National Technology Alliance

Human Language Technology Maturity Forecast
Final Report
24 March 2009

Rosettex™

Prepared for the National Technology Alliance by the Rosettex Technology & Ventures Group and its Partner TechCast LLC

TechCast
A Virtual Think Tank
Tracking the Technology Revolution

Publication Number: FR-001-091-032509-530

Report to the

Office of the Director of National Intelligence (ODNI)
Directorate of Science and Technology

Human Language Technology (HLT) Maturity Forecast

Submitted by

William E. Halal, PhD
Institute for Knowledge & Innovation
George Washington University
President, TechCast LLC

24 March 2009

SUMMARY

A broad assessment of highlights from this study shows a few major trends running through the data:

1. ***Now Commercialized:*** Speech Recognition (SR) and Image Recognition (IR) are commercially available now and are likely to reach 30% adoption levels soon (2016 and 2013 respectively). These markets are likely to be in the $5 billion range by 2010.

2. ***To Become Commercial Soon:*** Language Translation (LT) and Emotion Identification (EI) are expected to become commercial in 2013 and 2014 respectively, with 30% adoption in about one decade. The market for LT is expected to be large ($10.5 billion) while EI should be modest ($3.1 billion).

3. ***To Become Commercial Later:*** Non-Standard Languages (NSL) appear to be far less well-developed, with commercialization likely to require almost a decade. A small market size of about $1.7 billion is expected by 2010.

4. ***Complex Languages:*** For all technologies, commercialization of complex, non-Western languages, such as Arabic, Farsi, and Urdu, is likely to require roughly another decade.

5. ***Conclusions and Recommendations:*** The study concludes that these technologies are likely to be propelled along by a more general technological wave about 2015, so this is a crucial time to develop such capabilities. Possibilities are outlined for expanding on the system created for this study.

STUDY GOALS

TechCast was asked to study the following Human Language Technologies:

1. Speech Recognition (SR): SR is the computerized conversion of speech to create verbatim renditions for reading and analysis, to spot words or phrases, or to create computer-generated language responses.

2. Language Translation (LT): Also called Machine Translation, LT involves the computerized translation of one language into another for analysis, indexing, archiving, response, etc.

3. Image Recognition (IR): IR includes Optical Character Recognition (OCR), Optical Word Recognition (OWR), Optical Handwriting Recognition (OHR), and Interactive Character Recognition (ICR).

4. Non-Standard Languages (NSL): NSL is the machine interpretation of any languages not usually codified in existing dictionaries, for instance, combined languages (Spanglish, etc.), text messaging (BTW), and other novel forms of communication.

5. Emotion Identification (EI): EI is the machine identification of text, speech, or facial cues which reflect emotional states, e.g. anger, impatience, disgust, etc., usually according to a theory of emotion.

For each of these five technologies, TechCast was asked to scan for available background information, interview ODNI experts, create surveys to estimate maturity, review and correct the surveys, and have the TechCast Expert Panel complete the surveys. This report is a result of that work.

METHOD OF STUDY

The study was conducted using the research method developed by TechCast. TechCast experts scanned the Internet, periodicals, journals, and conducted interviews to summarize key trends driving each technology and obstacles opposing it. This information was clustered into key trends supported by numerous examples in bullet format. Data on maturity forecasts and market forecasts were also summarized in bullet form. This background information is included in the Appendix, which also presents the surveys conducted for the five technologies.

As can be seen, the surveys were designed to estimate when each technology would become commercialized, when it would reach the 30% adoption level, and the projected market size in 2010. The 30% adoption level is considered of particular significance because emerging technologies usually take off to enter mainstream use at that point. This is the stage in the technology life cycle when market demand reaches 30% of its saturation level,

usually expressed in terms of the potential number of users, products sold, or other measures that are relevant. Later, TechCast was asked to also estimate when these technologies would mature in complex, less-developed languages, such as Arabic, Farsi, and Urdu, and to obtain general comments. The surveys were conducted via the TechCast website and all 100 members of the TechCast expert panel were invited to participate by completing the surveys and offering comments. These experts are not authorities in Human Language Technologies, but they are knowledgeable about technology in general and they are experienced in using the TechCast method. For more on the experts who participated, see www.TechCast.org/About/Experts.

As noted in the Summary of Results below, the total number of responses (N) ranged from 27 to 31 for various questions. In addition to the survey data, two sets of comments were also obtained from the experts. The first set, which is most extensive, was made in response to a specific question asking for "general comments about the prospects for each HLT", and the second set was also made by the experts in response to a general request for comments. Both sets of comments are functionally equivalent.

It should be noted that the intent was not to produce a complete review of each field, but to summarize key trends in order to assist the experts in making sound judgments.

For more on this method, see www.TechCast.org/About/Method

SUMMARY OF RESULTS

Detailed results are contained in the Appendix and summarized in Table 1 below.

Generally speaking, the data appear to be sound and useful. Sample sizes (N) are adequate and the caliber of the respondents is high. The results discriminate well between data sets, with some varying significantly, especially when considering the modest standard deviations (SD) of 3-5 years. The forecasts are plausible and agree well with the supporting background scanning data.

Table 1
Summary of Maturity Forecasts and Market Forecasts
Means or Averages
Number of Respondents (N) = between 27-31,
Standard Deviation (SD) = between 3-5 years

Technology	Commercial Introduction	Take-Off Date (30% adoption)	Commercialization Complex Languages	Market Forecast in 2010
Speech Recognition	2009	2016	2018	$5.1 B
Language Translation	2013	2018	2019	$10.5 B
Image Recognition	1997	2013	2022	$4.9 B
Non-Standard Languages	2017	2023	2023	$1.7 B
Emotion Identification	2014	2019	2021	$3.1 b

ANALYSIS

Detailed analyses of these five technologies follow.

1. Speech Recognition

Survey results indicate speech to computer text technology is now commercially available and should reach the 30% adoption level at about 2016. Nuance, Google, Sybase, IBM, Microsoft, AT&T, and other companies are all improving products that are now considered highly accurate, at about the 99% level. The applications are vast, ranging from hands free interaction with all manner of machines to improved security. Automated voice synthesis is also likely to improve along with SR to provide complete two-way machine interaction. SR is especially advantageous for improving the interface with small mobile devices, and hardware-based systems are being developed that will improve performance at lower cost. The market is expected to be on the order of $5.1 billion in 2010. Below are typical comments from the TechCast experts:

“For European languages, we've already reached an acceptable rate of 99%.”

“I use Dragon Naturally Speaking and I find it effective.”

As with most language technologies, limitations persist. Noisy environments introduce errors, large and costly programs involving large amounts of data are needed to develop systems, and applications are daunting for complex languages like Chinese where subtle inflections and tones have great meaning. The survey data suggest non-western languages like Arabic, Farsi, and Urdu will not be recognizable for almost a decade. We are likely to see mounting use of SR in search engines, mobile phones, dictation, and conversational machine interfaces over the next few years. But frustrations are likely to be common as machines interpret human speech badly at times, continually pushing demands for improvement. Typical comments on the remaining challenges:

> "Chinese will be a problem for many years to come."
>
> "Natural human-computer interaction will be [common] but huge improvements are needed"

2. Language Translation

As the survey indicates, LT is useful now for routine, informal tasks but not for formal uses requiring precision. However, the rapid spread of globalization, e-commerce, and cultural exchanges is creating growing demands for translations among the hundreds of languages actively in use. Corporations and governments are adopting text-to-text translation systems for their growing need to translate documents, and large IT companies like Language Weaver, Systran Software, Google, and NEC are rapidly introducing LT packages that are reputed to number in the hundreds. Market demand is expected to be in the $10.5 billion range by 2010. Survey results indicate good quality commercialization is likely in 4 years (2013) and 30% adoption is likely to be reached in another 5 years (2019). The availability of good LT is expected to be a watershed event, fostering cross-cultural interaction as never before, for better and for worse. One respondent summed it up:

> "In a global marketplace the demand for this technology is vast."

