

Ten requirements for more cost-efficient and effective business decisions

WHITE PAPER:

**AMERICAS GEOCODING** 



Berk Charlton • Global Portfolio Director, Product Management, Pitney Bowes Software

Michael A. Schiff • Founder and Principal Analyst, MAS Strategies

Ten requirements for more cost-efficient and effective business decisions

### **ABSTRACT**

ORGANIZATIONS MAKE CRITICAL BUSINESS DECISIONS BASED ON LOCATION DATA, DECISIONS THAT IMPACT VIRTUALLY EVERY ASPECT
OF AN ENTERPRISE FROM MARKET ANALYSIS AND RISK ASSESSMENT TO TARGETING, PORTFOLIO MANAGEMENT AND NETWORK
INVESTMENTS. BUT BEFORE YOU CAN ANALYZE, EXTRAPOLATE OR PROFIT FROM LOCATION DATA, YOU NEED TO FIRST ASSOCIATE
EACH RECORD WITH AN ACCURATE LATITUDE AND LONGITUDE COORDINATE—A PROCESS KNOWN AS GEOCODING.

GEOCODING IS COMPLEX, AND MISTAKES IN GEOCODE ASSIGNMENT CAN LEAD TO POOR BUSINESS DECISIONS THAT IMPACT RISK,
RELATIONSHIPS AND PROFITS. FORTUNATELY, EXPERTS HAVE IDENTIFIED TEN FACTORS THAT CAN HELP DETERMINE WHETHER AN
ORGANIZATION'S APPROACH TO GEOCODING MEASURES UP TO TODAY'S BEST PRACTICES ON ISSUES COVERING ACCURACY,
FLEXIBILITY, FUNCTIONALITY AND EXPERIENCE.

WITH TODAY'S TECHNOLOGY, ORGANIZATIONS CAN EVEN ACCESS WORLD-CLASS GEOCODING SOLUTIONS ON DEMAND THROUGH CLOUD-BASED SOLUTIONS THAT OFFER 'PAY AS YOU GO' ACCESS WITH NO OVERHEAD EXPENSE. AS A RESULT, COMPANIES AND GOVERNMENT AGENCIES HAVE FOUND THAT IT'S NOW EASIER AND MORE COST-EFFECTIVE TO GENERATE MORE PRECISE INFORMATION—INCREASING CONFIDENCE IN YOUR DATA, AND YOUR DECISIONS.

ADDRESSES HELP THE UNITED STATES POSTAL SERVICE DELIVER MAIL. THEY DO NOT, HOWEVER, INDICATE IMPORTANT GEOGRAPHIC-RELATED ATTRIBUTES AND CHARACTERISTICS THAT CAN IMPACT RISK, CUSTOMER SATISFACTION AND PROFITS.

### **Overview**

Every business entity and government agency deals with address information.

A prospective customer, for example, may request service at 720 Primrose Lane in Chesapeake, VA. But what, exactly, can you tell from an address? Can you provide service at this location? If so, does your prospective customer qualify for a specific offer or special rate?

Addresses help the United States Postal Service deliver mail. They do not, however, indicate important geographic-related attributes and characteristics that can impact risk, customer satisfaction and profits. For example, organizations need an accurate, automatic way to:

- · Assign customers to territories for sales, service or pricing
- Determine the distance from an office or service junction
- Identify the Census Tract & parcel number for regulatory compliance
- Calculate the proximity to a coastline, hurricane path or nearest fire station



Today, geographic-based data drives social policy, network planning and market analysis. In business, these insights are essential for risk management, regulatory compliance, pricing and strategic planning. That's why so many organizations employ geocoding. Geocodes translate common reference points, such as customer addresses, into latitude and longitude coordinates that makes it easier to analyze data.

Although some people think of geocoding only in the context of maps, geocoding is also the enabling technology for spatial analysis—and makes it possible for organizations to determine the relationship between two or more locations.

With accurate latitude and longitude coordinates, organizations can calculate the distance between two points, the distance to a specific boundary and determine if an address is situated within a zone or territory. Unfortunately, most databases do not contain geocodes as part of the associated address data. When they do, there may be a high degree of uncertainty in the quality of those geocodes, which can lead to faulty business decisions and costly mistakes. On aggregate, these mistakes can lead to lawsuits, attrition and significant losses.

Ten requirements for more cost-efficient and effective business decisions

### Components of Business-strength Geocoding

Geocoding is very complex. There are several free or low-cost consumer-oriented applications that offer geocoding solutions—but these tools were designed to support general consumer needs.

