

Data Quality Goes International



Data Quality Goes International

Today more than ever, organizations are conducting business beyond the limited confines of their home countries and are actively pursuing opportunities across the globe. To enhance their ability to capture these opportunities – or simply to respond effectively to international government regulations – organizations need a data quality solution that is geared to meet the demands of international data quality.

Such a solution is critical, because the quality of data in international markets is a real concern, regardless of country or region. Poor quality data can not only negatively impact an organization's global efforts, but it can also contaminate the quality of other data if the organization isn't careful.

However, finding the right "global" data quality solution can be difficult, because all data quality solutions are not equal. There are a number of solutions that are being marketed as "international" solutions, but in reality, they offer little more than some support for postal information in a number of countries. Postal information can be important, but organizations should not be swayed by the number of countries that a solution can supposedly support.

Instead, companies should consider the depth of country support that the solution offers – and the depth within the countries or regions in which they are actively conducting business. Organizations should therefore look at how well the solution "understands" the language and culture of a particular region. They should also consider the solution's ability to support the script of the region – if it's the Asia-Pacific region, then Unicode is mandatory – and its support of postal data and government postal standards (e.g., AMAS, SERP, in addition to the U.S. postal standard CASS). Finally, organizations should consider its data enrichment capabilities, which will be critical in terms of filling out missing information or augmenting data records with value-added details.

The stakes in the global marketplace are high, and many organizations have stumbled because they weren't prepared to handle both the demands of the international marketplace and its opportunities. An international data quality solution is a requirement that should not be overlooked. However, even the right international data quality solution will provide limited results, and only go so far in supporting an organization in its global efforts, if the organization doesn't first get the rest of its house in order in terms of data quality.

The Global Economy

Today, more than any other time in history, organizations are conducting business beyond the limited confines of their home country or territory, and are actively pursuing opportunities across the globe.

This is not surprising, given the fact that governments across the globe are working to lower barriers to business in their countries, while at the same time working to open trade and other business in other countries. Of course, this is not to suggest that various trade barriers are coming down, or that governments are honestly trying to limit the boundaries to conducting business within their borders. While many governments are indeed working to open up the borders of their trading partners, these same governments are also fighting to restrict trade in some fashion within their borders, if only to protect domestic business. Nonetheless, many geopolitical borders are coming down, or at least being weakened, for the sake of business.

The movement to open up geopolitical barriers is also being facilitated by the movement of people, by the advancement of knowledge and the development of technology. Regardless of the underlying reasons for this monumental migration of

people over the past few decades, this movement is having a profound impact on the manner in which many organizations are conducting their business. Foreign labor at all levels is gradually changing the face of businesses across the world, and with the rise of outsourcing businesses can leverage the talents of individuals from virtually any place in the world.

Furthermore, the spread of knowledge and technology have been instrumental in the increase of globalization. Science has certainly become more global, and the benefits and discoveries of science are refining current businesses and creating new businesses that could not have been imagined even fifty years ago. There are a number of biotech organizations, for example, that in recent years have come into existence because of new scientific discoveries and the absence of physical boundaries to scientific study and progress.

Technology, in particular, is fueling the international movement of business and other areas of endeavor. In fact, the advancement of technology – the Internet, mass communications, computers, spreadsheets, databases, etc. – may have initiated some of the most profound changes not only in the manner of which organizations conduct business but also in the way in which their businesses transcend geopolitical boundaries. A generation ago, doing business in places like China or the Middle East was reserved for specialized industries or organizations large enough to afford the funds needed to open offices in these far-flung areas. Today, even the average person can routinely transact business with someone from any region of the world who has an Internet connection.

Technology is also responsible for the speed that trade and capital can move across the globe, and this in turn brings markets closer together and encourages the development of markets, products and services that would have been unthinkable in another era.

Finally, these and a multitude of other dynamics of globalization are altering the culture, politics and even environments of both organizations and countries. Not too many years ago, Western Europe was fairly homogeneous in terms of peoples and cultures (especially within limited geographical boundaries), but this is changing quickly. As a result, organizations are changing in order to survive and prosper in what could be an entirely new world within a few decades.

