WHITE PAPER

THE INSIDER'S GUIDE TO IP GEOLOCATION DATA AFTER COVID-19

A Closer Look at the Benefits of a Powerful Business Tool in a Changing Business Climate

neustar®

TABLE OF CONTENTS

OVERVIEW

IP geolocation data delivers foundational decisioning insights that power critical business applications in multiple industries, every hour of every day. These insights became even more important with so many people working from home during the pandemic–and experts predict remote work will remain 4 to 5 times higher than pre-pandemic levels. Understanding geolocation data and what it can help you accomplish is essential to the success of your online business—both today and in the future.

About This White Paper

This white paper presents the insights and information you need to better understand IP geolocation decisioning data, such as:

- How businesses use it
- Insights into accuracy, and how it can be improved
- The effects of changes in internet technology and usage
- Why new privacy regulations make it more valuable than ever
- Why the choice of provider is particularly critical

IP geolocation data, which maps the IP address of a device connected to the internet to a geographic location, has been helping inform online business decisions for almost 20 years.

Although it is no longer the new kid on the digital block, IP geolocation data has become an increasingly essential tool, addressing more varied business needs and applications than ever before—all of which arise from the growth of anonymous incoming internet traffic.

Yet many fundamental aspects of this decisioning data—how accurate it can be and how provider best practices can influence accuracy, as well as the capabilities it offers beyond location and the challenges posed by changing technology—are not well understood, even by organizations that rely on it.

As a market leader in geolocation decisioning data, our goal is to help our users better understand the value of this data—and the capabilities and challenges they may rarely think about—to ensure they're getting the greatest possible value in their online businesses.

Powering Business Applications

In most instances, IP geolocation data is received by making a request to a geolocation database. This database can be housed on premise at the customer's site or hosted by the geolocation data provider. This call returns data insights about the IP address, including details about where the IP is located, how the IP is connecting to the internet, and insight into the organization behind the IP address.

ONE DATA SOURCE, MANY APPLICATIONS

Created to Meet Business Needs

IP geolocation data was initially developed to help marketers attach locations to unknown website visitors, so they could serve localized content. Quova (acquired by Neustar in 2010), an industry pioneer and the forerunner of Neustar UltraGeoPoint, was among the first to provide this data.

The applications for geolocation data multiplied as innovators created new web-based businesses with geographic restrictions—on-demand content such as sporting events, for example, or gambling and gaming. These new applications also brought more serious penalties for breaches of content licensing or government regulations—and a greater urgency for getting location right.

Providers like Quova responded by expanding coverage and introducing new datapoints and capabilities, such as the ability to identify anonymizing proxies.

In short, committed providers have always worked in partnership with data-driven online business to ensure that geolocation data addresses evolving market needs.

If you use IP geolocation data, you already know how it supports your specific business application. But you may not know what else it can support in your enterprise.

With an estimated 25% of the US labor force continuing to work from home throughout 2021 as the pandemic recedes, today's connected businesses have a huge stake in knowing as much as possible about the devices and people logging on to their online resources.

IP geolocation data provides a critical element in that needed intelligence, delivering the geographic location of the IP address, information about the characteristics of the network connection, and, from some providers, insights about anonymizing connections.

This information is crucial to many different critical business functions, such as:

- **Media and content distribution,** by ensuring eligibility for access as the industry experiences powerful sustained growth, with an expected CAGR of 23.2% through 2025, as new players continue to launch in the aftermath of the surge of stay-at-home viewers
- **Customer experience,** by enabling geographically targeted, localized content and reducing friction throughout the customer journey
- **Cyber security and threat intelligence,** by helping to identify potential threats to the corporate infrastructure and preventing access by unauthorized users
- **Fraud prevention,** by helping to identify potentially fraudulent transactions – now even more important as fraudsters who had been sheltering in place found new ways to conduct nefarious activities
- **Legal and regulatory compliance,** by restricting or blocking unauthorized access to websites or content that could result in fines or penalties, while providing authorized users easy access
- **Gaming and gambling compliance,** by preventing access from users in restricted locations and ensuring only approved players have access to gaming platforms and events

With such a rich suite of applications, it's no wonder that IP geolocation data is in constant daily use in industries from financial services to retail and ecommerce, from OTT and streaming media to government.

THE INSIDE STORY ON ACCURACY

Where Do IP Addresses Come From?