Business-strength geocoding, on the other hand, is meant for problem solving. When organizations take a close look at the risks associated with a poor or inconsistent approach to geocoding, they soon find that accuracy, flexibility, functionality and experience can make all the difference.

Most business-oriented geocoding solutions consist of three primary components: address data standardization and validation, geocode determination and data enrichment.

- Just as the USPS cannot deliver a misaddressed envelope to
  its rightful destination, valid, well-defined addresses are
  important components of geocode accuracy. The USPS has
  established addressing conventions and guidelines
  governing punctuation and abbreviations, format, and
  address components for address standardization. The first
  step in augmenting an address with a geocode is to
  standardize and validate that address.
- When it comes to geocode determination, there will be times when it is not possible to deliver a geocode centered on a specific address. Business-strength tools will recognize this and apply consistent rules, automatically cascading to the next most-specific point of reference. In most cases, no geocode is better than a wrong geocode.
- Lastly, you will need an ability to append additional data sets, including demographics, purchasing preferences and lifestyle data. Plus, an ability to append spatial data derived from your analysis – such as territory assignment or flood zone determination.

The best business-strength tools will integrate all three of these components into a single solution, providing the most accurate results possible. In many cases, organizations can take advantage of on demand and cloud-based solutions that provide "pay-as-you-go" access with no overhead expenses.

### **Challenges with Geocoding**

Short of placing a global positioning system (GPS) at an address to determine its true longitude and latitude, there are several approaches to geocoding. These range from less precise methods based on ZIP Codes all the way to parcel centroid level geocoding. Organizations interested in doing more with location-based data need to understand the challenges associated with geocoding.

### **ZIP Code Limitations**

ZIP Codes are not necessarily stable over time. While a location's latitude and longitude will never change, the USPS will often modify a location's ZIP Code. Approximately 600,000 of the 40 million ZIP+4 Codes in the United States change each month. On top of that, many homes and businesses have been assigned addresses (e.g. for emergency 911 purposes) that are not recognized by the USPS for mail delivery.

Among the potential problems associated with ZIP Code geocoding, is that it may span county or other municipal boundaries. Some areas, such as Carmel, California, only have mail delivery to post office boxes, not to homes. At the other extreme are large office buildings in major metropolitan areas. These buildings may have multiple ZIP+4 codes, perhaps one for each floor. Additionally, centralized mail delivery locations, which are used in many townhouse complexes, may have their own ZIP+4 Codes and could be several hundred yards from where the residents of the townhouse physically live.

### Varying Degrees of Accuracy

Not all geocodes are the same. Parcel Centroid geocoding places the coordinates at the center of the parcel associated with the physical street address (these are often referred to as a point-level or rooftop geocodes.) Street Segment Address geocoding is based on knowing the geocodes at nearby intersections and then estimating the geocode of that address by interpolation. For example, if a block runs from 101 to 201 Main Street, then 149 Main Street would be approximately halfway between the two end points. Some geocoding engines go further and improve upon standard street level interpolation by inserting parcel centroid data into the equation—resulting in a more accurate interpolation.

# GEOCODING TOOLS SHOULD BE SOPHISTICATED ENOUGH TO ANALYZE MULTIPLE DATA SETS AND THEN COMPARE AN ADDRESS AGAINST ALL POTENTIAL GEOCODE CANDIDATES IN A SINGLE PASS.

Since ZIP Codes were designed to quickly sort and deliver mail, they are not nearly as accurate as Street Segment Address-level or Parcel Centroid-level geocodes. ZIP Code based geocodes assign the latitude and longitude associated with the midpoint of a ZIP Code area, and only approximate the location of any particular address. Geocodes based on fivedigit ZIP Codes, for example, will span an area much greater that geocodes derived from nine-digit (ZIP+4) codes—and users of this information often have no way of knowing how the geocode was determined.

### **Multiple Data Sets**

Geocoding engines that rely on a single data source cannot deliver accurate geocodes. In addition to the USPS data, one or more spatial data sets are needed to determine latitude and longitude coordinates as the spatial data sets also contain address information for non-USPS recognized addresses. Several data vendors provide the required spatial street network data (e.g. Tele Atlas, Navteq, and TIGER) and point data (e.g. Centrus Points, Tele Atlas Points). Geocoding tools must be able to parse, standardize and match the input address against all address dictionaries, score all potential match candidates, and only then determine the best match.