Data Quality and the Global Economy

Of course, these broad generalizations only scratch the surface in terms of the forces that are changing the world's marketplaces and the impact that these forces are having on countries, people and organizations. None of these statements can really convey the pace and the degree to which globalization is occurring. And it doesn't address the problems attendant on globalization, such as the social and cultural issues brought about by human migrations or the environmental issues caused by overdevelopment. Equally important, it doesn't begin to address the issues faced by organizations that either want to capitalize on the opportunities offered by globalization or simply address some of the new requirements in the global marketplace.

For organizations, globalization is increasingly an important factor of doing business. For those organizations that want to thrive in the new global environment and take advantage of opportunities created by the opening of new markets, new ways of looking at business technologies and processes are required. Among the most important, underlying technologies required is data quality.

Certainly, data quality technology is not new, but given the advancements in the technology and the need for high-quality data for optimal business operations, the

right data quality technology could make a difference between success and failure in the new marketplace. Indeed, the ability to address international markets fully and effectively is dependent not only on high-quality data, but also on the ability to support and understand the cultural nuances of foreign data.

One of the most common misconceptions regarding international data is that Western languages and cultures pose few problems for Western organizations, unlike the issues surrounding Asian languages and cultures. But while Western languages pose fewer problems than Asian languages (due to the complexity of character-based languages), there are many cultural issues inherent in Western languages that can cause problems, slow business and prevent organizations from succeeding in what are seemingly familiar markets. For example, the conventions governing the structure of names and addresses are different in Spain than they are in the U.S., and the failure to recognize these differences could increase the costs of doing business in both countries, and perhaps even sow the seeds of failure.

Globalization: An External and Internal Issue

But globalization is not simply an external issue. For organizations to capitalize on the international marketplace, they must also be able to integrate their internal structures, including business units, systems, applications and information silos. Internal globalization will allow an organization to capitalize on external issues across the international marketplace (including compliance) quickly and effectively, and in a way that no organization with such information silos can.

This white paper will address how organizations can use data quality to prepare themselves for the global market. It will consider how data quality technology can support foreign languages and cultures, both in Europe and throughout the rest of the world, including the Asia/Pacific region. And it will consider globalization as an internal issue – that is, unless organizations get their internal data in order, it will be extremely difficult for them to leverage effectively even the best of global data quality solutions, and hence make the most of their global efforts. (These efforts include avoiding penalties for failing to comply effectively with foreign government regulations.) While this paper makes no attempt to address every issue related to globalization, data quality should nevertheless be considered one of the most crucial issues facing any organization that wants to pursue international markets effectively.

Using Data Quality for Global Business

Conducting business globally should ideally present only a handful of difficulties for organizations, especially if that business is conducted over the Internet. With the Internet, along with a handful of global trading partners, an organization can expand its business to practically any location that supports a high-speed Internet cable or is within reach of its partners. Many organizations are prospering in this very manner.

However, the ease with which one can conduct business through these channels is deceptive, for while they do offer easy-to-reach opportunities, they also offer pitfalls that can trip up any business not sufficiently aware of the demands of the global marketplace. Conducting business virtually can be a simple exercise, even with different currencies – pounds versus dollars versus euros – or Web pages in different languages.

These matters are relatively easy to resolve – add currency converters and hire language translators – especially compared to the complications that can arise if the business is not virtual. The situation becomes more complex when an organization has to send mailings, deliver products or provide specialized customer service to customers outside its country of origin. In these instances, an organization might have

to grapple with language and cultural issues, as well as name and address formats that can be foreign to the organization.

Global Data Quality Extends Beyond Postal Information

To overcome these difficulties and to open channels for effective international business and communication, organizations must apply data quality technologies and processes that support business and communications not only in North America and Europe but throughout the rest of the world. Problems as simple as the conversion of currency and as complex as the parsing of multiple languages in a single document can be handled with the right kind of data quality capabilities – the sort of capabilities that can support language and cultural issues, understand non-Latin encoded languages and provide enrichment for foreign data.

