANALYSIS OF FRANCE'S CORE MARKET

As regards the 109 actors identified in France, they can be situated at different levels of the language tools market: several of them are historically positioned on the core of the language engineering market, others are found at the periphery of the language technologies market.

The French human language tools market : hard core and periphery



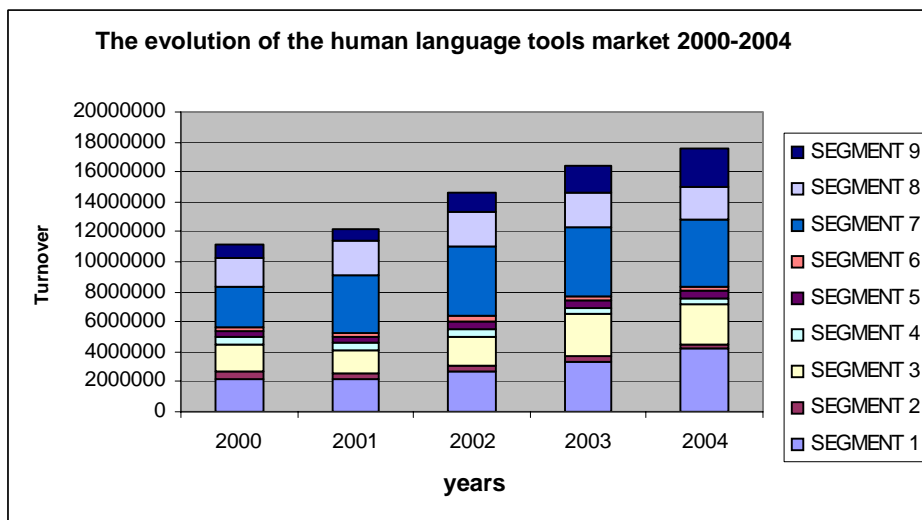
The analysis of the trends of the language technologies core market was performed using a representative sample of the 20 pioneer companies in language engineering in France and positioned at the hard core of the market. The earnings generated by these 20 actors positioned at the hard core of the market was 17,5 million euros in 2004, an increase of 6,93% when compared to 2003.



The earning capacity of the sampling remained in the red over this period. However, the turnovers generated when looking at the result improved considerably until there was a return to the black, which was expected in 2005 after 5 consecutive years of a drop.

The evolution of the generated earnings for each segment shows that 4 segments dominated and represented 90% of turnover for the period 2000-2004: segment 1 “Content management”; segment 3 “Automatic translation”; segment 7 “Voice and multimodal interfaces; segment 8 “E-learning” and segment 9 “Intelligent search engines”.

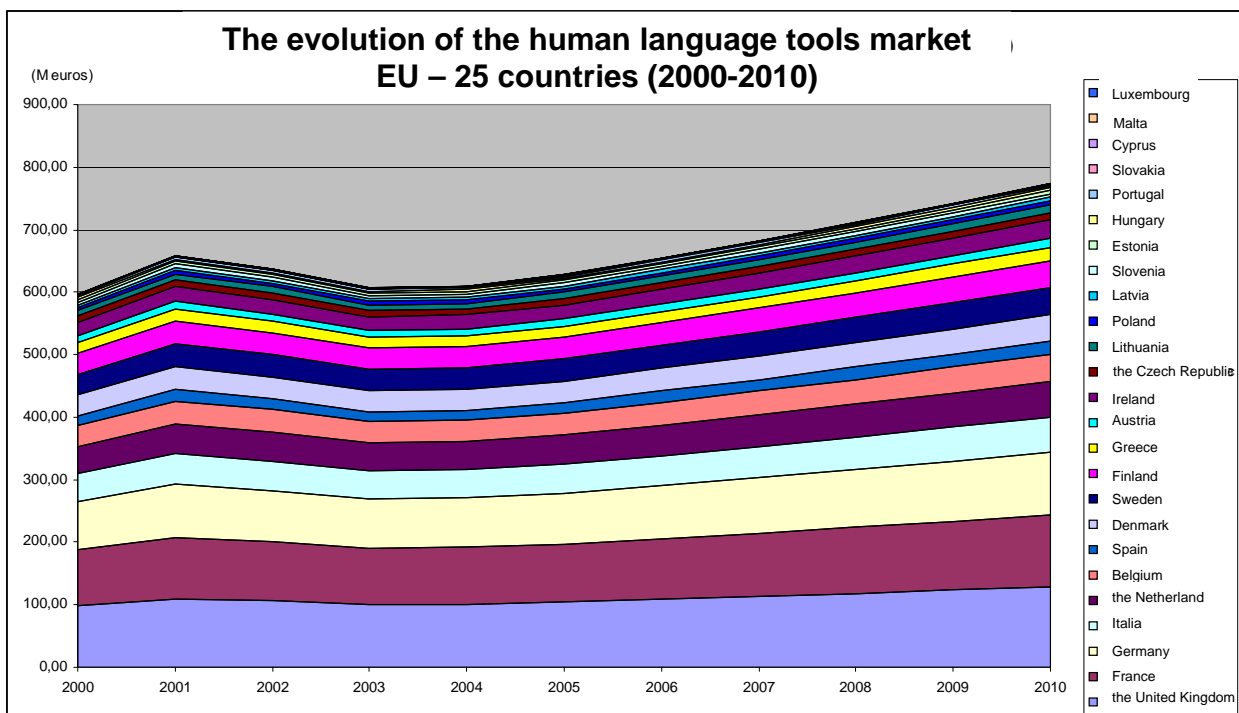
The average annual growth rates for the period 2001-2004 underline the positive trend marked by the 2-digit growth rates for these 4 application segments. The outstanding segment is Segment 9 as its average annual growth rate was three times greater than the average of the sector (31,42% against 12,08%) and its accumulated growth rate was + 179.31% over the period. The segments “Content management” (+94.24% accumulated over the period) and “Voice and multimodal” (+69.55% accumulated over the period) also experienced a strong growth rate. These 3 segments increased their share of the market over the period.



