

Email Verification

Best Practices for Marketers

This document contains Confidential, Proprietary and Trade Secret Information ("Confidential Information") of Informatica and may not be copied, distributed, duplicated, or otherwise reproduced in any manner without the prior written consent of Informatica.

While every attempt has been made to ensure that the information in this document is accurate and complete, some typographical errors or technical inaccuracies may exist. Informatica does not accept responsibility for any kind of loss resulting from the use of information contained in this document. The information contained in this document is subject to change without notice.

The incorporation of the product attributes discussed in these materials into any release or upgrade of any Informatica software product—as well as the timing of any such release or upgrade—is at the sole discretion of Informatica.

Protected by one or more of the following U.S. Patents: 6,032,158; 5,794,246; 6,014,670; 6,339,775; 6,044,374; 6,208,990; 6,208,990; 6,850,947; 6,895,471; or by the following pending U.S. Patents: 09/644,280; 10/966,046; 10/727,700.

This edition published August 2016

Table of Contents

An Introduction to Email Verification and Hygiene	2
Who should care about email verification and hygiene?	3
Where to use email verification and hygiene	3
How do you implement email verification and hygiene?	4
Implementation Best Practices	5
Email Verification and Hygiene Output	5
Recommendations	7
Online Real-Time Implementation	9
Handling Timeouts	9
Prompting for Retry	11
Batch Processing Implementation	11
How often should you verify email in batch?	11
Conclusion	12

An Introduction to Email Verification and Hygiene

Remember the phrase, “Garbage in, garbage out”? The saying has never been more relevant and applicable when it comes to email. It does not matter how compelling your content or offer may be, if you are sending to an email address that will not receive your message, they will not see it. Use email verification and hygiene to cleanse your list and ensure you are sending to only valid and safe emails.

Email verification and hygiene has three key benefits:

- Improve email campaign performance with better deliverability

Email marketing is an attractive channel for marketers. Savvy organizations not only utilize it, but learn how to optimize for greater efficiency. The reality is that the majority of in-house marketers do not think that their campaigns are running optimally due to their technology. Think of all the wasted time, effort and creative brainpower put into your email marketing plans, only to have lackluster performance due to technology challenges. According to Email Marketing Industry Census 2013, 58% of companies cited clean and up-to-date lists have the biggest impact on improving deliverability. Verification is a proven technique that allows you to increase deliverability and sender reputation by getting more of your hard work into inboxes. Ultimately, more delivered messages means additional opens, higher click through rates, increased conversions and more dollars in your pocket.

- Minimize risks associated with bad list hygiene

While email marketing seems inexpensive, bad practices cripple your organization. Save marketing budget by not sending to bad email addresses. Don't waste your valuable marketing dollars on email addresses that are going to bounce back and potentially hurt you. Even worse, you could accidentally send to a trap (perhaps via a typo trap) and be put on an SBL. If you are blacklisted then your email will not only be prevented from going through, but the overall cost of removal of a spam blacklist is quite expensive. It is a lengthy and tedious process that is detrimental to a company's reputation.

- Enhance customer loyalty and satisfaction

Better email marketing means increasing the ability to remarket to your customers and prospects. If you have invalid email addresses then you lose the chance to upsell and cross-sell in the future. According to Return Path's Email Intelligence Report for Q3 2012, roughly 18% of all commercial email in North America never reached the inbox. Verification alleviates the deliverability problem by getting emails to your customers and turning them into brand evangelists.

Who should care about email verification and hygiene?

If your job includes utilizing email to communicate with prospects, customers or partners, you need to consider email verification and hygiene as a significant component of your email strategy. Anyone who sends emails within a company should care about email verification and hygiene, especially:

- Marketing Professionals
- Marketing Technology and Operations
- Ecommerce / Retail Professionals
- Customer Service and Account Management
- Web and Database Management

Where to use email verification and hygiene

Use email verification and hygiene at any point you collect email addresses.

- Webforms and Landing pages
- Ecommerce Websites (checkout page and newsletter sign-up)
- Preference Centers
- Point of Sale (in-store email acquisition and eReceipts)
- List Acquisition
- CRM and Marketing Databases
- Call Center

The Email Verification and Hygiene solution is a powerful, cloud-based technology that includes two separate checks – one that verifies that an email address is valid at both the domain and recipient level, and a hygiene check to alert you to any malicious email addresses. To do this, the Email Verification and Hygiene solution goes through a five-step process to validate email addresses including several quality checks to ensure the right status codes are returned. Domains that have been traditionally hard to verify, such as Yahoo, can also be verified to the recipient level. Informatica's algorithms provide the most accurate verification of these domains. In addition to specific hygiene results, such as spam traps, moles, former traps, seeds, or pattern matches, information about disposable emails, emails containing vulgar phrases and role-based emails can be returned.