The challenges of accurate interpretation are enormous because all written languages have idiosyncrasies and change constantly, and because meaning depends on style, intent, context, and other subtle nuances that are hard to codify for machine translation. This is especially true for the non-western languages that are less well-developed and have not been studied for translation, which should require another decade to mature, if resources are available to support costly projects. LT will likely become far more widely used over the next few years, therefore, but problems interpreting subtleties will persist, requiring human supervision for applications requiring great accuracy. Respondents noted the challenge:

> "Automation with review by human translators is here – [but] real-time machine translation is a huge goal."
>
> "[There is an] inherent structural difficulty in some languages, full of symbolic dimensions and multiple meanings for a single word or sentence. …[This is] a very

difficult task … without putting the words in a social context. Moreover, Language … changes…"

"[A] core breakthrough will be pattern recognition with convergence of audio and visual."

3. Image Recognition

IR includes Optical Character Recognition (OCR), Optical Word Recognition (OWR), and Optical Handwriting Recognition (OHR). Of these, OCR has been well-developed for many years, and accuracy is now 99 percent for clean text. Our experts estimate commercialization began in 1997 and should reach the 30 percent level by 2013, with the market expected at $4.9 billion by 2010. Word recognition is not as well-developed, although we do not have a separate forecast for that technology. Handwriting recognition is far less advanced for obvious reasons, with our experts estimating commercialization at 2018. It should be noted that handwriting is increasingly uncommon, so this limitation may diminish in importance with time. Here are respondent's comments:

"Good fast-moving bet in the West; much slower elsewhere."

"Less handwritten content is being created, so the market won't grow much, but methods are perfecting."

A great deal of research activity is underway in IR because of growing demand. Kofax, EMC, Google, Nuance, SRI International, Internet Archive, and others are working on applications for recognizing text in candid video and photos, handwritten postal addresses, street signs, and name tags. Universities are working on recognition of non-Western languages, such as Chinese and Arabic, often focusing on OWR, but these efforts are not likely to bear fruit until 2022, like most HLT's for complex languages.

"A real operative OCR is a very difficult task …because of document heterogeneity."

"The potential is huge and will have a steep curve to over $20B by 2018."

4. Non-Standard Languages

This field serves as a catch-all for languages that are not codified: informal media (such as online chat), text messages, email, non-native speakers of a language, and dialects (such as Spanglish). There is a growing need for NSL because online communications are exploding and often use "netspeak" – BTW, IMHO, etc. One authority estimates 70% of all digital text is NSL. To help companies and governments deal with this, IBM's Indian Research Lab, Inxight, Nuance, Microsoft, SPSS, Anderson Analytics, ScanSoft, and other firms are developing a variety of programs that attempt to interpret NSL in text, phone messages, and other media.

"We are in a time where …language changes every day, every minute. So, the "formal languages" (Spanish, English, etc..) are disappearing as a solid reference."

"High accuracy will not be cost effective for a decade."

The challenges are huge, of course, because NSL is highly variable by nature, all languages are constantly changing, context is crucial, and large amounts of data are needed for analysis in costly programs. Some forecasters believe these barriers will be overcome in a decade or so, which agrees with our forecast that expects NSL to become commercially available in 2017 and reach the 30% take-off point by 2023. Demand, however, is expected to remain at $1.7 billion by 2010.

"Similar to LT for non-western languages. Needs convergence of visual and audio pattern recognition."

"Context is critical."

"Good translation may never be possible."

5. Emotion Identification

The accurate identification of emotions seems likely to become far more important for security reasons, to better understand human behavior, and as robots/avatars increasingly interact with humans. Applications often combine interpreting cues in speech, facial expressions, text, gaze, or other signs of emotion to produce more accurate results through triangulation. MIT, Phillips Research, Deutche Telekom, Prosody, Noldus IT, Affective Media, Autonomy, Emotion Mining Company, Emotive Systems, iMotions, Idiometrics, Clarabridge, RightNow, and universities are among the many players who are active in this emerging market. The expert survey found that EI will become commercialized about 2014 and enter the mainstream at 30% adoption levels about 2019 with a market demand in the $3.1 billion range. Comments raised interesting points:

"Key element for LT and NSL. Commercialization of EI … enables both."

"EI is a sure thing - BUT, technophobes will curb advances."

Researchers caution, however, that interpreting emotions is fraught with many possible sources of error, so EI is not likely to be used for critical purpose without human backup. And the interpretation of emotions in people using complex, non-Western languages is even more daunting, so that is not expected until 2021. Our experts cautioned about the crucial role of understanding cultural context:

"EI will be more difficult, smaller market and fraught with false positives/negatives."

> "… cultural patterns determine the way people express their feelings and interact with others, and this is … the central point to design an effective tool for analyzing emotional responses."

CONCLUSIONS AND RECOMMENDATIONS

The technological and economic forces driving these language technologies are likely to accelerate over the next few years. Today's constant drumbeat of cascading business failures is daunting, but our normal forecasting work suggests that continual advances in IT, the green revolution, advanced auto designs, surging e-commerce, and other new business sectors are poised to lead the global economy out of today's recession, producing a new economic boom at about 2015.

Technological progress is insulated from economic downturns because R&D projects often have long-term support from governments, foundations, and universities. Entrepreneurs may delay product launches during recessions, but they also discount temporary downturns in favor of long-term prospects for pent-up demand. The relentless power of technology can be seen in the fact that the dot-com bust of 2000 didn't faze the Internet, which soon entered a more sophisticated "participative" stage of Web 2.0 sites, like Facebook and YouTube, and even the election of Barack Obama. Today's recession appears to be a mere 2-3 year economic dip as globalization continues to be driven by the accelerating forces of technology. As Andy Grove, Chairman of Intel, put it so well - "Technology always wins in the end."

Computer power is almost certain to continue increasing under Moore's Law, costs will continue to fall dramatically, IT infrastructure will improve through faster broadband and wireless communications, and artificial intelligence will automate most routine aspects of human thought about 2020. These striking gains will be made possible by second generation computer designs, mainly optical, biological, and quantum computers, which should arrive during the end of the next decade.

Today's surging interest in green business should take off in 4-5 years, and governments are likely to take serious steps to curb global warming about the same time. Alternative energy sources – wind turbines, biofuels, nuclear plants, and solar cells – are growing 30-50% per year, roughly the same rate of Moore's Law. Green technology is roughly a $500 billion market and expected to reach $10 trillion in 2020, larger than autos, health care, and defense.

Almost all sectors of the economy are likely to be rejuvenated with high-tech advances in roughly the same time frame. A new wave of green autos powered by hybrid, electric, and fuel cell engines should enter the mainstream about 2013 - 2018, and we are likely to see "intelligent cars" that may even drive themselves.

The information technologies driving globalization are gaining momentum as publishing, entertainment, virtual education, and other forms of e-commerce reach the critical 30% adoption level where new businesses usually take off. And the huge populations of

China, India, Brazil, and other developing countries are moving in droves to PCs, the Internet, smart phones, and global media, for better or for worse. Our forecasts show that three-four billion people will soon inhabit a digital world that is smarter, faster, and interactive, creating online markets of several trillion dollars.

The year 2015 seems to mark the serious beginning of all this innovation because it is the next inflection point in the 35 year cycles that roughly govern U.S. markets. The roaring twenties were the peak of a 35 year cycle that ended with the Great Crash and Depression. The Eisenhower boom of the sixties started about 1945 and was followed by the Reagan boom that began with his election in 1980. Today's collapse marks the end of the Reagan 35 year cycle, and it is likely be followed by the "global boom" outlined above starting about 2015.