3.3.EVOLUTION OF THE EUROPEAN LANGUAGE TECHNOLOGIES MARKET

The language technologies market in the 25-countries European Union zone evolved from 610 million euros in 2004 to 628 million euros in 2005 and should attain 774 million euros in 2010.

The curve in the trend in the graph below once again proves that the market fell after 2001 and that it progressively has made an upward return to its earlier level between 2005 and 2007. The years 2008-2010 appear to be good years as regards turnover growth rates and economic earning capacity in that the generated earnings of the sector's companies are again finding equilibrium, with certain actors even having made profits as of the year 2005.



The analysis of the evolution of the 9 segments over the period of 2000-2010 in France was carried out based on the study of the representative sampling of the market. The generated earnings by the market seems to be situated between 100 million (low estimate) and approximately 115 million euros en 2010 (high estimate).

The language technologies core market in France is estimated to attain 45,8 million euros in 2010 (high estimate).

The growth rates noticed in the language tools core market are higher than those generated by the market's average (12,08% from 2000 to 2004 for the hard core market against 4,70 for the overall average).

Likewise, the core market's economic earning capacity is often less than that generated by the market: the turnovers generated by the actors of the hard core market are often less than the average turnovers generated by the actors structurally more important and positioned on the market's periphery which can explain this growth differential.

The maturity of certain application segments and the stability of the actors of the hard core market illustrates the new digital era, the truly growth leverage for the sector.

The perspectives as regards the evolution of the language tools market depends, above all, on its intensity of competitiveness which is getting stronger and stronger: the evolution of the supply, the demand and the technological environment is engendering market concentrations - a systematic confrontation of the competitiveness of the technologies and the tools available.

For these past 10 years, the mergers and acquisitions within the market have become more and more frequent between competitive actors and can be explained by:

- the market's frenzy for the Internet in the 1990s with a race to be the first entrant, which enabled partnerships to be developed and mergers between competitive actors to be validated ;
- the economic stagnation of the NTIC sector at the beginning of the years 2000 due to the bursting of the Internet bubble and the loss of market value of several supply actors ;
- the strategic multinational positioning of the technology suppliers on a language tools market that is becoming global (the joining of European and American actors), today materialising by the penetration of American actors on the European market ;
- the will to control the technologies on the demand side, inclined to buy out companies that have developed knowledge management tools in a highly strategic field.

The market's dynamics are now turning towards demand-side logic: the integration of tools into existing architectures, a return on investment, a continual demand for technological evolutions with the use of multimedia and the mobile Internet becoming widespread and the need for multilingual communication.

The offer is still scattered and fragmented: the technology providers companies' sizes still vary considerably and their national representativeness is not always easy to identify. Moreover, from one year to another, the perimeter fluctuates and sometimes even tremendously: some companies disappear (liquidations, buy-outs or mergers), others are created (spin offs of large account companies, research centres, start-ups).

The years 2006-2010 could represent the fitting growth rate expected by the sector's different protagonists.

CONCLUSION

The convergence of technologies is quite noticeable in the communications field. Computerisation, the telecommunications and the audiovisual fields have been federated by digitalisation. Digital content now involves text, sound, and images. The multimedia era is taking shape with the deployment of interactive kiosks, next generation mobile telephones, entertainment consoles and terminals and embedded systems. High-speed network connections are becoming more and more widespread. Applications involve multimedia documentary intranets, cross-language content management, *e-business*, voice/multimodal, *e-learning*, automatic translation...

Different factors have participated, these past years, in the development of the European language technologies market:

- the different digitalisation projects in the public and private sectors ;
- high-speed, wireless computerised network ;
- the accrued globalisation of the language technologies market with an increase in international partnerships (strategic, marketing, technological) and the increase created by the need for instantaneous translations, whether they be voice or text, linked to the social-cultural factors – multilingualism being one of them ;
- open source, at the same time being:
 - a development opportunity: more and more stable, evolving and highly efficient, open source solutions are still free and enable to capitalise on the recognized models and standards.
 - and a competitive limitation that forces cost decreases : the associated developments are sometimes long and costly, which can put a brake on the development of competitive, European open source tools insofar as where the emerging countries that master these computerised languages, that capitalise a technical know-how and that have learned to acquire recognized competences, can become serious contenders for multilingual projects and in the development of knowledge management technology at the least cost.

Several factors incite decision-makers to integrate innovative solutions into their company to intelligently manage digital content:

- the digital convergence of computerisation and information and communication technologies ;
- the integration of multimedia content ;
- high-speed, wireless Internet ;
- web 2.0 and web 3.0 applications that promote editorial and social contributions of users-subscribers of contents and services.

The progressive use of NTIC leaves us to predict that the language tools market will open towards the general public. The need is felt to take marketing actions to optimize the appropriateness of the supply and demand.

Bureau van Dijk Information Management - Martine DEJEAN : mde@bvdim.com - Antoine RAULIN : ara@bvdim.com