### Sequential Processing

Some geocoders do not integrate USPS data with third-party spatial data sets. They simply apply different data sources and address dictionaries sequentially, and stop when a match is found. This type of sequential processing leads to errors, as the tool may identify a match through one data source while a more accurate match is available elsewhere. This is known as a "false positive". Geocoding tools should be sophisticated enough to analyze multiple data sets as a single address dictionary and then compare an address against all potential geocode candidates in a single pass.

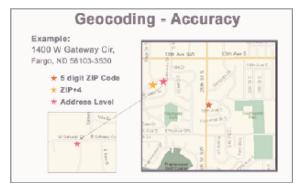
### **Outdated Data**

Given the pace at which ZIP Codes change, new street names are assigned and new addresses created, the data driving any geocoding solutions needs to be updated frequently. It's imperative that your geocoding solution can support monthly updates to provide the highest number of matches and the most accurate results.

The number of physical addresses continues to grow as a result of new developments and the conversion from Rural Route and PO Boxes to house number style addresses. However, due to the development schedules and priorities of the data providers, spatial street vector and address range coverage may be incomplete. In cases where multiple matches cannot be resolved (for example 50 Main St could be 50 S Main St or 50 N Main St) most geocoding software reverts to ZIP Code geocoding when address-level geocoding is not possible. Solutions that rely on more frequent data updates can determine and append the most current 5 or 9 digit ZIP Code and provide a more precise geocode.

While these challenges exist, however, there are easy, effective ways to avoid potential problems. Given the value that an accurate geocode provides, organizations would be well served to search for geocoding solutions that can overcome these obstacles.





Ten requirements for more cost-efficient and effective business decisions

### **High-value Business Applications**

Geocoding solutions turn addresses into latitude and longitude coordinates, making it easy for organizations to understand the relationship between two or more points.

With an accurate geocode, organizations can gain insights, perform calculations and automate processes. These relationships—uncovered through spatial analysis—are generally classified into three categories: distance between two locations (e.g. nearest store), distance to a boundary (e.g. coastline), and determining if an address is situated within a zone (e.g. tax district). The last category is also referred to as a "point-in-polygon" application. The polygon is defined by the latitude and longitude coordinates of a series of line segments that form the polygon.



The specific value that geocoding provides an organization varies from industry to industry. For example:

### All Industries

Geocoding helps analyze data in the context of location so organizations can make smarter decisions regarding targeting, networks and risk.

- Marketing
  - targeting, segmentation and modeling
- Field service routing
  - installation, maintenance & repair services
- · Sales territory assignment
- · Tax jurisdiction assignment

### **P&C Insurance**

Risk varies greatly based on location, and even a few hundred feet can mean the difference between a profitable policy and a costly claim.

- · Risk management
  - proximity to landslides, earthquakes fault lines, volcanoes and other hazards
  - likelihood of terrorism, wildfires, wind storms, hail, hurricanes and tornados
  - located in or near flood plains or areas prone to coastal flooding
  - distance to nearest fire station and fire hydrant
  - overall risk accumulation/aggregation
- Territory and rating determination
  - Pricing
  - Driving distances
  - Straight-through-underwriting
  - Wind pool credits determination

### Telecommunications/Utilities

With networks that cross traditional geographic boundaries, location-based insights add value to nearly every functional group.

- Service eligibility
  - Accurate location of reported bad coverage
  - Service availability/coverage locator
- · Network management
  - Asset management
  - Network design and maintenance
  - Call before you dig

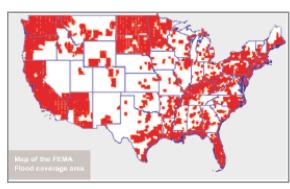


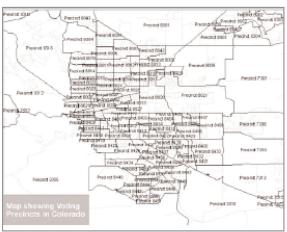
# WITH AN ACCURATE GEOCODE, ORGANIZATIONS CAN GAIN INSIGHTS, PERFORM CALCULATIONS AND AUTOMATE PROCESSES.

### Municipalities & Public Sector

Geocoding and mapping tools have applications across government agencies.

- Constituent services
  - Eligibility for services and programs
  - Voting precincts and polling locations
  - Nearest office, government resource
  - Crime mapping
- Tracking trend data to anticipate needs
- Economic development and urban planning





### **Financial Services**

Geo-based market segmentation plays a critical role as proximity to branch and competitive branch locations can affect both sales and profits.