This scenario frames the forecasts of this study in a more meaningful context. All of these technologies (except NSL) are forecast to take-off at roughly the same period between 2013 to 2019, so they are likely to be propelled along to maturity by the coming economic boom. The combination of advanced IT and the imminent explosion of e-commerce in particular is likely to increase the level of communication around the globe dramatically, and with it the security threats of hackers, viruses, cyberwarfare, and terrorism. Exploding e-commerce will also make huge amounts of data available for the analyses typically needed to develop accurate HLT programs. And advances noted above in IT will provide the needed computer power at lower cost, while advancing Artificial Intelligence.

We suggest building on the system designed for this study to create a capability to track critical technologies of interest to the ODNI. A model has now been created, populated with scanning data, and surveys conducted, so it would be relatively easy to continue this work by updating the scanning and periodically conducting new surveys. The system could be improved upon by redesigning it for ODNI use, expanding the scope to cover other technologies or other critical issues of interest, inviting ODNI professionals to participate, and possibly to involve other intelligence agencies and outside groups like TechCast as well.

APPENDIX

Speech Recognition (SR)

By Dexter Snyder, with contributions by Florence Reeder

SUMMARY

Speech is a natural, rapid, and common act, and so SR offers a more convenient hands-free and eyes-free form of communication and control. Commercialization of SR is being driven by increased computer power, more sophisticated software and hardware-based systems, the need for cost saving services, and increased demand for security and intelligence. As the technology expands to a wider range of applications, SR is becoming a major enabler for simplified interfaces with computers and phones in cars, health care, and other fields. Companies, like Google, with the ability to create large voice databases, are working toward a competitive advantage through providing superior systems for individual users. Bill Gates recently said, "Within five years, people will discard their keyboards and interact with computers using … voice controls." (CarnegieMellonToday.com 2/21/2008)

DRIVING TRENDS

Advances in Technology

1) Nuance is the leading SR software provider, with 2007 sales of $700 million. Their product, Dragon NaturallySpeaking, returns 99% accuracy and about 160 words per minute speech-to-text in a controlled environment with no training. Nuance collected a large database of speech samples from people with different languages and accents and used it to improve the technology. (PC World, 7/18/06)

2) Google uses large voice databases to train its SR systems and offer more accurate responses. As one example, its free Voice Internet Search Service stores the voice samples of millions of users. (google.com/goog411)

3) Sybase and Toyota have been awarded patents on open-conversation SR-based interfaces. Ben Reaves, a Toyota executive, said: "Free-form speech-based interfaces will be an essential feature in future car interfaces with hands-free eyes-free interaction requirements" (Sybase Press Research, Nov. 12, 2007)

4) SR in healthcare is a disruptive innovation for those companies able to integrate it into the business process. Nuance leads this field. (Gartner 2008)

5) IBM and Microsoft's English SR programs are expected to raise accuracy from 95% to 99% by 2010 and match the human voice soon after. An IBM scientist said "We have … crossed the threshold where users will accept it." (Scientific American, 5/16/05)

6) ABN Amro, a large Dutch bank, has 95% of their customers conduct transactions by phone using SR because people find it faster and easier. (TechnologyReview 7/20/06)

7) Electronic auto systems, like GM's OnStar, respond to voice commands. The Lexus system knows over 100 commands and guides the driver with voice and visual directions. (New Scientist, 12/7/03; Wired, 9/2/04)

8) DARPA funds the Global Autonomous Language Exploitation Program at $50 million/year to have IBM, SRI International, and BBN Technologies develop automated transcription and translation of foreign speech and text. (darpa.mil/ipto/programs/gale/gale.asp)

Mobile Applications

1) Google introduced voice search on its G1.Phone mobile application, which also allows the company to gather data to improve the system. (TechnologyReview Nov 20, 2008)
2) AT&T offers speech recognition on mobile phones with a program called Watson.
3) Nokia and Sony are creating intelligent agents operating on 3G cellphones that make purchases, among other tasks. (New Scientist, 6/15/03)
4) Vlingo Company is marketing an SR interface for mobile phones. "Small platforms need speech recognition," said an MIT researcher. (TechnologyReview Sept/Oct '07)

Hardware-Based Systems

Today's SR is based on software. For speaker-independent, large-vocabulary, continuous speech, the demand on computing power is immense. Real-time recognition and 90% accuracy require 100 MB of memory and 1000 MHz of CPU. Hardware-based solutions are coming:
1) The Carnegie Mellon "In Silico Vax" Team is building a working field programmable gate array that handles a 1000-word vocabulary in near-real time. This is the most complex speech recognizer rendered completely in hardware. Commercial versions can enable sophisticated SR applications for any portable device. (ece.cmu.edu/~rutenbar/)
2) The NSF is sponsoring research to move SR from software onto chip designs, making it 1000 times more effective. (Kurzweilai.net, 9/14/04)
3) SR will improve with advances in pattern recognition embedded in hardware. As examples, silicon neural network chips are based on a model for brain cortex function, and low-power analog SR circuits are being based on a model for human cochlear function. When commercial, these processors will enable SR capabilities rivaling those of humans. (Dan Hammerstrom, Portland State)

Applications for Complex Languages

1) The DARPA GALE Program expects 95% accuracy for translating Arabic and Chinese speech. In the first phase, accuracy for both increased from 35% to 65%. John Olive, DARPA/ITPO, projects 95% accuracy by 2012. Performance of GALE for Arabic and Chinese is being evaluated at NIST, with results available March 2009. (mt-archive.info/AMTA-2006-Olive.pdf).
2) A growing family of large speech databases is becoming available for training. They include GlobalPhone (Interactive Systems Laboratory, Carnegie Mellon U.); Carnegie Mellon Sphinx (cmusphinx.sourceforge.net/html/cmusphinx.php); Entropic Systems/Cambridge University HTK (http://htk.eng.cam.ac.uk), and IBM (sciencedaily.com/releases/2001/10/011011065502.htm)
3) A major challenge involves speaker variability and nonstandard dialects. Spoken Arabic dialects vary dramatically from the standard form, creating a current need for domain and context information to capture meaning. (MINDS 2006-2007 Report of the Speech Understanding Working Group, ODNI supported)
4) Mandarin Chinese is a tonal language, unlike English. The pitch of a syllable determines its meaning, in conjunction with the sequences of consonants and vowels. Progress is being reported to build models capable of accurate SR.

OBSTACLES

Noisy Environments

A major hurdle that blocks broader application is that noisy environments prevent accurate SR. (MINDS 2006-2007, Speech Understanding Working Group, NIST/ODNI)

1) Audience Corporation recently announced a technology to pick voices out of background noise using a Fast Fourier Transform (FFT) that takes cues from pitch, spatial location, and onset time of the speech. (www.audience.com).
2) The Honda robot (Asimo) has an SR system that understands three humans talking at once with 70-80% accuracy. It locates each voice using an 8-microphone array on Asimo, then cleans up the signals before decoding. (New Scientist, June 10, 2008)
3) Through its Advanced Speech Encoding Program, DARPA is developing a system to deliver robust SR in noisy combat situations by sensing muscle signals from the vocal cords when a person mouths speech. If successful, this approach could be used in acoustically similar environments, such as manufacturing. (darpa.mil/ipto/programs/ase)

Complex Languages

Nuance, a major supplier of SR software, notes that costly training data is needed to build an application for a new language, and no single company can afford the investments. The company considers Arabic, Chinese (Mandarin and Cantonese) and Spanish to be Priority 1 languages meriting investment.

Market Forecasts

1) SR in automated directory assistance and speech-enabled internet is growing 22% per year and is expected to reach $2.7 billion by 2009.
2) SR sales for mobile applications such as cell phones and car navigation should reach $250 million by 2011. (Economist, June 2007).
3) SR for healthcare transcription is globally estimated at $170 million in 2008, and predicted to double by 2013 (Datamonitor Report, 2008).
4) Biometric applications, estimated at $100 million in 2008, are projected to grow to $600 million by 2012 (Opus Research, May 2008, Voice Biometrics Conference 2008).
5) Combining the above segments, the global market is likely to be about $2.5 billion in 2008, reaching $3.5 billion in 2009 and $5 billion about 2010 or soon after.