Unfortunately, some organizations believe the only necessary tool for effective international business is a database of foreign postal information. International postal information is indeed valuable, but these databases tend to be limited in the type and range of information that they provide. They are also rarely certified by legitimate postal authorities. Certification by postal authorities is important, because this certification ensures that the information is accurate and detailed to a certain level – assuring that street names, addresses, postal codes and other factors are correct.

The most important postal certifications currently are the United States Postal Service's CASS (Coding Accuracy Support System) for U.S. address information, SERP (Software Evaluation & Recognition Program) for Canadian address information, and AMAS (Address Matching Approval System) for Australian address information. There are other legitimate postal certifications, such as in Europe, and some that go beyond merely certifying the accuracy of the postal information. The USPS, for example, not only provides certified information on U.S. streets and addresses that are not yet on maps, but also provides this information for address changes of individuals and businesses that might not yet be in other directories.

Still, postal information – which should be considered alongside a vendor's other capabilities – provides only a limited range of information and does not offer specific information regarding cultural entities and the use of languages within these entities.

Understanding Unique Cultural Behaviors

To do more than move mail from one location to another, deeper international support is required. This kind of support should include solutions for specific countries or regions – solutions that “understand” the rules governing the cultural behavior of a specific region. These capabilities should encompass geographic and language-specific rules and standardization conventions for an understanding of spelling, syntax, phonetics, vocabularies and related issues. Without these capabilities, it becomes difficult for organizations to support customer relationships in a host country. It also becomes difficult for them to develop business opportunities in the host country or simply to comply with the rules and customs governing the interactions between organizations and with government entities.

Also critical is support for the visual format of native text or character sets. Even though English is the unofficial business language worldwide, most business is transacted in the native language of an organization's host country. This, however, creates many problems for Western organizations that are expanding beyond their geographies and for others that are pursuing business in the West.

If organizations are only conducting business in North America and Western Europe, then support for the single-byte character sets that are characteristic of such Western languages as English, French, Spanish and German would usually be sufficient.

However, business is growing significantly in Asia and the Middle East. China, for example, boasts the world's second highest GNP, and business there is growing at a faster rate than most North American and European nations. With this kind of growth, there is an increasing need to support multi-byte character sets that are typical of languages such as Chinese, Japanese, Korean, Hebrew, Thai and Arabic.

The Importance of Unicode and Data Enrichment

Supporting these character sets (in addition to single-byte character sets) can be achieved through the support of Unicode. Unicode is an industry-standard coding system that enables texts and symbols from diverse writing styles to be consistently displayed by computers (or perhaps through the International Organization for Standard [ISO] ISO 10646, which is also an industry-standard coding system for diverse languages). Without Unicode, data quality processes cannot be effectively applied to multi-byte languages. Not only will the value of high-quality data not be extended to such text, but the quality of an organization's other data could be contaminated with the introduction of data whose quality cannot be assured.

Finally, to round out the solution, there should be strong support for data enrichment. International data is no different than any other data when it comes to enrichment – that is, the usability and value of data can be increased with relevant supplemental information. This information might be geocoding, census tract numbers, country codes, phone numbers, time zone appends, gender information, entity information, country-specific information, etc. Lifestyle information from third-party providers can also be valuable for international business. This information could be critical in terms of understanding customers and pursuing business and other opportunities. Indeed, why market a Ferrari to prospects in Sierra Leone when the GDP per capita is only around \$800 USD and more than half the country is below the poverty line?

Ultimately, to support the increasing demands of the global economy, and to capitalize on opportunities in such increasingly important areas such as the Asia/Pacific region, prospects should consider the international capabilities of the data quality solutions they are evaluating. Regardless of whether data is in Thai, English or some combination of languages, its quality should be good enough to ensure success and an underlying understanding of the business and its customers.