- Branch expansion
- · Performance goals by branch and product
- · Sales potential of new products for each market and channel
- · Fraud detection and risk scoring
- CRA and HMDA analysis

#### Healthcare

Healthcare must be accessible, so companies need to demonstrate that they can adequately cover members and provide quality care.

- Mapping provider networks against member addresses
- · Gap analysis
  - availability of specialists
- Fraud detection
- Health monitoring—including epidemics—across geographies

### **Shipping Logistics**

Geocoding provides a clearer, smarter picture of just how to deliver goods to customers, with an ability to append weather, traffic and construction updates and proximity to important locations such as distribution centers or stores.

- Determining the true territory or zone an address is in
- Plotting how addresses are located relative to each other
- Determining ways to optimize all aspects of delivery from point of origin, to loading of goods, to routing deliveries, to tracking of items in transit

Ten requirements for more cost-efficient and effective business decisions

### Retail/Real Estate

Location impacts most every strategic decision facing organizations in the retail and real estate industries.

- · Site selection
- · Market expansion
- · Managing store performance
- · Trade area determination

# The ten 'must-haves' for optimal geocode results.

Given the wide-spread impact that geocoding can have on business decisions, organizations need to verify that their geocode technologies measure up to industry best practices. Ten factors in particular can determine how cost-efficient, effective and consistent your solution performs.

### 1. Integrated Address Quality

Accurate addresses are a prerequisite for accurate geocoding. Leading-edge solutions will offer the ability to cleanse data, standardize addresses and validate that source addresses are correct before applying geocodes.

While some geocoding vendors have partnerships with third-party name and address data quality vendors, it is more effective when these capabilities are integrated in one platform. This way you can standardize and correct address information in a single pass using multiple parsing and matching algorithms— giving you the ability to potentially resolve and standardize addresses a stand-alone algorithm might have rejected.

Having to conduct two or more separate passes (against multiple data sets) will also slow down data processing and potentially introduce erroneous matches and geocodes.

### 2. USPS-Certification

The USPS has developed rigorous standards and testing through CASS (Coding Accuracy and Support System) to ensure that vendor software can cleanse, standardize, and match common

address entry and formatting errors—and then accurately return the correct address, 9 digit ZIP Code, and other postal attributes. CASS-certified solutions also provide for Delivery Point Validation (DPV $^{\infty}$ ) and Locatable Address Conversion System (LACS), other USPS technologies that help validate the accuracy of address information down to the physical delivery point.

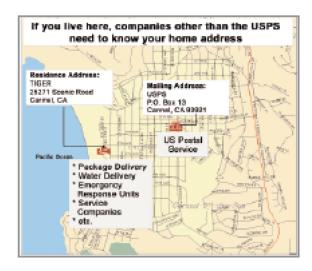
- Previously, address-matching software could only validate that an
  address fell within the low-to high address range for the named
  street. By incorporating DPV, leading geocoding solutions can
  resolve multiple matches and determine if the actual address
  exists. DPV is also useful to identify if a location is vacant and can
  identify commercial mail receiving agencies (CMRAs) and
  private mail boxes.
- LACS converts rural addresses to city-style addressess. For
  example, an address like RR 1 Box 32 might be converted to a
  street-style address like 141 Morrison Ave. LACS also links your
  geocode engine to USPS updates when existing city-style
  addresses are renamed or renumber, or when a new delivery
  address is established.

### 3. Beyond USPS Data Sets

USPS data by itself, however, is not enough. It is vital that your geocoding solution can standardize and validate addresses against data sets that combine the most current USPS and non-USPS address data in a single, standardized address dictionary.

There are millions of households, for example, that receive mail at a Post Office Box; so the Smith Family may receive mail at PO Box 1934 even though their actual physical location is 31 Elm Street. The USPS will have no record of Elm Street in its database, even though it exists in street segment databases, such as Tele Atlas.

By combining information from the USPS with a street segment database, one benefits from a more complete universe of address data that provides more accurate matches. This process is known as conflation. Since any address verification and/or geocoding technology requires "fuzzy" matching – realizing the input addresses are not always perfect – this can lead to "false positives." If 123 Main St is entered and the USPS file only shows a South Main Street,



for example, the software will assume this is correct. By conflating Tele Atlas data with USPS data, the more complete universe will show that there is both a North and South Main St. In this case, no match would occur because the match is ambiguous, which is better than making the wrong match.