MATURITY FORECASTS

1) Ray Kurzweil projects that computers as we know them will start disappearing as early as 2010, to be replaced by small electronic CPUs embedded in all manner of devices. (SpeechTek 2008)
2) DARPA's Global Autonomous Language Exploitation Program is planned to develop a capability for handling Arabic and Mandarin Chinese by 2009.
3) Intel's Technology Group forecasts "strong capabilities" in transforming speech-to-text by 2015. (NewScientistTech, 2/22/05; 2005 Intel Developer Forum)

Survey Results

1	When will effective Speech Recognition (SR) in general become commercially available?* Enter the year from 2000 to 2030 when you think commercialization is likely to occur. If you think it has already occurred, enter the year when you think it happened.	**RESULTS** N = Number of responses SD = Standard deviation Mean – 2009 N - 28 SD – 4 Mean – 2009 N - 28 SD - 5
2	When will commercialization occur in non-western languages that are more complex and/or less well-developed, such as Arabic, Farsi, and Urdu?* Enter the year between 2000 and 2030 when you think this is most likely to occur.	**RESULTS** Mean - 2018 N - 28 SD – 4
3	When will effective SR reach the take-off point of 30% adoption to enter mainstream Use?* Enter the year from 2000 to 2030 when you think 30% adoption is likely to occur.	**RESULTS** Mean - 2016 N - 28 SD – 3
4	What will be the market for SR in 2010? Our data show large applications like LT (both human and machine) at about $14 billion today, medium markets like SR and EI about $2 billion, and small markets for IR and NSL less than 1 billion.* Please estimate the market size for SR at 2010 Under $1 billion	**RESULTS** **% N**

		7	2
	Between $1-5 billion	57	16
	Between $6-10 billion	32	9
	Between $11-20 billion	0	0
	Over $20 billion	4	1
5	Please offer any general comments about the prospects for SR.	**RESULTS** Moore's Law will make this happen sooner than later This is already happening! Commercial for Western domain-based use. Must combine with visual for everyday and non-western apps It will contribute to the burgeoning human renaissance. Chinese will be a problem for many years to come. Many PC users expect this soon - will give major productivity gain Thanks to DragonFly this tech is well underway - Almost ready for Prime Time. The benefits of SR will need to transcend human resistance to using it in public places. Natural human computer interaction will be via SR, but huge improvements are needed for viability For European languages, we've already reached an acceptable rate of SR; 99%. I use Dragon Naturally Speaking and I find it effective. SR is difficult to reach at an 100% efficient level, but with a big commercial potential	

Comments

Enric Bas (03/17/2009)
In my view reaching a high level in SR, which could lead its generalized use without significant (and dangerous) mistakes, is -even its big potential - a very difficult task since it´s even harder to deal (managing it through machines, I mean) with the "emotional" side of communication in speeching than in reading/writing. Anyway, the potential of SR as a tool is just amazing and an exponential jump in communication (and in consequence, in social) terms

Art Shostak (03/06/2009)
As Microsoft has for many years creamed the best of China's IT young turks and placed them in China-based R&D Centers I expect major developments here soon coming the Middle Kingdom. Should reward all of us in short order.

Art Shostak (03/06/2009)
This is a sure thing, in that it meets a Tower of Babel challenge known since homo sapiens first encountered others whose speech puzzled them. As AI brings along the "intelligence" of inanimate objects - complete with "speech" in different native languages - this tool will become all the more strategic.

Language Translation (LT)

By John Sagi, with contributions by Florence Reeder

SUMMARY

Although LT (or "Machine Translation") has been used since the Cold War to translate text to text, a surge of globalization, electronic commerce, immigration, cross-cultural activities, and even the war on terrorism are driving its rapid development. Computer hardware is adequate but the software and methodologies are still lacking, especially for complex languages such as Arabic, Farsi, and Urdu that have received scant attention. The state-of-the-art is good enough for routine, informal text translation in tourism, news articles, and chatting, but it is not yet adequate for legal documents, military orders, and other tasks where precise meaning is crucial. Commercial use of LT includes internal use by corporations, "shrink-wrapped" products for off-the-shelf use in consumer markets, and large scale government applications.

DRIVING TRENDS

A Multi-Lingual World

1) English is the primary language of choice online by 30.5% of all Internet users, while the other 69.5% prefer their native languages. Moreover, local cultures are treasured, so a variety of languages is likely to remain widely used. (Internet World Stats, 2008)
2) Global companies require communication across languages to work in foreign markets. For example, Ford has been using an LT system for a decade to translate assembly instructions from English to German, Spanish, Portuguese, and Dutch. Ford, FedEx, and other companies use hybrid translation systems combining rules and statistics to translate 80% of their workload. (ComputerWorld 8/13/07)
3) There are currently 23 official languages in the European Union (Europa, 2008), and most of the world's 266 nations have their own language.
4) LT often features prominently at academic conferences. For example, Asia Online is using LT to "bring the world's content to non-English language speakers in Asia". (Jupiter Research 2008).
5) Demand for translation of more difficult languages has increased, particularly those of the Middle East. AppTek's popular PC-based LT software, TranSphere, is being extended for the Urdu language. (www.aramedia.com)
6) Manassas, Virginia-based Ciyasoft, specializing in Arabic languages, has developed a suite of products that reportedly scan and recognize documents in Farsi, Dari, Pashto, and Urdu with up to 95% accuracy. (Ciyasoft, Inc.)

Global E-Commerce Growing

1) Worldwide mobile telephone subscriptions reached 3.3 billion -- equivalent to half the global population (Reuters 11/29/07)
2) Verizon sends over 18.5 billion text messages in the U.S. alone every month, a growth of 250% each year for the last two years (CellSigns, 2008).
3) Facebook, the largest social network in the world, had 132 million visitors in June 2008 and is growing 150% per year. (Seekingalpha.com, 2008)
4) Google added 10 new languages to Google Translate in 2008, bringing the total to 23. (Jupiter Research, 5/30/08)
5) NEC introduced English-Japanese translation of text on mobile phones, able to translate

50,000 words in real time. SpeakLike offers instant message translation between English, Spanish and Chinese for a fee of ten cents. (Springwise, 2/13/08)
6) The Japanese have translation software on their PCs.

Much Government Activity
1) India is taking the lead in language translation, and expects to have systems that will translate between any languages in the world by 2010. (telegraphindia.com 11/15/04)
2) The U.S. government spent $430 B on international programs to combat terrorism.
3) John Negroponte, former Director of National Intelligence, published millions of pages of captured Iraqi files in Arabic on the Internet "in an unprecedented effort to… disseminate raw intelligence." (Boston Globe, 18 March 2006)
4) Over 660 thousand people from Mexico, India, Philippines, China and Viet Nam were admitted to the U.S. in 2007, requiring new efforts to translate languages more conveniently. (The Department of Homeland Security www.dhs.gov; NCLEA.gwu.edu)
5) The UN's World Intellectual Property Office translates 23 international treaties for 179 member states, with correspondence increasing 35% annually. (Bob McCallum, Translation Technology at the United Nations, www.crim.fr)

A Large Industry
1) There are 3000 human translation businesses in the U.S. alone. (Schreiber Translations, 2008; U.S. Census Bureau 2008).
2) The leading statistics-based LT vendor, Language Weaver, includes among its government customers the Office of Advanced Technology and Programs, the National Science Foundation, DARPA, In-Q-Tel, DoD, and SAIC. (LangageWeaver.com)
3) The leading rules-based LT vendor, Systran Software, has large Internet companies among its clients, including Google, Yahoo and AOL.
4) Language-Fon, a leading international cell phone translation service, boasts of "immediate access to over 150 languages via a toll-free number." It is supported by over 1,200 professional, certified linguists." (Languagefon.com)
5) Noted linguistics author John Hutchins observes that "there are over 1000 different LT packages on sale….and the demand is great." (Hutchinsweb.com)