Using Data Quality Technology for Internal Globalization

Unquestionably, there are challenges and benefits in the global marketplace. While many organizations are actively pursuing these challenges, most organizations are increasingly having to address global issues – if only because the Internet has expanded the customer base of most organizations, parts and product suppliers are now global, it is likely that any organization with partners will have some international connections through these partners.

The previous sections of this paper focused on the benefits and challenges of the international marketplace. Leveraging global opportunities is not always an easy task, and there is no guarantee that a “pot of gold” awaits at the other end. To succeed, organizations must have an understanding of not only the language of a particular region, but also the culture and cultural forces that govern the region. They also need to understand and contend with foreign (and domestic) regulations of all kinds, which by themselves can derail even the most far-sighted efforts to expand beyond the North American market.

Nevertheless, the odds of international success are obviously not insurmountable, and organizations that are carefully calculating both the opportunities and the demands of

the international marketplace will consider data quality tools and processes as integral pieces of any international effort. Regardless of language, culture, geographical location or government regulation, data quality is a crucial requirement. No international effort will fulfill its promise if it is founded on poor quality data.

But simply implementing a data quality solution is not enough by itself to overcome the challenges of poor quality international data. For one thing, the solution needs to be truly international in scope, providing capabilities that extend beyond some foreign postal information. This postal information is certainly an important aspect of any international data quality solution. But a broadly effective solution must also provide a deep understanding of a region's language and culture, its city and street-level information, and international data enrichment capabilities.

These are complex requirements, but the demands of the international marketplace are themselves extraordinarily complex.

However, the true keystone for any organization that wishes to succeed internationally is not simply international data quality capabilities. While critical, organizations will still have difficulty succeeding internationally if they don't look inward and globalize their own domestic business first.

Global Data Quality from Within

One common problem among many organizations is that they don't consistently perform some tasks on a global scale internally. They might implement an application or a set of procedures to suit the needs of one part of the organization, but they neglect to consider the impact this implementation might have on the organization as a whole, particularly if there are divisions overseas that are affected. For example, some companies remediate critical business problems in one operating unit without necessarily extending this remediation to other departments in the organization – departments that seemingly don't need such remediation. Of course, they might also try to control costs by incrementally remediating a problem without considering the potential impact of this remediation on the organization as a whole.

In terms of data quality issues, many organizations leverage data quality technology for particular operating units or systems. But for many reasons, including short-term costs, they refuse to leverage it across the rest of the organization. While the attention to data quality issues provide benefits regardless of where it is used in an organization, the greatest benefits are reaped – and success on an international level is ensured – when this focus is extended to every part of the organization. Data quality, therefore, should be applied consistently throughout the organization. Data quality must be a global issue internally.

Achieving Internal Consistency and Synergy

The problem here is relatively straightforward – individual operating units, systems or applications in which data quality technology might be leveraged exclusively within an organization are rarely the only areas in an organization that require high quality data. If the organization is to reap the maximum benefit from its data – and if the organization is to completely understand itself and account for its business effectively and financially – then all of its operating units, systems and applications need the same care extended to individual systems.

The benefits of data quality technology and high quality data can be diluted if some parts of an organization are beneficiaries of the technology, while others are not.

No department, system or application is truly an island. Few operating units are entirely cut off from every other operating unit, such that data from outside its

borders does not flow through it. This flow of data opens up the possibility of data contamination, limiting the overall effectiveness of whatever data quality technology or efforts are being made. This consequently limits the value of the data to the organization as a whole. But even if the flow of data can be controlled, which in some instances is the case, the organization again loses the benefits that come from consistent, high-quality data across the enterprise.

Organizations should never think of data quality as a departmental or application-specific issue. Instead, they should consider data quality as an organization-wide or global issue. The degree to which data quality is considered a global asset will directly impact the organization's international efforts. If the organization does not get its domestic data in order, then it will also be extremely difficult to get its international data quality efforts in any reasonable order.