How do you implement email verification and hygiene?

As we continue our discussion of cloud-based email verification and hygiene, it should be noted that there are two ways to implement:

- Real-time
- Batch

Real-time verification means emails are verified at point of collection, either on a web form or at other points of acquisition, such as a retailer's point of sales (POS) system or ecommerce shopping cart. This is typically done as customers are entering their email addresses and other information. This implementation involves inserting web services code into your system(s) to prompt the customer in the case of an invalid entry.

Not all who use real-time verification have set up customer-facing implementations. Many will verify emails after collection, but before the data is synchronized with their CRM system or marketing database. Think of this verification process as transactional, as the marketing or purchase transaction happens, email verification and hygiene is occurring. The code for these implementations can usually be added to any of the involved systems or developed with a simple middleware solution.

The biggest benefit to verifying email addresses at point of collection or prior to sending is to increase your list size by prompting for invalid emails at acquisition point. If someone gives you an invalid email address, your chances of getting through to them via the email channel are 0%. Most invalid email addresses are typos - either self-inflicted or entered by a customer service representative typing an email address. If you use real-time email verification to catch errors before users complete your form (or other process), you increase the chance at getting a new opportunity to grow your list.

In some cases, it makes more sense to verify email addresses in a batch process, or bulk format. Use the Informatica API to develop a back-end process that routinely passes email addresses to us on a nightly, weekly, monthly or quarterly basis. Informatica can also receive a file to batch process via email or an SFTP upload.

Depending upon your business, it may make sense to use the batch option in conjunction with real-time verification. For example, verifying emails at point of acquisition and then a batch process of the entire database on a quarterly basis to verify all data to find what email addresses have changed. Approximately 30% of emails change year over year, and you want to be able to clean those from your email system on a regular basis.

Other times, the batch option is a phase one for Informatica customers, who want to begin the verification process immediately but have some limitations (e.g. technical) elsewhere. They begin with batch processing files and add real-time verification at point of data acquisition when possible.

“Marketing campaigns create 60% of all spam trap hits, a considerably higher percentage than any other source.”

- Return Path

Ultimately, you should verify email anywhere and everywhere you collect addresses, so you may use several of these methods:

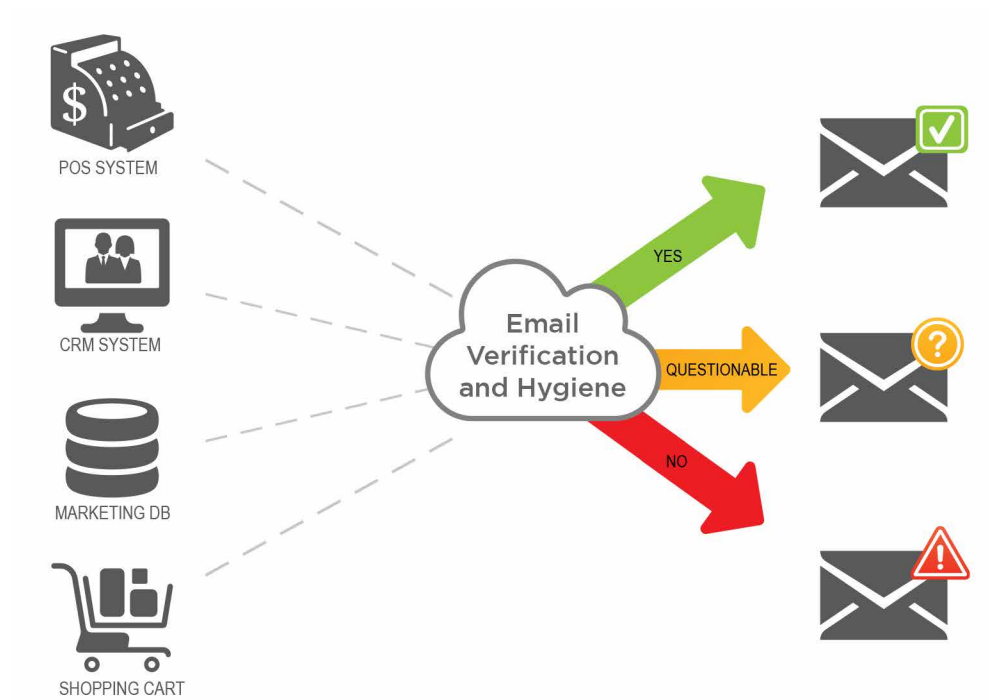


Figure 1.