OBSTACLES
Beyond Current State of the Art
1) The American Translator's Association notes that translation is more than replacing a word with its equivalent but must convey the style, tone, and intent of the speaker. (Yndigotranslations.com)
2) The most promising statistical protocol, Statistical Machine Language Software (SMTS), requires massive amounts of pre-translated documents to apply its "brute force" approach. The most promising rule-based protocols require thousands of linguistic syntax and semantics rules.
3) Keith Devlin, head of Stanford's Center of Language and Information, says LT will never equal the human linguist. (Scientific American, March 2006)
4) Futurist Joe Coates says, "In literature translation, linguistic nuances are still beyond the capability of machines." (Coates et al. 2035)
5) Journalist Jim Fallows observes, "Some [languages] are just impossibly hard… and

constantly changing." (Fallows, One-Button Translation)

Cultural Differences

Languages are different. Researchers at MIT and the Defense Language Institute studied translations among Spanish, Farsi, Arabic, Russian and Korean, and found "interesting relationships between difficulty of the text and various measures of machine translation." (Clifford & Jones, 2005; www.mt-archive.info) Soget, a leading human translation service specializing in Arabic, Farsi and Urdu, finds that machine translations are poor and sometimes incomprehensible. Many languages are rarely written, have not been developed into LT systems, and are often inconsistent. Arabic can be stored 20 ways in a computer. (www.Soget.com)

MATURITY FORECASTS

1) Ray Kurzweil thinks 25% of the world's population will communicate via automated language translation by 2012. (Language Weaver Blog 7/22/08).
2) Other forecasts suggest universal language translation services should be available by 2015. (Technology 1/29/04)
3) IBM expects to have accurate translation systems developed about 2010. (Technology Review 2/04)

MARKET FORECASTS

1) Annual sales of all types of translation globally is estimated to be $12B for 2007, $14 B for 2008, and is expected to reach $24 billion by 2012. Europe now generates more revenue than the U.S. (U.S. Census Bureau 2008; Scientific American March 2006; Language Weaver)

Survey Results

1	When will effective language translation (LT) in general become commercially available?* Enter the year from 2000 to 2030 when you think commercialization is likely to occur. If you think it has already occurred, enter the year between 2000 and 2030 when you think it happened.	**RESULTS** N = Number of responses SD = Standard Deviation Mean - 2013 N - 31 SD - 4
2	When will commercialization occur in non-western languages that are more complex and/or less well-developed, such as Arabic, Farsi, and Urdu?* Enter the year between 2000 and 2030 when you think this is most likely to occur.	**RESULTS** Mean - 2019 N - 31 SD - 5
3	When will effective LT reach the take-off point of 30% adoption to enter mainstream use?* Enter the year from 2000 to 2030 when you think 30% adoption is most likely to occur.	**RESULTS** Mean – 2018 N - 31 SD - 4

4	What will be the market for LT in 2010? Our data show large applications like LT (both human and machine) at about $14 billion today, medium markets like SR and EI about $2 billion, and small markets for VA and NSL less than 1 billion.* Please estimate the market size for LT at 2010.	**RESULTS** **%**	**N**
	Under $1 billion	10	3

Between $1-5 billion	19	6
Between $6-10 billion	26	8
Between $11-20 billion	32	10
Over $20 billion	13	4

5	Please offer any general comments about the prospects for LT.	**RESULTS** It is happening now. Core breakthrough will be pattern recognition with convergence of audio and visual. It will certainly make life easier for the human species! Effective to me means > 90% accuracy Automated language translation with speech synthesis will bring this technology to the consumer. Invaluable for R&D; indispensable for cutting-edge research In a global marketplace the demand for this technology is vast. Hopefully this will accelerate development. Automation with review by human translators is here - real-time machine translation is a huge goal. As a translator and user of translation software, I fin the prospects very exciting. Translating engineering documents is easy. Translating poetry is nigh impossible.

I think languages that lend themselves to LT, like English will dominate market fast.

Managing both structural difficulties and pattern changes (mainly induced by the use of IT) is key.

Comments

Enric Bas (03/17/2009)

First point is the inherent structural difficulty of some languages, full of symbolic dimensions and multiple meanings for a single word or sentence. So,sometimes it´s a very difficult task approaching an efficient translation without putting the words in a social context. Moreover, Language is something -as any cultural pattern and social institution- alive, so it changes according with social change... this is an additional problem in LT mainly due to the fast and radical changes induced by the extensive use of IT in communication (messenger, msn, facebook, etc..) which is producing important changes in the way of expressing (mainly writing) and understanding concepts.

Adam Gerber (03/11/2009)

As a translator and user of translation software, I find the prospects very exciting. However, in order for machines to reach parity with human translators in nuance and meaning, they must be able to translate paragraph by paragraph. Currently, machines are translating at the clause level or sentence level at best. Thus, machine translations often end up sounding like a non-native speakers with much loss in idiomatic, cultural, and nuanced use of language. We're getting closer, but we're not quite there yet.

Art Shostak (03/06/2009)

Appetite for this is VERY great, especially as it has payoff in pop culture uses in IT and can also be seen as another valuable aid to the ending of fear and zenophobia among the peoples of the world. Advances here may prove as rapid as my colleagues suggest - thereby faster than what I entered.

Image Recognition (IR)

By Devin Fidler, with contributions by Florence Reeder

SUMMARY

There is a wide range of Image Recognition (IR) technologies, and their development varies markedly. Optical Character Recognition (OCR) of cleanly printed Latin text, for example, is a mature technology today, with accuracy levels approaching 99%. OCR is rarely a stand-alone function but is often absorbed into a broader group of "intelligent document capture" technologies that read and correct information from scanned images of text based on context (IDC Worldwide Capture and Image Management Software Forecast and Analysis, 2008). However, recognizing clean print in many non-English languages is often much less well developed. Recent advances have been made in the recognition of text in candid video and photo images. Optical Handwriting Recognition (OHR) has been implemented for limited contexts like postal addresses or forms designed for handwriting recognition, but it is generally beyond the state-of-the-art. Mature image recognition technology would approach human-level recognition of printed and handwritten texts in any major language

DRIVING TRENDS

General Advances in the Field

1) Both commercial document management systems and search technologies are widely available that included capabilities for text recognition. The large imaging companies involved include Kofax, EMC Captiva or Datacap and by search companies like Google. (The Gartner Group, 2008; ECM Connection, 2004)
2) Google has recently adopted technology to permit search of scanned documents in PDF files. This opens up a large number of previously inaccessible documents such as academic papers, government reports, etc. (Computer World, 10/31/08)
3) SRI International has released a text recognition technology that can find and read street signs, name tags, and billboards in candid video and photographic images. (SRI International, 2008)
4) Internet Archive uses participants ("crowd-sourcing") to transcribe words that are difficult to recognize. This can be seen in the software that generates strings of distorted characters that some websites force you to recognize and type in order to establish that you are a person and not a computer. (Technology Review TR35, 2007)
5) A product called TypeReader advertises the ability to convert "a scanned image of a 700-page book into an editable Word file in a startling 6 minutes." This translates into the ability to perform enterprise-level OCR. (PC Magazine, 8/12/08)
6) A method has been developed that can train software to better recognize the handwriting of specific individuals. This technology has recently been applied to the letters of George Washington. (Document Recognition and Retrieval, 2007)
7) Google is sponsoring an open source project, called "Ocropus," with the goal of developing a high-level handwriting recognition system that can convert handwritten documents to computer text, but currently in English-only. (CNET news, 4/11/07)
8) Existing OCR becomes much more effective as computer processing speed and memory increase. Higher quality scans yield more accurate results but take longer and consume far more computational resources. (IMT, 2008)