IP address allocation begins with the Internet Assigned Numbers Authority (IANA) and the five regional internet registries (RIRs) responsible for managing the global space.

The RIRs allocate their IP address blocks to ISPs, universities, governments, and organizations that are large enough to administer their own IP networks, which in turn subdivide them into smaller blocks to be dynamically assigned for use on active networks.

How accurate is IP geolocation data? The short answer is: There is no short answer.

The real answer is: it's hard to say. Accuracy depends on a number of different factors. Not a great marketing slogan, but that's the simple truth—and any other IP geolocation provider that suggests otherwise, without disclosing that they are using opt-in or invasive location collection methods, is exaggerating. (Another simple truth.)

Why it's hard to say, #1: The internet itself. Below the country level, the internet is not just non-geographic; it's almost anti-geographic. IP block allocations belonging to carriers aren't necessarily aligned to established geographic boundaries, and carriers can reallocate their assigned IP addresses at any time. Providers such as Neustar work hard to ensure they remain on top of these kinds of changes, through direct relationships with carriers, for example.

Nevertheless, IP addresses are in a constant, fluid state of change, at times with no geographic rhyme or reason to any of it. And there is no absolute geographic truth set for IP addresses—a strong IP geolocation database is as close as you can get.

Why it's hard to say, #2: Connections. In addition to the limitations that arise from the way IP addresses are dynamically allocated, the type of connection between a device and the internet can also have a major influence on the accuracy of a geographic determination.

Many common types of connections, including mobile gateways and regional or international proxies, obscure the true location of the end user as they connect thousands of clients from disparate locations through a relatively small allocation of IP addresses.

Similarly, internet connections proxied through a hosting facility will generally yield the location of the data center, not the location of the actual user.

Just imagine the number of users connecting through Starbucks locations at any one time!

Why it's hard to say, #3: Non-invasive data collection. IP geolocation data is not an opt-in service. It does not involve the use of technologies such as GPS in mobile devices to pinpoint a user's location within a few meters.

In fact, geolocation data doesn't require users to opt in to anything. That's a disadvantage for precision—but a significant advantage in an era of growing skepticism about invasive opt-ins, as well as greater concerns and regulations regarding data privacy.

Bottom line: How accurate is it? Neustar UltraGeoPoint provides location data that's accurate to the city or postal code level. It's accurate enough to provide decisioning data to drive:

- Geofencing of content delivered (or accessed) virtually anywhere in the world
- Legal and regulatory compliance for restricted content
- Compliance for gaming and gambling
- Improved customer experiences through localization and friction reduction

It cannot locate a user to a specific business or home. Period.

What about those confidence codes? Some providers, including Neustar, include a **confidence factor** as part of their IP geolocation dataset. In our case, it's derived from a number of data points associated with the IP address and the connection, including routing type, ASN, hand-mapped input, and other factors.

Higher confidence factors indicate a higher number of inputs that agree upon the current assigned location.

Neustar UltraGeoPoint provides location data that's accurate to the city or postal code level.

WHY YOUR PROVIDER MATTERS

HOW PROVIDERS AFFECT ACCURACY

One Practice to Avoid

When an IP address can't be resolved to a specific location, it shouldn't be arbitrarily assigned to one - say a farmhouse in Kansas that happens to be near the geographic center of the 48 contiguous United States.

When one provider had 600 million IP addresses that could only be resolved to the US, it identified that home as the location for all of them. This resulted in countless visits from police and sheriff's officers, FBI agents, federal marshals, IRS collectors, ambulances, and even angry internet users.

Ask your provider how it handles that situation!

Many data feeds are basically cut and dried. Black and white. Either/or. Not so with IP geolocation data. Its accuracy—and therefore its value—is materially affected by the policies and practices of your provider.

Because there is no definitive, absolute truth set for IP geolocation data—and because the data is affected by multiple factors (see page 5)—data from different providers can differ in its accuracy and value.

Every IP geolocation database is in a constant state of change, with some degree of uncertainty as an inherent characteristic. Your provider's policies and practices regarding that data are crucial in maximizing its health and value and minimizing its uncertainty.

Specifically, does the provider amplify its data sources to gain additional insights and data points, and does it have established procedures to refine and improve the accuracy of its data?

Amplifying data with third-party sources. Providers with a commitment to greater accuracy enrich their geolocation database with relevant IP and location data from premium third-party sources.

