A HARTE-HANKS WHITE PAPER

Managing Global Data

HARTE HANKS®

Insight. Passion. Results.

Managing Global Data

Executive Summary

The world is getting smaller. Modern technology is enabling companies to operate globally in order to take advantage of an expanded customer base and greater breadth of suppliers. This helps drive down costs and increase potential revenues, and ensures that regional boundaries no longer pose limits on a company's reach.

Marketing to all this potential business poses its difficulties, however. And one of the most fundamental challenges is simply managing all of the available data – different sources, character sets, address formats, etc. – to reduce redundancies, control costs and, most importantly, ensure relevant and timely communication with prospects and customers to drive profitable interactions.

This white paper details the array of challenges companies must address in order to manage global data efficiently, the best practices for addressing those challenges, and real-world case studies demonstrating the benefits gained from global data management.

The following topics are included in this white paper:

Multi-national Data – Vive la Différence!

Data are considered to be global when they are collected from multiple countries then combined and managed centrally. The data have values or formats specific to their country of origin and may be in a native language or character set.

When handling multi-national data, hidden problems soon become evident. Of course, languages are a concern, but the problems are more complicated than the several thousand languages in the world. Consider name and address information. There are currently 239 countries, principalities and territories, 130 different address formats and at least 36 personal naming conventions throughout the world. In addition, different countries have different cultural influences that impact how addresses are created and understood.

For all these reasons, companies holding and maintaining global data need to find a way of managing customer information that meets the many challenges but is cost effective and easy to deploy. The challenges are real, but it is possible to overcome them and achieve significant cost and performance benefits with the right tools and expertise.

Japanese

１−７−４５
プランタン朝霞台２０２号
朝霞市浜崎

Taiwanese

台大政治系
台北市紹興南街6號台大女五舍
224室 台北市

Chinese

上海英特华投资咨询有限公司
上海市黄浦区九江路333号1608
室 黄浦区

Global Data Management

For simplicity's sake, many companies choose to store each country or region's data in a separate database. This overcomes much of the confusion caused by global data because the issues are managed locally. Data quality is generally handled regionally as well, using country-specific solutions.

However, regional databases severely limit a company's reach as firms can only do business in those countries where they have a local presence. Knowledge cannot be easily shared and the whole implementation is costly, with a new 'build' required every time a new country or region is added.

Even if a company has regional offices that span the globe, the regional approach forces it to behave like a global organization doing business on a local level. It is impossible to understand customers as a whole. And when customers are other businesses, firms must deal with them on a country by country basis, and struggle to understand their needs on a global level. Global data management enables a more complete customer view.

Centrally managed databases

The power of global data comes from managing all data from one central location. Only then can firms start to understand their customers comprehensively. For example, the opportunities for up-selling are enhanced because judgments about global customers will be more thorough and, therefore, more accurate.

Also, by focusing resources centrally, a higher level of expertise can be developed than that which would be possible across many regional offices. Best practices, developed in more advanced markets, can be shared with developing markets. With today's economy moving focus from established regions to emerging territories, this knowledge sharing can provide competitive advantage when it is most needed.

Additionally, a centrally managed database offers companies the benefits of a single customer view, extended marketing prospects and cost efficiencies.

Of the respondents interviewed for recent market research, 71% agreed there was a significant trend to centralize multi-national contact data.

Single customer view

An essential concept of customer relationship management is the single view of a customer – the idea that all customer information is brought together in one place for analysis and evaluation. When customers are international, especially when they are global businesses, a single customer view is especially important. With the single view, businesses are better able to identify buying trends, as well as customer behaviors across national boundaries, and deliver customer experiences that satisfy needs and grow customer value.

Extended marketing prospects

When an organization is able to manage global data, it is no longer restricted to the countries in which it can have a regional presence. Effectively, the world becomes the company's marketplace. New and developing markets in countries that until recently were out of reach for regional marketing initiatives become available, opening new opportunities for more rapid growth.