9) Text recognition systems are often integrated into enterprise-level document management. For example, one vendor's integrated package includes, "diverse B2B integration services, such as integrated desktop messaging, bidirectional fax optical character recognition (OCR)/electronic document processing, etc." (Gartner, 5/28/08)

Applications for Complex Languages

1) Carnegie Mellon University's Million Book Project has partnered with organizations in China and Egypt to improve text recognition tools for both Chinese and Arabic to facilitate large book archiving projects in those regions. (Million Book Project, 2009) Carnegie Mellon has also created software that reads Chinese street signs and quickly translates them into English with up to 80% accuracy using a palm-size computer equipped with a small camera. (Technology Review, 3/03)
2) Researchers at Tsinghua University have announced the development of a text recognition system that processes Mongolian, Tibetan, Uygur, Kazak, Korean and some Arabic in addition to Chinese. (People's Daily Online, 1/30/07)
3) Optical Word Recognition technologies that seek to comparatively identify the image of a whole word, rather than reconstructing individual characters, are often regarded as a promising alternative to traditional OCR in complex cursive languages like Arabic, Farsi and Pashto. (Burrow, 2004)
4) Hindi, the second most populous language, is not covered by OCR packages.
5) Arabic is the fifth most written script, yet there are limited options for OCR and OHR in Arabic. Funding sources are becoming available, but it is likely to be years before tangible results are achieved. A handwriting recognition researcher argues that the cursive script, frequent omission of small marks that change word meaning and varying shapes of letters will make Arabic handwriting very difficult to process. (Burrow, 2004)
6) Researchers rank Chinese handwriting into three rough categories: regular or hand- print scripts have a handwriting recognition accuracy reaching 98% of characters, fluent scripts approach 82%, and individual cursive scripts characters are correctly recognized only 70% of the time. (Srihari et al, 2007)

Advances in Contextual Analysis

1) Historically, most scanned print text documents had to be hand labeled and tagged to extract relevant information. However, the field of intelligent document recognition is emerging to automate this task and allow greater depth of analysis into unstructured text data. (KM World, 7/07)
2) Nuance Communications is adapting its speech recognition applications to use parts of speech and grammar to recognize individual characters in scanned print documents. (IMT, 2008)
3) The Fuji Xerox Palo Alto Laboratory has been developing software designed to analyze the structure and content of untagged scanned documents and to extract key phrases that summarize individual sections. (Technology Review, 5/6/08)

OBSTACLES

1) A major hurdle that blocks broader application is that noisy documents prevent accurate OCR. (MINDS 2006-2007, Image Understanding Working Group, NIST/ODNI)
2) There is a general consensus that OCR engines can be “99.5% accurate at reading Latin characters under perfect conditions of clean paper, perfectly aligned, and printed in a known font. (ComputerWorld, 2002) But trouble arises with other languages, handwriting, highly stylized fonts, and unique layouts”. (BusinessWeek 2006)

MATURITY FORECASTS

1) An OCR industry executive predicted that by 2013, cell phones will be equipped with OCR programs sophisticated enough to allow blind people to read documents. (Journal of Visual Impairment and Blindness, 2003)
2) Microsoft has projected that high-level online handwriting recognition and translation among Latin text languages will be available around 2018. (Microsoft, 2008)
3) A computer scientist forecasts that human level handwriting recognition will not be available until around 2020, after very difficult "hard A.I." applications have been developed. (Singh, 2008)

MARKET FORECASTS

1) It is estimated that OCR and other “document capture technologies” that extract information by scanning images will grow at a compounded annual growth rate of 16.5% through 2011 from $670M in 2007. (Harvey Spencer Associates, 8/18/08)

Survey Results

1	Good Optical Character Recognition (OCR) technology is now commercially available. Please estimate when you think commercialization first occurred.* Enter the year from 1990 to the present when you think commercialization of OCR occurred.	**RESULTS** N = Number of responses SD = Standard Deviation Mean - 1997 N - 27 SD - 4
2	When will effective Optical Character Recognition (OCR) and other well-developed forms of Image Recognition (IR) in general reach the take-off point of 30% adoption to enter mainstream use?* Enter the year from 2000 to 2030 when you think 30% adoption of OCR is most likely to occur.	**RESULTS** Mean - 2013 N - 27 SD - 5

3	When will Optical HANDWRITING Recognition (OHR) become commercially available?* Enter the year from 2000 to 2030 when you think commercialization of OHR is most likely to occur.	**RESULTS** Mean - 2018 N - 27 SD - 5
4	When will commercialization of OCR generally occur in non-western languages that are more complex and/or less well-developed, such as Arabic, Farsi, and Urdu?* Enter the year between 2000 and 2030 when you think this is most likely to occur.	**RESULTS** Mean - 2022 N - 27 SD - 5
5	What will be the market for all forms of IR in 2010? Our data show large applications like LT (both human and machine) at about $14 billion today, medium markets like SR and EI about $2 billion, and small markets for IR and NSL less than 1 billion.* Please estimate the market size for IR at 2010.	**RESULTS** **% N**
	Under $1 billion	11 3
	Between $1-5 billion	56 15
	Between $6-10 billion	22 6
	Between $11-20 billion	7 2
	Over $20 billion	

		4	1
6	Please offer any general comments about the prospects for IR.	**RESULTS** **The potential is huge and will have a steep curve to over $20B by 2018.** **Good fast-moving bet in the West; much slower elsewhere.** **Less handwritten content is being created, so the market won't grow much, but methods are perfecting** **A real operative OCR is a very difficult task, in my opinion, because of the doc heterogeneity**	

Comments

Enric Bas (03/17/2009)
AI is key to developing an effective OCR tool because of the high complexity of some languages and heterogeneity of documents.

Non-Standard Languages

By Devin Fidler, with contributions by Florence Reeder

SUMMARY As various Human Language Technologies (HLTs) advance, more applications are being developed to adapt these technologies to non-standard languages (NSL). NSL is an umbrella term that covers any language not codified in standard format, making it difficult to process using mainstream language technologies. This includes non-written languages that are poorly understood; those that have not been developed for machine analysis; those found in informal media, such as online chat, text messages, e-mails, blogs, web pages; dialects (e.g. Spanglish); and communications by non-native speakers of a language. At present, machine interpretation of NSL relies on a patchwork of technologies that are tailored to specific problems and are labor intensive. Mature machine interpretation of NSL would rely on far less human input and would be fully integrated into standard HLTs, allowing them to be used by more people in more situations.

DRIVING TRENDS

General Advances in the Field

1) Rule-based HLTs and statistical HLTs each have strengths and weaknesses. New machine translation technologies combine these two systems to provide a more flexible approach. (Computerworld 8/13/07).
2) Inxight has developed a program that applies text analysis and NSL to thirty different languages, including NSLs in Spanish, Chinese and Farsi. (Inxight, 2008)
3) An institute in Japan has developed a technology that can automatically identify 100 different languages on Internet applications. (The Nation, 10/23/06)
4) Nuance is developing SR technology to recognize dialects, including Spanish dialects. (IEEE Transactions on Audio, Speech, and Language Processing, 11/07)
5) Computational linguists have devised programs to interpret “netspeak.” (e.g. btw). (Journal on Document Analysis and Recognition, 12/07)
6) An annual workshop studies the machine analysis of NSL in multiple languages, especially on the Internet and in text messages, (2007 and 2008 Workshops on Analytics for Noisy Unstructured Text Data)
7) A system to process Spanglish is being developed at the University of Texas and can be used on other bilingual dialects. (Language Information Technologies, 7/18/08)
8) Microsoft's Beijing Lab has introduced an "English Writing Wizard" to help non-native English speakers detect errors in their text. (Technology Review, 7/26/04)
9) IBM has developed an automatic technique for partially cleaning up NSL from text messages for data mining. (Workshop on Noisy Text Analytics, 2008)
10) The Interlingua project has developed an automatic translator that is able to translate NSL from email, newsgroups, and chat forums that contain both Catalan and Spanish (Journal of Computer Mediated Communication, 11/03)
11) Among the leading developers of unstructured text analytics applications are SPSS and Anderson Analytics. (Text Analytics Summit, 2008)
12) ScanSoft has released SR software with learning capabilities that gets better at recognizing and interpreting a person’s accent. (CNET News.com 5/17/04)