### 4. Validated Geocode Results

Analyzing geocode results based on positional accuracy doesn't provide the full insight needed to make critical business decisions. Even when source addresses are fully validated, the geocoding process needs to ensure that the address is located at the right spot. Some solutions return geo-coordinates without providing any details regarding the accuracy level—even when no close match is found.

Geocoding tools should return detailed match codes that indicate the portions of the address that match the source data, as well as detailed location codes that indicate the accuracy of the assigned geocode. These codes can then be incorporated into rules-based processes and automated decision making. For example, a commercial insurance company may have a rule that says when the value of an insured risk such as a property exceeds \$10M, there must be an exact match to the street name (including directionals) and the geocode must be accurate at the street level or better—otherwise it will be kicked out for exception processing.

### 5. Multiple Levels of Accuracy

There will be times when it is not possible to deliver a geocode centered on a specific address or parcel. Based on the available data, the solution should be able to return latitude and longitude coordinates at successive degrees of precision, including parcel centroids, parcel point interpolation, street address interpolation, street centroid, ZIP+4 centroid, ZIP+2 centroid and ZIP Code centroid as well as geographic centroids such as city and county. The tools you use should recognize this and apply consistent rules, automatically cascading to the next most-specific point of reference. This approach ensures the best-available information.

### 6. Data Flexibility

One should expect the flexibility to handle different types of databases and input addresses, including structured/unstructured, residence/business, etc. Most importantly, organizations need the flexibility to utilize and switch among multiple data sources. In addition to the conflated address dictionaries highlighted above, organizations may also have proprietary data sources. For example, a company may own street data or point of interest data in one town, or, may have access to data from a local municipality. These custom User Dictionaries should work in tandem with USPS and street segment data sources as the goal should always be to provide for the best decision possible.

### 7. Data Enhancements

The ability to support data enrichment is also important. Addresses and geocodes can be the keys to files containing additional data that could be added to customer records, including census bureau demographics, income levels, purchasing preferences and lifestyle data.

Several third-party sources provide demographic and lifestyle data that can be appended to individual consumer or household records. The lowest level about which the U.S Census Bureau publishes demographics is the Census Block level. There are over eight million Census Blocks, which typically encompass a city block in urban areas, but could be several square miles in rural regions. Census Block data includes number of residents, number of housing units, median age and median age by sex, number of residents in various age brackets, number of households, average household

9

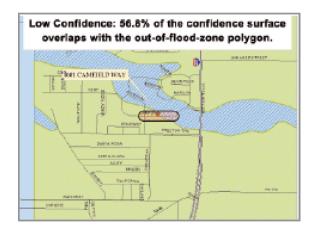
Ten requirements for more cost-efficient and effective business decisions

size, number of households with children, median household income, family income, along with several other fields. Marketing organizations can incorporate this data into their spatial analysis in order to determine where to open new stores, conduct targeted mailing campaigns, or advertise products and services.

#### 8. Geoconfidence

As each geocoding technique is subject to varying degrees of precision, best-in-class geocoding solutions not only return the latitude and longitude, they also measure the degree of confidence in the results and the maximum possible error.

When using street segments, for example, the maximum possible error would be the physical length of the street segment on which the address is located. This might be 264 feet in New York City to over a mile in some rural areas. A desirable feature of any geocoding product is the ability to generate a "geoconfidence surface" or buffer area that can be used to define, and perhaps with the help of mapping software display, the smallest possible region in which the address must exist. These buffers represent more than a simple circle with the most probable location at its center, rather they are custom polygons based on the level of accuracy that can be derived from the data. When used in conjunction with point-in-polygon analysis, organizations can assess the probability of whether an address falls in one area or another. If a buffer area spanned multiple floodplain boundaries, for example, the tool should indicate the percentage lying within each floodplain and outside of it.



Some geocoding systems have the ability to use confidence surfaces to indicate the maximum and minimum possible distances between two locations—such as distance to a central office, or the closest and furthest possible distance to a coastline.

Products that provide these capabilities have a distinct competitive advantage over those that do not. Suppose for example that you have a low-quality geocode (e.g. one based on a five-digit ZIP Code alone). If you calculated that entire 'geoconfidence area' fell within your target zone, you could still be 100% confident in the end results.