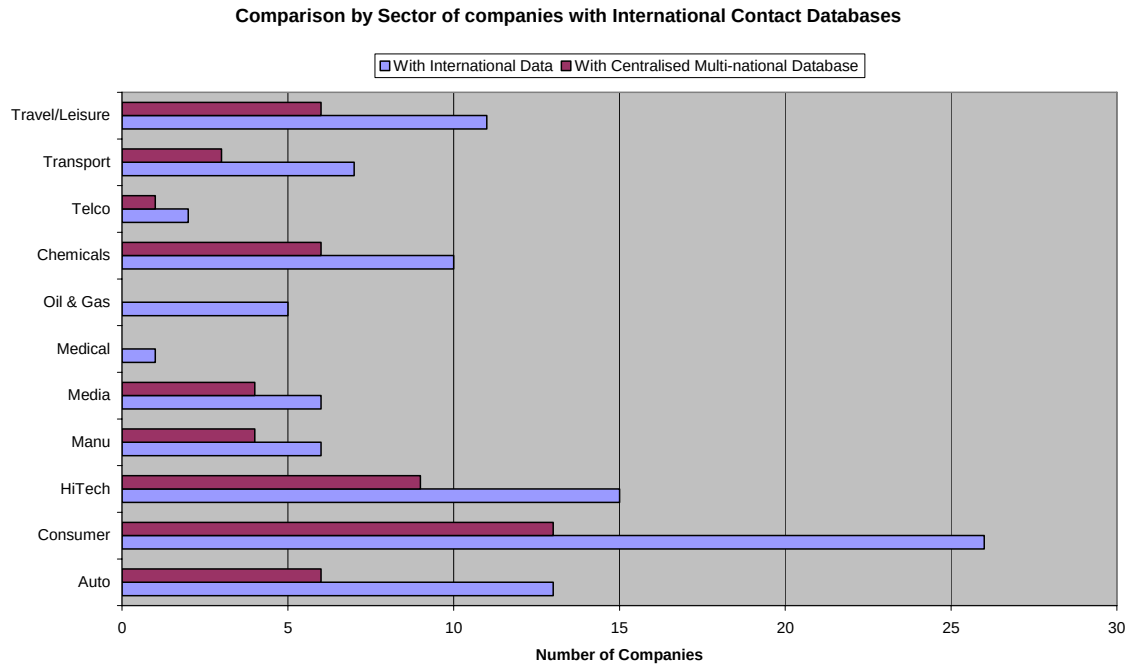
Cost efficiencies

The opportunities for cost reduction are great when data are managed centrally. IT costs can be lowered dramatically by combining local and regional data sets into one central database. Centralizing marketing operations can decrease staffing costs as the need for regional marketing professionals diminishes.

Objectives and challenges

Harte-Hanks commissioned research on the subject of global data management. The research looked at 102 large companies holding international customer data. Approximately half of the databases were global and the remainder regional.

The following graph shows the split by industry and also by the global or regional nature of the companies' data management.

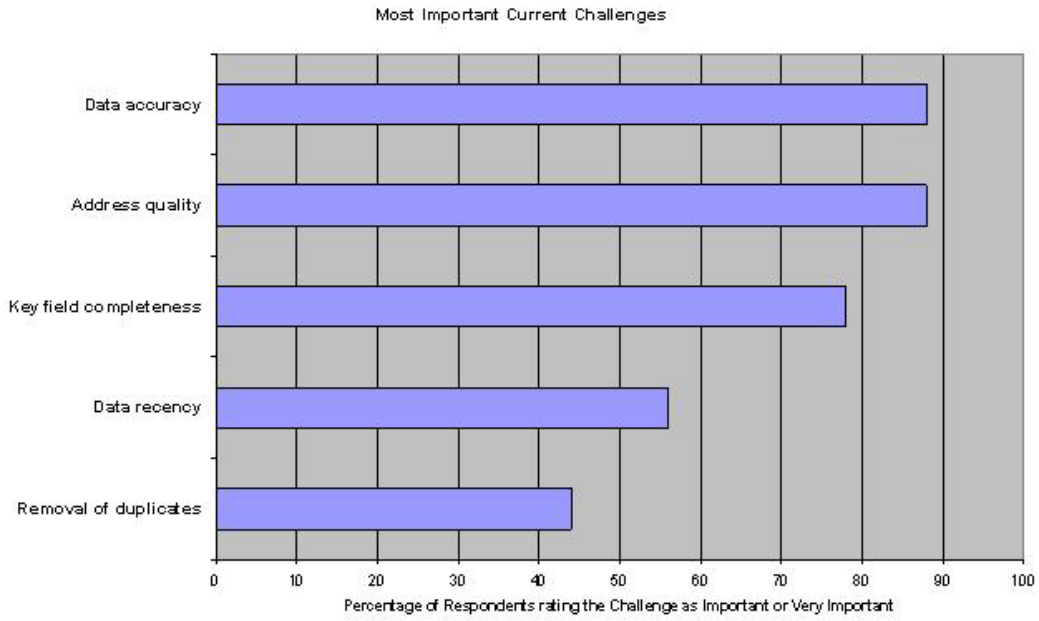


Respondents were asked for their objectives for creating multi-national databases. The primary responses were:

- Economies of scale
- Improved data quality
- More effective use of corporate data

The principal uses of their databases are marketing communications, campaign management and customer analytics.