Need to Translate Large Quantities of Material

1) According to a 2008 survey conducted by Forrester Consulting, 41% of US companies with 20,000 or more employees read or analyze outgoing email, which often includes NSL. (Information Week Online 5/20/08)
2) More than 35 million adults in the United States are native speakers of a language other than English, and there are approximately 15 million speakers of the Spanglish dialect (Journal of Hispanic Behavioral Sciences, 05).
3) Linguist David Crystal argues that “netspeak” represents an important new kind of language use, and predicts that it will have a "much wider stylistic range as the medium is adapted to suit a broader range of communication." (New York Times, 12/13/01)
4) A computer linguist from the IBM Indian Research Lab estimates that up to 70% of all digital text is “noisy,” either because it is in a difficult format (e.g. instant messages, blogs) or because it is abbreviated or ungrammatical. (Subramaniam, 9/26/07)

Business Intelligence Uses

1) A commercial market is emerging around machine analysis of internet NSL. Several companies license business software to analyze millions of unstructured internet media to determine the demographic profiles and opinions of individuals discussing a given product. (EContent, 5/1/06)
2) The IBM Indian Research Lab has developed software that generates business intelligence based on machine interpretation and analysis of phone calls, email, and text messages in call centers. (Wireless News, 3/24/08 (Workshop on Noisy Text Analytics)

OBSTACLES

Great Complexity

A member of IBM’s Indian Research Lab speculates that applications which require accurate correction of NSL “may be an exercise in futility”. For example, some text messages require knowledge of context, which makes correction difficult even for human editors. (New Yorker, 6/23/08; International Association for Pattern Recognition, 04/07)

Large, Costly Programs Required

A language systems researcher at MITRE thinks the development of NSL is slow because large amounts of human-analyzed data are needed, which is labor-intensive and costly. (Interview) Another said that breakthroughs in NSL are rare because they are “more about perspiration than inspiration." (InfoWorld, 11/4/04)

MATURITY FORECASTS

1) British Telecom’s in-house futurist Ian Pearson predicts that NSL will be sophisticated enough to route messages to specific recipients based on analysis of their content by 2020, (Wireless Life 2020, 2006)

MARKET FORECASTS The unstructured text analytics market, the largest application in NSL, is estimated at $200-250 million annual revenue in 2007, with a 25-50% annual growth rate. (Intelligent Enterprise, 6/20/08)

Survey Results

1	When will effective Non Standard Languages (NSL) in general become commercially available.* Enter the year from 2000 to 2030 when you think commercialization is likely to occur. If you think it has already occurred, enter the year when you think it happened.	**RESULTS** N = Number of responses SD = Standard Deviation Mean – 2017 N - 29 SD - 5
2	When will commercialization occur in non-western languages that are more complex and/or less well-developed, such as Arabic, Farsi, and Urdu?* Enter the year between 2000 and 2030 when you think this is most likely to occur.	**RESULTS** Mean - 2023 N - 29 SD - 5
3	When will effective NSL reach the take-off point of 30% adoption to enter mainstream use?* Enter the year from 2000 to 2030 when you think 30% adoption is likely to occur.	**RESULTS** Mean - 2023 N - 29 SD - 4
4	What will be the market for IR in 2010? Our data show large applications like LT (both human and machine) at about $14 billion today, medium markets like SR and EI about $2 billion, and small markets for IR and NSL less than 1 billion.* Please estimate the market size for NSL at 2010.	**RESULTS** **% N** Under $ 1 billion 59 17

	Between $ 1-5 billion	34	10
	Between $ 6-10 billion	7	2
	Between $ 11-20 billion	0	0
	Over $ 20 billion	0	0
5	Please offer any general comments about the prospects for NSL.	**RESULTS** What "effective NSL in general" means is not clear. Similar to LT for non-western. Needs convergence of visual and audio pattern recognition. High accuracy will not be cost effective for a decade. A far less significant matter than the others in the Survey: trivial information. For monitoring of Internet and text messaging, NSL sounds useful, for other languages, low demand Context is critical. Good translation may never be possible. The difficulty of developing an efficient NSL is proportional to the need of a tool like this (top)	

Comments

Enric Bas (03/17/2009)
Coming back to my first comment, I would like to outline the fact that most of languages (maybe not the ancient ones like latin, greek, or other not widely spreeded) are just live beings.. they have mutations both referred to the structure, the lexic or the meanings of words and sentences. The dynamics of this change depends on the dynamics of social change in communities. We are in a time where -mainly in societies using extensively ITCs- language changes every day, every minute. So, the "formal languages" (spanish, english, etc..) are just dissapearing as a solid reference; they are becoming, like the rest of social institutions, "liquid" (mutable, complex and with diffuse limits) according with Zygmunt Bauman. It´s just basic to work on it: if we pretend to understand the expectancies (fears,whishes..) and behaviour (opinions, acts..) of people whe HAVE to know what they mean in the communication process they use to express them.

Art Shostak (03/06/2009)
As the war/peace relevance here is minor, and the profits to be made from making progress here (in commercial matters) is minor, I do not see this as a site where major R&D effort is likely for years to come.

Emotion Recognition

By Richard Varey, with contributions by Florence Reeder

SUMMARY A variety of sensors and cognitive-physical theories are enabling computers to recognise emotions from facial expressions, speech, text, gaze, and physiological conditions and respond appropriately. Applications include recommender systems that evaluate preferences in voting or consumer tastes; security systems to detect lying and suspicious behaviour for intelligence and police work; shopping systems that detect frustration in consumer behaviour; medical system to diagnose stress; personalised robotics; and e-learning and training systems. Some research suggests that interpretations of EI data can be spurious, which raises questions about accuracy and meaning. Because of cultural differences that are hard to interpret, advances are largely valid only for Western societies, while progress is limited for Arabic, Farsi, Urdu, and other more exotic cultures.

DRIVING TRENDS

Research Advances

1) MIT researchers are developing a voicemail system called Emotive Alert that labels messages according to the caller's tone of voice and identifies which are most urgent. (The Washington Times, August 2007)
2) Philips Research has developed a robot-human interaction system that recognizes emotions and responds appropriately. (TechnologyReview.com); Researchers at the University of Hertfordshire are working on robots that interact with people emotionally. (www.news.bbc.co.uk)
3) Nicu Sebe at the University of Amsterdam has used an EI program to determine that the Mona Lisa is 83% happy, 9% disgusted, 6% fearful, and 2% angry. (Time, October 2006). His team is working with Unilever to map a subject's face into regions that determine their level of pleasure, and plans are to launch a consumer version that reads consumer responses to products. (www.engadget.com, July 2007)
4) The W3C Emotion Incubator Group, with representatives of 16 institutions from 11 countries, studies "the prospects of defining a general-purpose emotion language to be usable in a large variety of technological contexts." (www.w3.org/2005).
5) EI is used in “conversational”’ or “narrative user interfaces,” where synthetic characters (avatars) converse with users. (Fraunhofer IDG, 2004).
6) Dr. Kate Hone at Brunel University is developing EI systems that analyze people's facial expressions, tone of voice and even sweat levels using webcams, microphones and a special mouse. (New Scientist, January 2006).
7) Deutsche Telekom is developing an intelligent avatar tutor that reads emotions and responds appropriately. (http://emotion-research.net