Reaping the Benefits of Global Data Quality – Next Steps

So what can be done in order to enhance an organization's efforts to reap the benefits of the international marketplace while, at the same time, enabling it to meet the demands of both cultures and governments across the globe? Organizations should start by committing to these three overriding steps to succeed on a global scale:

1. **Organizations should get their entire house in order, in terms of high quality data.** They should formally acknowledge the impact that consistent, reliable and accurate data has on business, and implement a data quality solution that meets company objectives. Even if the organization has no interest in international business, operating on the highest quality data possible is the only way to operate optimally.
2. **Organizations can begin to implement internationalized solutions that support their global efforts.** This comes with the understanding that high quality data is crucial to all parts of the business. Beginning with this high quality data, companies can tackle the data problems and issues outside of their boundaries more efficiently and cost-effectively. At the very least, they eliminate half of the data problems that they will encounter in their efforts.
3. **Organizations should commit to applying global data quality standards internally.** The last thing any organization interested in the international marketplace wants is to battle two separate culture wars – one that exists outside of the organization, and the other that resides inside due to the effects of poor-quality enterprise data. To succeed globally, an organization's global data quality efforts should be extended internally. When this happens, the entire business operates on the highest possible quality of data.

Conclusion

This paper has examined how organizations could enhance their efforts in the global marketplace by using the right kind of data quality solution. Data quality is crucial to all business efforts, foreign and domestic. Therefore, any organization that wants to position itself effectively to succeed internationally, capturing both the opportunities of the international marketplace and avoiding the pitfalls of some of the international regulations, should make finding the right kind of data quality solution a top priority.

One of the core problems confronting organizations in every international market is the quality of its data. In many regions the quality of this data can be poor, and not only will this impact an organization's global efforts, but it can also contaminate the quality of its other data if the organization isn't careful.

But while virtually every data quality company markets some kind of international data quality solution, not all data quality solutions are equal. And not all international data quality solutions are worth the cost and effort made to market them.

Finally, there is the question of updates. In many instances, this information is updated on a monthly basis, but in industrialized areas this is certainly not frequent enough. How frequently should the postal data on China be updated? And if the postal data is not up-to-date, how useful will the database be?

Moving Beyond Postal Information

Postal information is indeed important, but it is only one small aspect of an effective international data quality solution. Some valuable postal information, or postal support, should be the support for postal standards such as AMAS, SERP and the U.S. standard CASS. These and other standards ensure that the organization has access to the highest level of quality postal information, and ensure accurate mailings and even postal discounts.

But more important to any international data quality solution is the depth of support that it provides. A dozen countries with deep support is preferable in most cases to a solution that has some postal data on over 200 countries. What is meant by "depth" is how well the solution "understands" the language and culture of a particular region or country. This is especially true if this country is in the Asia/Pacific region, to the degree to which it can support the region's language character set (which requires Unicode support).

Additionally, international data quality solutions should be evaluated for data enrichment capabilities. This is because information from both outside and inside an organization can be critical in terms of filling in the blanks of data records and adding high-value third-party information that can be used.

Global Data Quality Begins at Home

Finally, this paper also pointed out that one of the surest ways of using a strong international data quality solution to limited effect is failing to provide the same high standards of data quality internally. Global data quality begins at home. If an organization operates on data whose quality is suspect, or if there are not sufficient procedures and processes in place internally to ensure the quality of the organization's domestic data, then the effort at leveraging a global data quality solution will be wasted. Data quality begins at home, and no organization that fails to understand this will be able to leverage the benefits of high quality data, either internally, externally or internationally.

As noted at the outset of this paper, the stakes in the global market place are high, and many organizations have stumbled because they weren't prepared to handle both the demands of the international marketplace and its opportunities. An international data quality solution is a requirement that should not be overlooked. However, even the right international data quality solution will provide limited results, and only go so far in supporting an organization in its global efforts, if the organization doesn't first get the rest of its house in order in terms of data quality.