Not surprisingly, results show that data quality and accuracy are key concerns, as represented in the graph below.



To make the most of centralization, companies need to manage data in a way that is sympathetic to the many different challenges that global data present. Failing to do so will be costly.

Many of the issues experienced by companies managing global data are the same as those with local data. These include:

- Difficulty generating good target lists from the database
- Poor campaign response rates
- Matching and merging multiple data sources
- Lack of data/reports needed to manage the business
- Difficulties in developing a single view of the customer

As noted above, data quality and accuracy are critical concerns, cited by nearly 90% of respondents as important or very important. These concerns are present for virtually all businesses, but companies with multi-country databases have found the challenges to be even greater. This is due to the challenges inherent to global customer data, such as:

- Worldwide there are 239 countries, 5,000 to 10,000 languages, 130 address formats, and 36 personal name formats.

- Almost every language has specific scripts, writing directions and diacritical marks.
- Personal names, company names and job titles differ in terms of order, casing, gender coding and titles, etc. by language/country/region.
- Addresses vary in terms of required content, order, numbering, casing, postboxes and postal coding by language/country/region.
- The availability of address validation reference data changes country by country.
- Other data formats – dates, time, numbers – can be diverse based on language/country/region.

Further, the issues of quality, accuracy and usability are greatly compounded by the complexities intrinsic to multi-national databases. The areas in the research rated the highest in terms of key issues and lessons learned were:

- Cultural issues
- Attitudes toward data ownership
- The importance of consistent data structures

Cultural and country-specific issues

Language barriers are often much easier to overcome than cultural ones. Many countries have strict cultural protocols surrounding names and addresses. Customer satisfaction often relies on the presentation of a customer's personal details.

Matching of names and addresses varies from country to country. In some countries the change of a single letter in a name could be a misspelling, in others, names are so similar that a single letter could have a great impact. Parsing rules will need to reflect this.

Lack of knowledge of local customs and idiosyncrasies can impact the ability to mail to customers effectively. Some addresses look unusual to foreigners and many seem to contain extraneous information. It is often the case that this extraneous information is actually an acceptable part of the address.

Local knowledge and excellent language skills are the only way to overcome these issues.

Attitudes toward data ownership

A major concern of global organizations is a lack of accountability and ownership for data. IT departments often complain that they do not know who to contact when they have questions about data elements within their systems. Business people feel that IT departments take liberties with their data. Changes are made to data without proper controls and without clear communication to concerned parties.

Consistent data structures

When systems go global, companies often make accommodations in the data structures for the sake of simplicity and economy. Rarely are the requirements of the entire world's addresses considered when building the requirements for a new system.

As address formats differ from country to country, it is not unexpected that the types and naming of standard fields required for these addresses varies as well. Some countries have provinces where others have states. Some addresses show the street number after the street name. The positioning of the post code in an address is also variable.

Some countries apply more than one address as standard. This is true for France where there is a mailing address and a physical address for a single company.

Problems also arise when applications using data are constrained by the number of fields available and global data are lumped into inappropriate or ambiguous fields. Direct mailings can have incorrect addresses as a result and campaigns are often plagued with issues around segmentation.

Mailing

Company A
Zone Industrielle
1 Chemin du Pilon Cedex
ST MAURICE DE BEYNOST
01700
France

Physical

Company A
1 Chemin du Pilon
ST MAURICE DE BEYNOST
01703
France

Some additional challenges

In addition to the reported concerns with data structures, data ownership and issues specific to countries or cultures, companies seeking to integrate their global databases face additional challenges when it comes to postcodes, character sets, privacy and achieving a unified, 360° customer view.

Specific fields

Postcodes do not exist for all countries. So, it would be difficult to make 'postcode' a mandatory field in a database that contains global data.

Additionally, field length requirements can vary greatly depending on the language and culture of a country. In Germany, a company may need to have more than 100 characters for a job title alone. This is due to strong cultural protocols that must be upheld to show due respect for customers.

Multiple character sets

The current focus on emerging markets due to the 'credit crunch' is exacerbating the need for dealing with character set issues. But, across the globe, there are numerous countries that use character sets that are not in the 'roman' style. Asian countries like China and Japan are obvious examples, but even countries like Germany and France present issues due to the

presence of enhanced characters like the um-laut (ü) in Germany and letters with the accent aigu (é) in France.

Code pages are tables of mappings that describe a character set for a particular language. There are many different character 'code pages' globally, a fact that doesn't lend itself to global markets. For example, data in Central and Eastern Europe use very different code pages from those of Western Europe. Russia is also emerging as a large marketing opportunity, increasing the demand for companies to manage Cyrillic data as well.