Commercial Applications

1) Patents have been issued for a speech analyzer to determine emotional states and for an man-machine interface that simulates emotions (www.patentstorm.us).
2) Prosody, the study of patterns of stress and intonation, has been widely used to identify emotions and accents. (www.isca-speech.org)
3) Applications of EI are proliferating for use in employee selection, workplace ethics, customer analytics, public safety, national security, and monitoring of machine user

behaviour (T & D, 2004, cover story).
4) Affective computing is being applied to learn customers attitudes by monitoring the content of telephone, web site, and other interactions (Bank Technology News, May 2008; Bryan Bergeron, The Eternal E-Customer: 2001, McGraw-Hill)
5) Noldus Information Technology Inc. announced FaceReader - the first automated software capable of automatically detecting and analyzing facial expressions. (Product News Network, October 2007).
6) SnapFeel detects emotions in text messages in real-time, while Nemesysco uses voice analysis in flight simulators to detect mental stress.
7) Affective Media will soon be selling software that detects drowsiness and frustration as a driver's talks to the in-car navigation system and then attempts to wake the driver up or calm him down. (The New Scientist, January 2005).
8) Tanja Schultz at the Language Technologies Institute, Carnegie Mellon University, studies how SR algorithms can function in multilingual environments using non-verbal cues, such as emotion, focus of attention, and information about the speaker.
9) Autonomy's Etalk systems can monitor call center conversations in real time to detect when a customer is starting to get angry and call in a supervisor to take over.
10) Emotion Mining Company has a patented process for measuring conscious and subconscious emotions for marketing, branding, team building, and to facilitate change initiatives. (http://www.EmotionMining.com)
11) Wisconsin Physician Services Insurance Corporation uses emotion-detection software to identify unsatisfied callers. (Chabrow, InformationWeek, 4/4/2005).
12) The Danish firm iMotions developed Emotion Tool to infer emotions based on the excitement an image creates. (www.pr-inside.com and www.imotions.dk)
13) Emotiv Systems has launched a system to detect facial expressions of video game players, measure emotional states, and perceive a player’s conscious thoughts in real time. (Thomas Majewski, www.associatedcontent.com, March 2007).
14) Idiometrics, a leader in facial recognition, announced the release of FaceTrace to measure emotions from facial expressions. (Press release, August 24, 2008).
15) Clarabridge Company is the fastest-growing firm in the EI market. Its clients include Gaylord, H&R Block, Intuit, Lowe's, Marriott, Oracle, Sage, TNS Global, The Gap and Gallagher Benefit Services Inc. (Business Wire 2008).
16) RightNow Technologies uses EI features to gauge customers' opinions by applying emotional ratings to text-based customer communications. (www.cio.com.au).

OBSTACLES

General Difficulty

Research has yet to discover facial expressions which uniquely correspond to specific emotions. Few speech sounds, for example, express only a single emotion but usually involve several, even discordant ones. Accuracy is also hampered by unusual factors in the situation, the speaker’s state of mind, and so on. Research has shown that actual intentions are often more important than emotional displays and require considering social and interpersonal context to reliably predict intentions. For example, an agitated traveller arriving at an airline check-in desk may have argued with a loved one, and not be intent on violence or smuggling. (Ward & Marsden, Interacting with Computers, 2004). To counter growing frustration with automated call centers, SouthWest Airlines, Siemens, and other corporations

are staffing their customer service lines with trained operators. (The Philadelphia Inquirer, December 2007)

Cultural Differences
A multicultural understanding of human behaviour is arduous because of the spontaneous nature of speech, the complex ways cultures express emotions, and the unique attributes of each person. Observing people using automatic emotion detectors is often resisted on grounds of human rights, civil liberties, and privacy. (The Culture of Emotions, Dr Lewis Opler).

MATURITY FORECASTS
1) A researcher at MIT said "Machines with emotional intelligence will be built within the next five to ten years." (IDG News Service, October 1999)
2) TechCast forecasts that advances in speech recognition, artificial intelligence, and computer power suggest the keyboard and mouse may yield to an "intelligent interface" about 2014 in which users simply converse with smart computers.
3. BT's future timeline forecasts that a toy will be able to respond to its owner's voice with a variety of emotions by 2006-2010. Sony's Aibo dog robot can already simulate anger, fear, surprise, dislike, sadness and joy. (www.mirror.co.uk, July 2005).
4) An authority thinks software that can sense emotion, hesitation, aggression, hostility, anger, etc. will appear in more common applications within five years. (Lance Winslow, www.WorldThinkTank.net).

MARKET FORECASTS
1) Donna Fluss, president of DMG Consulting, says speech analytics including EI is expected to grow 100% in 2007 and 2008 (1to1 Media 2006; Carlson MarketingWorldwide).
2) Daniel Hong, an analyst at Datamonitor who specializes in speech technology, estimates that EI generates $2 billion annually. (Wired, April 2008).
3) Speech analytics, including EI, is just entering the commercial market. Gartner Research estimates that fewer than 1% of call centers are using the technology. (Davies, Unravel the Complexities of Call Center Speech Analytics)

Survey Results

1	When will effective Emotion Identification (EI) in general become commercially available.* Enter the year from 2000 to 203 when you think commercialization is likely to occur. If you think it has already occurred, enter the year when you think it happened.	**RESULTS** N = Number of responses SD = Standard Deviation Mean - 2014 N - 28 SD - 4
2	When will commercialization occur in non-western languages that are more complex and/or less well-developed, such as Arabic, Farsi, and Urdu?* Enter the year between 2000 and 2030 when you think this is most likely to occur.	**RESULTS** Mean - 2021 N - 28 SD - 5
3	When will effective EI reach the take-off point of 30% adoption to enter mainstream use?* Enter the year from 2000 to 2030 when you think 30% adoption is likely to occur.	**RESULTS** Mean - 2019 N - 28 SD - 4
4	What will be the market for EI in 2010? Our data show large applications like LT (both human and machine) at about $14 billion today, medium markets like SR and EI about $2 billion, and small markets for IR and NSL less than 1 billion.* Please estimate the market size for EI at 2010.	**RESULTS** **% N** Under $ 1 billion 25 7

Between $ 1-5 billion	54	15
Between $ 6-10 billion	21	6
Between $ 11-20 billion	0	0
Over $ 20 billion	0	0

5 Please offer any general comments about the prospects for EI.	**RESULTS** Key element for LT and NSL. Commercialization of EI precedes and enables both. Exotic and alluring, EI is a sure thing - BUT, technophobes will curb advances. I think EI will be more difficult, smaller market and fraught with false positives/negatives EI efficiency depends on our ability to understand (and to interprete) cultural patterns

Comments

Enric Bas (03/17/2009)
Interpreting the emotional by using ICTs is a controversial thing. Everybody is feared up to be "accessed" til the point of knowing its deepest feelings (that would be a kind of massive personal/social control). But at the time, ICTs are promoting a kind of exhibitionist behaviour within youngest people, who are not afraid to show their (real or constructed) emotional state by writing (emoticons, thoughts) or showing private photos and even their emotional network: the social networs, like Facebook, are a good example of this trend. One more point I would like to outline is the fact that cultural patterns determine the way people

express their feelings and interact with others, and this is for me the central point to design an effective tool for analyzing emotional responses.

Phillip Nelson, Ph.D. (03/08/2009)
We can radically accelerate progress in these 5 fields, and any field that similarly has many different excellent projects racing to market. We need to find a way of coordinating efforts and quickly sharing breakthroughs and best practices across these teams (and nations). There are already partial solutions. We need to push a systemic solution in all critical fields to radically accelerate technology development. The pay-off would be huge.

Phillip Nelson, Ph.D. (03/07/2009)
See Wed. 8 p.m. Fox HD Channel "Lie to Me" for an exciting education on the uses, implications & limits of EI. Integrate with SR to accelerate both.