Implementation Best Practices

Email Verification and Hygiene Output

Once Email Verification and Hygiene has been successfully implemented and email addresses have been run through the service, output data is provided. Our service provides a multitude of data that can be returned - all of it is useful, but the three data points used today are status code, reason code and hygiene result.

Status codes are high-level codes that are provided to give you an overall look at both email verification and hygiene. These status codes are as follows:

200 Email Valid

210 Domain Confirmed

220 Analytics in Progress

250 Email Valid, Potentially Dangerous

260 Domain Confirmed, Potentially Dangerous

270 Analytics in Progress, Potentially Dangerous

300 Email Not Valid

310 Not Verified

Reason codes are specifically related to the email verification portion of the service and speak to the validity of an email address. The reason code provides granularity into what was found when conversing with the mail server. For example, an email address of darlene.nyce123456@gmail.com will provide you a status code of 300 Email Not Valid with a reason code of 307 Mailbox Rejected, which lets you know that this was a valid domain, but the mailbox or user was not valid for this domain. Overall, these reason codes provide insight further than "valid" and "invalid" and can be particularly useful when you see status codes such as 220 Analytics in Progress.

Hygiene results are specifically related to the hygiene portion of the service, informing you if the email address is safe or malicious in nature. There are several types of categories that email addresses may fall into; they are as follows:

Safe to Send

- Safe US - These email addresses are based in the US and are safe.
- Safe International - These email addresses are based internationally and are safe.

Dangerous / Do Not Send

- Spam Trap - These email addresses have been determined to be spam traps or honey pot email addresses. They are directly owned, controlled or monitored by a 3rd party DNSBL site. Emails sent to any of these addresses are considered unsolicited and unsafe.

Be Cautious

- Mole - These email addresses are not directly owned or controlled by a 3rd party DNSBL sites, however, historical evidence establishes a relationship to one or more DNSBL sites. Often times, these tend to be anti-spam supporters collaborating with DNSBL sites by submitting their email address for use or by using anti-spam software on their personal computers. You should be cautious when sending messages to these email addresses.
- Former Trap - Email addresses in this category have come back as safe without any current ownership by a 3rd party DNSBL site. However, they have been previously considered spam traps. For this reason, we keep these in quarantine until they clear a 3 month probation period. You should be cautious when sending messages to these email addresses.
- Seed - Seed emails are third party oversight emails specifically used for monitoring resources, many times for deliverability monitoring. Also included in this category are any litigious addresses, those who seed their email addresses for the sole purpose for litigation and habitual complainers, those who have an extremely high rate of spam complains. You should be cautious when sending messages to these email addresses, especially if you have not employed a third party to monitor your system via a seed list.
- Pattern Match - These email addresses are typically key words, phrases or patterns that may be created by a bot that could be used for trap purposes. Also included in this category are role-based email addresses that do not go to a specific person, such as info, support, sales. You should be cautious when sending messages to these email addresses.

“Sending to even one spam trap can destroy a sender’s reputation.”

- Return Path

Recommendations

For most of your email addresses, you will be provided a definite answer - either a valid, invalid or malicious/dangerous email address. A status code of 200 Email Valid means the email address is definitively valid and safe to send. A status code of 300 Email Not Valid means the email address is definitively invalid. A status code of 310 Not Verified means the email address has been identified as a spam trap. These codes are easy to interpret because valid emails should be emailed and invalid and trap emails should not.

Unfortunately, for a small percentage of email addresses, the path is not as clear. Sometimes it is impossible to definitively determine if email addresses are valid. The question then becomes: what do I do with these email addresses? The answer depends on a number of aspects including:

1. What are your overall goals?
2. How much risk are you willing to accept?
3. What is the likelihood that they are invalid?
4. What other information is available to help make a decision on these email address?
e.g. acquisition method, prior activity

There are five status codes that do not provide a definitive answer:

- 210 Domain Confirmed

These are email addresses where the domain has been verified, but user-level verification has not been successful. Most of these will be reason code 211 Server Will Accept. This is a common response for business domains as they are often set up in a way that allows for a catch-all email address. All user names (both valid and invalid) are returned as valid in this set up. Instead of returning a false positive, we return this as a 211 Server Will Accept. All email addresses in the 210 Domain Confirmed status code are safe to send to from a hygiene perspective. A percentage of these will bounce, but these will not contain any malicious emails addresses.