For example, gaining insights into IP geolocation data from mobile devices can be challenging (see page 11). A provider may therefore purchase data from mobile carriers, providing important additional insights such as:

- Mobile device data
- Wi-Fi hotspot data
- Opt-in data from apps

These data resources improve insights into the location of devices connecting through mobile gateways. Similarly, a provider may have relationships with ISPs to obtain information about new or reallocated IP address blocks.

When incorporated into the data synthesis process, this kind of high-quality supplemental data expands coverage, improves insights into IP address blocks, and ultimately strengthens the accuracy of geolocation placement.

Geo-feedback and updating processes: What does your data provider do when you discover a location is inaccurate? Responsible providers encourage a two-way flow of information with their customers. They not only welcome feedback about inaccuracies; they also have an established procedure to use it to improve data accuracy.

A provider can give its clients a simple tool, such as an online form to report an IP address that has been inaccurately located, then ingest or review the information via a formal process that ensures it's incorporated into the database.

Five to seven percent of residential IPs change on a weekly basis.[1] Strong providers therefore update their geolocation database regularly and often, incorporating client-provided geofeedback along with data from all other sources. The Neustar UltraGeoPoint database is updated weekly at a minimum, with daily ad hoc updates delivered upon request.

Five to seven percent
of residential IPs change on a weekly basis.[1]

1 Based on a Neustar study of residential IP addresses.

WHY YOUR PROVIDER MATTERS

UNMASKING ANONYMOUS CONNECTIONS

In the physical world, smart crooks avoid surveillance cameras to prevent identification. In the online world, they use IP anonymizers to hide their location. Criminal intent isn't the only reason someone might want to remain anonymous online. But the ability to identify connections made through anonymous internet proxies is a critical capability for many users of IP geolocation data.

When geographically regulated services such as online gaming and gambling sites began using IP geolocation data to create geofences that excluded prohibited participants (often US citizens), determined gamers looked for a way around it.

Similarly, when bad guys are trying to defraud a busy ecommerce operation, they need an easy, effective tool to cover their tracks.

Enter the anonymous proxy. And because the fraudsters who hide behind them can bring serious consequences to legitimate online businesses—fines and penalties in the first case, top-line losses in both cases—it's important to be able to identify these anonymous visitors when they log in.

Identifying proxies. Naturally, they don't announce themselves. A provider must have superior network-level information and signals to identify suspected proxies and other anonymizers. The provider then needs to test suspected proxies to validate those that are active, at which point it can add active proxies to its geolocation database, and if necessary, flag them as high risk.

But inactive doesn't mean gone forever—nor, for that matter, will an active proxy necessarily remain that way. Providers must regularly retest both suspected and validated proxies to ensure they are accurately accounted for and identified.

Beyond yes or no. "Proxy" is a broad term that includes everything from a shadowy resource shared on the dark web to commercial VPNs used by legitimate, privacy-conscious consumers. It's in the interest of every online business to be able to differentiate between them, blocking the bad guys while providing access to legitimate users.

Some providers include additional data points regarding the attributes of proxy connections to enable businesses to make distinctions in how they handle anonymized users. UltraGeoPoint, for example, includes:

- **Date last active,** the most recent date the proxy was confirmed to be active
- **Internet proxy type,** specifying the type of network or protocol used by the server to proxy the connection, such as http, Tor, web, or SOCKS
- **Internet proxy level,** describing the degree of concealment provided by the proxy: transparent, anonymous, distorting, or elite, the last of which provides the highest level of obfuscation regarding the user's originating IP address

These additional distinctions help decision engines deliver more nuanced and informed dispositions for anonymous traffic.

WHY YOUR PROVIDER MATTERS

THE CHALLENGES OF CHANGE

Version 6 of the IP address specification (IPv6) is now in widespread use. Mobile devices now account for roughly half of internet traffic worldwide. How providers are handling these significant shifts in internet access is critical to the value of the IP geolocation decisioning data they offer.

The internet is a seething sea in a constant state of change. No news there.

These two particular changes, however, have significant implications for IP geolocation data. It's important for users to understand why and learn how their provider is responding to them.

IPv6 adoption now stands at 35 percent globally.[2] IPv6 location data needs to be part of your decisioning strategy—if not for the here-and-now, then for the very near future.