The current focus on emerging markets due to the 'credit crunch' is exacerbating the need for dealing with character set issues.

As these character sets are so different, bringing them all together into a single, centrally-managed database must be handled with technology using software tools that store and manage data using Unicode, which is language independent. Unicode code pages contain mappings for all languages and, as described later, greatly simplify the use of multiple character sets within a single database.

Privacy

Privacy regulations do not exist in some countries and are rigorously upheld in others. The laws around privacy are constantly changing and the number of legal challenges is increasing. Multi-national databases that share information with the US are particularly awkward in this regard. Every company should seek appropriate legal advice on privacy issues. Business rules should be agreed and measures should be in place to enforce this advice.

Central to almost every data privacy regulation in existence is the need for accurate data and accountability for data storage and use. Centralizing data provides greater control and measurability. Locally stored data, in particular data stored on individual laptops in spreadsheets, pose greater risks of privacy breach.

Global single-customer views

The true 360° view of all customer interactions has long been the 'holy grail' of marketers. Achieving this goal on a global scale is even more sought after, and companies who accomplish it have real advantages over their competitors.

However, they must match and consolidate data from many countries, with different formats and languages. As discussed below, they accomplish this through careful parsing of the data, using rules that recognize the differences on a country level.

Tackling Global Challenges with Local Knowledge

Research has shown that the benefits of centralizing multi-national databases can be great, and trends towards these types of databases are on the rise. However, this same research shows that concerns about data quality are growing as well, and companies must also contend with the additional issues described above.

To tackle the challenges of global data management effectively, organizations must apply a set of best practices to the acquisition and use of local knowledge, incorporating technology and processes as needed to ensure an effective, efficient process.

The following steps detail a best practice solution.

1. Perform a data audit

Before setting out to improve the quality of global data, it is important to understand any inherent problems that may exist within the data and the processes that create and update the data. Don't rely on a data dictionary to tell a complete story about the data. Let the data tell the story directly.

Don't rely on a data dictionary to tell a complete story about the data. Let the data tell the story directly.

A data audit should be undertaken to provide valuable insights into data quality, including:

- Uniqueness – Are there unwanted duplicates within the data?
- Completeness – Is all the information needed about customers available? Are key fields completed?
- Validity – Is the data format correct? Does it meet its technical specifications? Are agreed values used? Are mandatory fields enforced? Are appropriate privacy rules in place?
- Accuracy – How closely do the values within the data match reality?
- Referential integrity – Where required, do key fields match values in other, related tables? Are those fields consistently in existence?

Once the audit is completed, a business review should be undertaken to look at the business use and requirements for essential information. This includes understanding the business rules and definitions that should be applied to key data in order to assure that the database meets the needs of the organization.

2. Use Unicode data

The best solution to all types of character set issues is the use of Unicode for storing and processing all global data. Unicode is a method of indexing all types of characters and storing them in a way that is language independent. The benefit to global data is that all data can be

handled as if they were in the same language, thus overcoming the challenges of multiple character sets.

3. Re-engineer the data

The results of the audit and the business rules discussed during the review will highlight areas in need of improvement. Some of the data will need to be re-engineered, particularly fields where field sizes are inadequate. Often data elements need to be moved from one field to another to prepare the data for the next steps in the process.

4. Standardize the data

Good global data management should always include standardization of 'equivalent' pieces of data, particularly abbreviations, using country-based rules. For example, Ltd and Limited are the same and should be standardized to one or the other, using clearly defined rules, to aid matching. Likewise, in Argentina and various other Spanish-speaking countries, Sociedad Anónima = S.A.

5. Parse the data

Parsing means breaking data down into component parts. A parser splits a data record into pre-defined parts so that comparisons can be correctly made with other records, either within the same data set or with external data such as enhancement data. Careful parsing results in highly effective matching because the whole world essentially becomes one country.

Careful parsing results in highly effective matching because the whole world becomes one country.

For example, name and address fields should be broken into pieces such as contact name, business name, premise number, street name, street type, locality name, region name, postal code, and country name.

When working with global data, it is important to parse the data using country-specific rules to address all the different formats and languages involved. Once properly parsed, data that spans many, many countries will take on the appearance of being just one country, and matching becomes much simpler.

6. Validate addresses

Parsed addresses should be compared to postal address files to find a matching entry. Once a match is found, the postal address can be used to append additional elements to an address. Sometimes items of an address can be changed to correct typographical errors and other data capture mistakes. In some cases, even a company name can be improved.

It is important to use reliable reference sources to undertake validation at whatever appropriate level is available for each country.