- 220 Analytics in Progress

These email addresses are truly unknown - both from a domain and a user perspective. In these cases, there is an email server for the domain, but for some reason the service was unable to elicit a definitive response. This happens for a number of reasons, such as grey listing, throttling, temporary server issues, etc. The Email Verification and Hygiene engine will continue to try to resolve these email addresses for 24 hours. If you decide to not mail to these email addresses or include them within your database, remember that you always have the option of retrying them later. All email addresses in the 220 Analytics in Progress status code are safe to send to from a hygiene perspective. A larger percentage of these may bounce, but these will not contain any malicious emails.

- 250 Email Valid, Potentially Dangerous

Email addresses in this category are all valid email addresses. They are in this category because their hygiene check returned a cautious result. A cautious result is a mole, former

trap, seed or pattern match. It is recommended that you email to these with caution. If you acquired these emails organically and show recent activity/confirmed opt-in, then you should feel safe emailing these email addresses. If you do not, proceed with caution.

- 260 Domain Confirmed, Potentially Dangerous

These are emails where the domain has been verified, but user-level verification has not been successful. Most of these will be reason code 211 Server Will Accept. This is a common response for business domains as they are often set up in a way that allows for a catch-all email address. All user names (both valid and invalid) are returned as valid in this set up. Instead of returning a false positive, we return this as a 211 Server Will Accept.

This status is essentially the same as described above for 210 Domain Confirmed, EXCEPT the hygiene check returned a cautious result. A cautious result is a mole, former trap, seed or pattern match. If you acquired these emails organically and show recent activity/confirmed opt-in, then you should feel safe emailing these email addresses, as long as you are aware that a percentage of these emails will bounce. If you do not, proceed with caution.

- 270 Analytics in Progress, Potentially Dangerous









These email addresses are truly unknown - both from a domain and a user perspective. In these cases, there is an email server for the domain, but for some reason the service was unable to elicit a definitive response. This happens for a number of reasons, such as grey listing, throttling, temporary server issues, etc. The Email Verification and Hygiene engine will continue to try to resolve these email addresses for 24 hours. If you decide to not mail to these email addresses or include them within your database, remember that you always have the option of retrying them later.

This status is essentially the same as described above for 220 Analytics in Progress, EXCEPT the hygiene check returned a cautious result. A cautious result is a mole, former trap, seed or pattern match. If you acquired these emails organically and show recent activity / confirmed opt-in, then you should feel safe emailing these email addresses, as long as you are aware that a larger percentage of these may bounce. If you do not, proceed with caution.

There are five status codes that do not provide a definitive answer:

- If more email addresses is the priority, then email to anyone who does not come back definitively invalid or a spam trap. After one bounce, suppress any of those email addresses with unknown statuses. (Email to everything except status code 300 & 310.)
- If risk management is your top priority, then only email to those with definitively valid email addresses. (Email to only status code 200.)
- Many are going to fall into a middle of the road approach.

It is always best to look at a number of factors when making a decision on whether or not to send an email message to an email address. In this middle of the road approach, acquisition method and date of last activity are also recommended to make the best decisions possible.

STATUS NUMBER	STATUS DESCRIPTION	SEND IF ACQUIRED...	AND HAVE ACTIVITY...	SUPPRESS
 200	Email Valid	Send		
 210	Domain Confirmed	Refer to Hygiene code and prior sending data if applicable*		If email bounces 1x
 220	Analytics in Progress	Organically with user-based submission evidence		If email bounces 2x
 250	Email Valid, Potentially Dangerous	Organically with user-based submission evidence	within the last 90 days	Refer to Hygiene code and prior sending data if applicable*
 260	Domain Confirmed, Potentially Dangerous	Organically with user-based submission evidence	within the last 60 days	Refer to Hygiene code and prior sending data if applicable*
 270	Analytics in Progress, Potentially Dangerous	Organically with user-based submission evidence	within the last 60 days	Refer to Hygiene code and prior sending data if applicable*
 300	Email Not Valid	Do not send		
 310	Trap	Do not send		

* Suppress unknowns, invalids, parked domains, moles, former traps, and seeds. Send only to safe and pattern match. If you have any evidence of engagement, transactional history, or purchase behavior from specific email addresses, continue sending to those particular email addresses.

Figure 2.

Online Real-Time Implementation

Handling Timeouts

How you handle timeouts is extremely important when you are implementing email verification in a customer-facing situation (e.g. on a web form, checkout page or POS system). The Email Verification and Hygiene web service accepts a timeout value up to 120 seconds.