The new IPv6 address protocol was accepted as a standard more than 20 years ago. But it always seemed to be on the horizon and wasn't initially deployed until 2012. Since then, however, its adoption rate has accelerated exponentially, accounting for 60 percent of addresses in some countries.

The challenge IPv6 presents for geolocation providers is the sheer volume of its address space compared to IPv4.

	IPv4	IPv6
Address length	32 bits	128 bits
Address space	4 billion possible	340 undecillion possible

That's 340 followed by 36 zeroes—an address space 10 billion times larger than its predecessor. It's simply not possible to gather information for such a vast address space on a daily basis.

It's important for users to understand how their provider is responding to these challenges.

2 Percentage of users accessing Google over IPv6, March 13,2021

A Comprehensive Strategy for IPv6

Neustar began testing data collection and location assignment algorithms for IPv6 in 2015, starting with publicly available information. We learned we needed to supplement that data with multiple additional sources, including carriers, thereby enabling our database to include coverage for more than 99 percent of all allocated IPv6 addresses.

We also learned that addresses are reassigned from one device to another less frequently under IPv6. This discovery raised the possibility of creating historical maps of IPv6 addresses to help validate the location of other addresses within the same block.

As a result, IP geolocation data providers must adopt new strategies to track and account for addresses under this protocol. A successful approach will be based on extensive knowledge of address allocation practices—and real-world experience working with IPv6 addresses.

If your provider hasn't already shared its plans for dealing with IPv6, ask it to.

Mobile gateways as internet entry points. There is nothing inherently difficult about geolocating a mobile gateway, the connection between a mobile carrier and its users.

But challenges do arise when the mobile device user is actually mobile. Traffic can be handed off quickly as the user moves through space, and the device is not always assigned the same IP address.

This fundamental issue affects every provider and user of IP geolocation data.

Of course, it is possible to develop location solutions for mobile devices that rely on user opt-ins—providing access to either an app that can share internal GPS data or a device ID to facilitate location tracking. These solutions, however, are well outside the IP geolocation toolset.

What geolocation data providers can and should do, however, is:

- **Identify traffic from mobile gateways** with a high degree of confidence, so that users can decide if they want to treat it differently and respond accordingly.
- **Provide actual location data with an indication of confidence,** so each business can decide how to handle it based on its specific use case(s).
- **Work with businesses** to share best practices in refining decisioning solutions for mobile traffic.

In addition, providers can improve the location accuracy of mobile traffic by augmenting their resources with critical, high-quality data purchased directly from mobile carriers (see page 7).

WHY YOUR PROVIDER MATTERS

PRIVACY AND GEOLOCATION

Setting the Bar High, Right from the Start

Data privacy is not a new concern for Neustar. We have adhered to our principles of Privacy by Design for more than 20 years, incorporating respect for personal privacy into the design, development, and delivery of our data products and services since our founding. That includes UltraGeoPoint.

We take our responsibilities seriously as a steward in the collection and protection of consumer data, to support the delivery of trusted connections between companies and people without sacrificing personal privacy.

The General Data Protection Regulation (GDPR) in the EU, and now the California Consumer Protection Act (CCPA) in the US, have radically changed the regulatory landscape for data privacy. Conscientious providers ensure their IP geolocation data is compliant with both—delivering a powerful tool that is fully respectful of consumer privacy.

The two sweeping privacy acts on either side of the Atlantic are just the beginning. Multiple US states and more than 128 countries are in the process of establishing their own data privacy laws or regulations.

If the growing concern about online privacy among consumers weren't enough to make data-driven businesses reexamine their data practices, the fines, penalties, and other consequences imposed by these current and pending regulations are.

Today, no company with an online presence can afford to be uncertain about privacy.

The GDPR regulates privacy consistently across all industries and businesses operating in the EU, protects personal data, and creates standards for obtaining consent to access, process, or store data. It also introduced the idea of Personal Information (PI).

The CCPA broadens privacy regulations to require businesses and organizations to state explicitly how consumer data is used, shared, and kept secure and to live up to those promises while not acting in a way that would surprise or be unfair to consumers. It also provides consumers with specific rights relating to their personal data. Although technically only in force in California, as a practical matter, it applies to every company doing business in the US.

The shift to PI. GDPR refocused interest from Personally Identifiable Information (PII) to Personal Information (PI)—and that one little letter makes a huge difference.