In Spain …

John Smith
Cra N-152
(Barcelona A Puigcerda), KM 13
Montcada I Reixac
Barcelona 8110
Spain

Means …

This residence is 13 kilometres along the highway

Address validation is variable across the world as there is tremendous variation in postal address quality from country to country. Some countries have very specific formats and regulations, others have virtually none. As noted earlier, postal codes exist in many countries, but not all. Some countries are very detailed in big cities but have few controls over rural areas.

7. Match the data

Standard data management tasks such as de-duplication, data enhancement, and merging & purging require effective matching of name, address and other data. To do this within one centralized database requires exceptional parsing to overcome the differences that exist.

Organizations that parse well also match well, because the data have been given the best possible chance at matching.

8. Apply local knowledge

What looks right to someone in the UK, might look completely wrong to someone living in France.

Technology alone will not provide a complete solution. Companies need to consider cultural differences, local customs, and country-specific addresses. For example, what looks right to someone in the U.K. might looks completely wrong to someone living in France. A best practice solution should combine technology with regional experts in order to get the optimal results.

Local experts should undertake data reviews to identify cultural issues. Parsing should then be tuned to reflect this acquired knowledge.

9. Consolidate for a single customer view

Once your data has been audited, parsed and matched, it is ready to be consolidated together. Business rules will help decide what data to choose from the various sources likely to be available. One source may be given precedence over others, and marketers may want to apply rules regarding currency or quality. In the end, the organization will have a global database with a consolidated, more complete view of each customer, ready for marketing and for reporting globally.

A Note about Global Data Governance

Creating a multi-national database of customer details and then managing that data over time requires a serious amount of involvement from the business. It is not something that should be left to the IT department to handle. Instituting Data Governance as a core competency of the business on a global level is a good way of ensuring the right people are involved and the global data are understood and maintained with the needs of the business in mind.

Organizations should begin by identifying and empowering data owners, who will be responsible for stewardship and change management for all key data elements. This must be done on a global level and there may be the need for ownership forums that include regional data owners to assure quality on a multi-national level.

Global Data Management Case Studies

The two case studies that follow demonstrate the real-world benefits that organizations gain from effective management of global data. In each case, the firms struggled with multiple data sources spanning the globe, incomplete and/or inaccurate data (including multi-byte data), and an inability to match data together – issues that hampered their ability to market products and services effectively to their customers. The application of global data management strategies and technologies helped each to achieve critical data-driven goals.

Global High-Tech Security Company

This firm needed a b-to-b marketing database to help it identify customer segments and improve campaign velocity, first for the APAC region and then globally.

Challenges:

- The volume of data was extensive and challenging
- The ability to segment data and pull lists was lacking
- Several previously attempted solutions were not successful
- A sophisticated tool to manage campaigns was required

Solution:

- Detailed data discovery provided the company with a first real insight into its own data, previously only available in pre-defined views
- A new marketing database incorporated global data management to deliver a single-customer view with a campaign management front-end
- Several years of historic information was incorporated into the database

Benefits:

- Campaigns are faster to implement, easier to control and simpler to track and report
- The company can quickly and effectively expand into new markets, an imperative important in these trying economic times

Global IT Research and Advisory Firm

This firm needed a global marketing database with campaign management capabilities to manage information for their many events.

Challenges:

- Many regional and local databases with separate rules and controls
- No automated response management slowed response time, resulting in lost leads
- Poor data quality and many duplicated contacts reduced efficiency and impacted the customer experience

Solution:

- A global data management solution was used to build the marketing database, which included several years of customer history
- A global single-customer view formed the basis of the campaign management application
- More than thirty different sources of data from all corners of the world were integrated

Benefits:

- The implementation of the marketing database provided much higher quality lists, de-duplicated and marketing ready
- The firm now has the ability to segment data, run and track campaigns, with rapid response management, all on a global basis

Reaping the Rewards of Managing Global Data

By adopting best practices and using the right tools to manage global data, organizations can create a more complete view of their customers around the world, and rapidly capitalize on opportunities to expand into new or under-served markets. The centralized approach, drawing on localized knowledge, further offers greater efficiency through reduced technology and staffing costs, more rapid learning and adoption of best practices across geographical regions, and consequently greater speed to market.

For More Information Contact Harte-Hanks:

In the US:
Phone: (800) 456-9748
E-mail: *contactus@harte-hanks.com*
Web: *www.harte-hanks.com*

Outside the US:
Kathy Hunter, Vice President - Global Data Solutions
Phone (UK): +44 (0)1344 325500
E-mail: *kathy_hunter@harte-hanks.com*
Web: *www.harte-hanks.com*

Usage Notice