### 9. Multiple Deployment Options

While some business situations can be satisfied by assigning geocodes to a group of addresses in a batch environment, other situations require this determination be made to an individual address in real-time. A product that can be deployed in both batch and real-time modes serves a broader range of business applications—including future needs. Deployment flexibility is further enhanced if the software is available in both thick client/ server and thin-client Web-based versions. On-demand applications via Web Services Application programming interfaces (APIs) should be available so that the geocode software can be called from other software applications.

With SaaS and OnDemand solutions, typical integration efforts may take hours instead of days due to the availability of multiple APIs including web services. With no investment in servers, no ongoing maintenance and no software updates to install, organizations can save time and money month after month. Benefits include:

- Lower total cost of ownership
- · Quicker time to value
- Flexibility to rapidly scale based on market opportunities
- Automatic access to the most current—and most accurate—information
- More efficient collaboration

### 10. Experienced, One-stop Service

While simple in concept, geocoding applications can have subtleties that are not obvious to those implementing them for the first time. Vendor expertise, including a track record of successful implementations across a variety of business cases, can prove invaluable both during the planning and implementation stages. Experienced vendors can often add value by suggesting the appropriate data augmentation databases and providing suggestions as to how other clients with similar issues have utilized geocoding technology. A vendor's financial stability should also be examined in determining future support and product stability and reliability.

Overall, solutions need to be simple to use and flexible enough to meet different business requirements. A single technology platform that matches up with your overall corporate objectives can help ensure that a consistent standard will be applied in every market. Likewise, maintaining one platform reduces cost of ownership and can speed up system integration. A single interface also simplifies training and education, and makes it easier for your company to gain the skills and capabilities needed to achieve a competitive advantage.

Even if two vendors are using the same underlying geocoding database, match rates will vary due to differences in parsing and matching techniques. Conducting a benchmark designed to reflect your business requirements will provide valuable information in helping you select one technology over another.

### Conclusion

More than 70% of all business records include a location component, so it is not surprising that location-based analysis is in such high demand nowadays. While a physical address is of paramount importance for postal delivery purposes, it does not tell you where that address is located relative to other locations, such as a sales territory, flood plain or specific tax district—or distance to the nearest office.

Spatial analysis serves as a basis for business decisions in marketing, operations and a wide variety of industries such as insurance, telecommunications and the public sector. But before you can

analyze, extrapolate or profit from location data, you need to associate each record with an accurate latitude and longitude coordinate—a process known as geocoding.

Geocoding can be complex, and mistakes in geocode assignment can lead to poor business decisions that impact risk, relationships and profits. Organizations need to consider several factors when choosing a geocoding solution, including accuracy, flexibility, functionality and experience. Today's leading enterprise geocoding tools combine address quality, multi-level geocoding and spatial analysis in a single solution that can apply the same standardized rules in every market—so you can act with confidence.

### ABOUT PITNEY BOWES SOFTWARE

Pitney Bowes Software offers a unique combination of location and communication intelligence software, data and services that can be used throughout an organization. We combine our deep industry knowledge with our strategic analysis offerings and apply our expertise to help you take action that leads to better, more insightful decisions. You will get a more accurate view of your customers, and integrate that intelligence into your daily business operations to increase revenue, improve profitability and enhance operational efficiency. Call 800-327-8627 or visit www.pb.com/software for more information.

### ABOUT MAS STRATEGIES

MAS Strategies specializes in helping vendors market and position their business intelligence and data warehousing products in today's highly competitive market. Typical engagements include SWOT analysis, market research, due diligence support, white papers, public presentations, and helping organizations evaluate tactical and strategic product and marketing decisions. MAS Strategies also assists user organizations in their data warehouse procurement evaluations, needs analysis, and project implementations. For more information about MAS Strategies, visit its Web site at www.mas-strategies.com.



### Every connection is a new opportunity™

### **UNITED STATES**

One Global View Troy, NY 12180

1.800.327.8627

pbsoftware.sales@pb.com

#### CANADA

26 Wellington Street East Suite 500 Toronto, ON M5E 1S2

1.800.268.3282

pbsoftware.canada.sales@pb.com

### EUROPE/UNITED KINGDOM

Minton Place Victoria Street Windsor, Berkshire SL4 1EG

+44.800.840.0001

pbsoftware.emea@pb.com

### ASIA PACIFIC/AUSTRALIA

Level 7, 1 Elizabeth Plaza North Sydney NSW 2060

+61.2.9437.6255

pbsoftware.australia@pb.com pbsoftware.singapore@pb.com