PI is a much broader concept that is applicable to anything that can be used to identify, describe, or link directly or indirectly to an individual or household. It includes persistent IDs, even if they're pseudonymous, such as cookie IDs, IP addresses, a hashed email and geolocation data—along with all associated attributes.

IP geolocation data has never been intended to locate consumers to a specific home or business or to identify a specific consumer. In part, its value has always been based on the fact that it doesn't require a consumer opt-in to deliver critical decisioning insights to data-driven companies. That's more true now.

But geolocation data can include PI as defined by GDPR—unless your provider ensures that it is stripped from the data you receive.

Make sure your data provider is fully compliant with the laws and regulations that apply to your business. The potential consequences are simply too serious to ignore, as are the benefits of privacy-compliant geolocation insights.

Today, no company with an online presence can afford to be uncertain about privacy.

WHY YOUR PROVIDER MATTERS

THE HUMAN FACTOR

Human input is not typically considered an essential day-to-day contributor to "big data" products. But IP geolocation data is not a typical big data product—and human contributions significantly improve its value.

No process is 100 percent accurate. It's true when mapping any digital data to the real world—particularly when the process involves the inherent ambiguities of translating IP addresses to physical locations. (Even the unambiguous process of digitally mapping the real world required Google to create a Ground Truth Team.)

Resolving these ambiguities—along with the errors reported by users via the geofeedback process described on page 8—can still be most accurately accomplished through informed, experienced human judgment.

That's why providers such as Neustar, which are committed to accuracy and currency, involve a team of highly trained analysts in maintaining their IP geolocation databases.

Our Network Geography Analysts are intimately familiar with regional routing technologies and experienced in the complex topologies of networks. They are involved in refining and improving our geolocation data on a daily basis, which entails:

- **Reviewing location validation requests** and quickly collaborating to review the data and assign resources to make any required changes
- **Analyzing and incorporating geolocation data** into the database, such as address block mappings provided by ISPs
- **Researching router-naming conventions** to identify the cities in which routers are located
- **Creating and refining mapping rules** to associate IP addresses with locations—for example, when a large block of IPs is reallocated to a different state, or when an ongoing issue for a specific carrier in a specific region is uncovered
- **Analyzing and identifying internet routing patterns and historical trends** to build network maps with network "anchor points" in the physical world
- **Identifying IP addresses used as anonymizers or VPNs** by manually digging through networks to find them

Their work contributes directly to improved accuracy and currency in geolocation data.

MAPPING THE FUTURE

Does a 20-year-old data product—challenged by technological change on the one hand and the growing prevalence of GPS-equipped devices on the other—have a future? When it delivers critical intelligence to support instantaneous decisioning in a wide range of critical online business applications, all while respecting stringent privacy regulations, it certainly does.

IP geolocation data is not the exciting new data feed that IT departments are clamoring for.

But in providing essential insights that enable an online enterprise to make informed decisions about how—or whether—to handle each visitor the moment they engage, geolocation data plays a vital role in:

- **Improving customer experiences** with localized content and reduced friction through the customer journey
- **Enabling media and content distribution** in compliance with geographic and territorial restrictions and helping to prevent fraud
- **Protecting gaming and gambling organizations** by preventing illegal access from restricted locations
- **Preventing fraud** by identifying potential fraudsters
- **Supporting cyber security and threat intelligence** by identifying potential threats and preventing access by unauthorized users

No other data currently available—or on the horizon—can provide the same foundational insights to drive robust security solutions, ensure regulatory compliance, and support more positive customer experiences.

As long as those functions remain important, so will IP geolocation data.

LEARN MORE

To learn more about how Neustar UltraGeoPoint can support your business needs, **click here**, email us at **security@team.neustar**, or call us at **1-855-898-0036** in the US and at **+44 1784 448444** in the UK.

ABOUT NEUSTAR

Neustar is an information services and technology company and a leader in identity resolution providing the data and technology that enables trusted connections between companies and people at the moments that matter most. Neustar offers industry-leading solutions in Marketing, Risk, Communications, and Security that responsibly connect data on people, devices and locations, continuously corroborated through billions of transactions. Neustar serves more than 8,000 clients worldwide, including 60 of the Fortune 100. Learn how your company can benefit from the power of trusted connections here: **www.home.neustar.**

WP-SEC-210878-04.28.2021