Best practices obviously dictate that you do not wait 120 seconds, but how long do you wait? This depends on a number of factors:

- Your customer base
- Where they are interfacing with you (e.g. in person, online)
- What the interaction includes (e.g. multi-step form, multiple page form, single entry field)

It is recommended that you set a timeout of 5-15 seconds. Ultimately, the goal is to set the timeout as high as you feel you can go without sacrificing the user experience. This way you will get the valid data that you need, while not affecting conversion. It's important to remember in a typical situation, you aren't going to hit that timeout value.

The timeout value is the upper limit you are willing to have the customer wait for a response. The average response time for email verification is between .5-1.5 seconds. However, because we are reaching out to each ISP at the time of verification, we are at the mercy of the various ISP servers responding to our query. For instance, Gmail can be a bit slower and if you have a lot of customers with Gmail addresses, you will want to set a timeout to ensure you are verifying those addresses.

If you decide that you can't support a 5-15 second timeout value on your system, you can certainly set it lower. You have two options if you timeout before getting definitive results:

1. The response you will get is a status code of 220 Analytics in Progress and reason code of 221 Timeout Too Short. The engine will continue to try this email address after we return this code. If you make an additional API call (even within 10 seconds), you will most likely get a more definitive code.
2. If your system supports multi-threading or AJAX, you can continue to run the call without sacrificing the user experience. Run email verification and pass a longer timeout to our system (i.e. 60 seconds), include a shorter client-side timeout within your system. If the client-side timeout occurs, keep the thread running as a child to the original process and return the status information of that email to your database. The end-result is a verified email without affecting the next step for the customer. Additionally, it is recommended that each of these calls be separate non-blocking calls, as there are other issues that could potentially hang you up (i.e. network issues elsewhere) and you would not want them to hold up your customer.

While running a multi-thread environment helps increase the number of verifications without affecting the customer, one of the biggest benefits of using real-time verification is to allow users to correct any mistakes before submission. In order to maximize that benefit, it is recommended setting up your form to give the process time to run while your customers finish filling out the form.

For example, instead of placing the "email address" form field at the end of your form, ask the customer for his or her email address first. Use the cursor moving off of your email address field as the trigger for the email verification Web service to run. By the time your customer gets to the end of your form and hits the submit button, you should have a verification response that prompts for a new email address if necessary.

Figure 3 illustrates two different email form placements. The left side, labeled "BAD EMAIL FORM PLACEMENT:", shows a "SIGN UP" button at the top, followed by fields for "FIRST NAME" (Alex), "LAST NAME" (Siminitzki), "ADDRESS" (8273 Monarch Birch Lane), "CITY" (Cambridge), "STATE" (MA), and "ZIP CODE" (02140). The "YOUR EMAIL" field (example@email.com) is at the bottom. A red arrow points to the email field. The right side, labeled "BEST PRACTICE FOR EMAIL FORM PLACEMENT:", shows a "SIGN UP" button at the top with a green checkmark. Below it is the "YOUR EMAIL" field (example@email.com). Below the email field are the "FIRST NAME" (Alex), "LAST NAME" (Siminitzki), "ADDRESS" (8273 Monarch Birch Lane), "CITY" (Cambridge), "STATE" (MA), and "ZIP CODE" (02140) fields. A red arrow points to the email field. Both forms have "Sign up" and "Cancel" buttons at the bottom.

Figure 3.

Prompting for Retry

As mentioned above, one of the biggest benefits to real-time email verification is being able to prompt a user to put in a new email address if the entry is invalid. The real challenge is finding that middle ground where you get the most valid data possible without sacrificing conversion. After all, you want the user to correct any mistakes, but do not want to be annoying enough that he or she gives up altogether.

It is recommended to prompt customers for invalids in the following situations:

- Any time a customer service representative is verbally asking for a customer's email address and typing it in, such as retail POS systems or call centers. It has been shown that 30-40% of these email addresses will be invalid. Validating at time of collection allows a representative to try again and correct any mistakes. Ask for retry once and regardless of the second verification allow the transaction to continue.
- Any time an email is collected for transactional messages, such as e-receipts, order confirmations and shipping notifications. Here you definitely want to prompt for retry at least one time and possibly more. You want to make sure they will receive important transactional messages that they expect, as well as ensure the purchase is from a valid user while not dissuading them from their purchase. If you decide to prompt multiple times, it is recommended that you be more liberal with your response code interpretations, perhaps only prompting multiple times if the email address returns definitively invalid.
- Any time your main objective is to collect email addresses. If you are creating web forms that are specifically designed to collect email addresses, it does no good to have invalid email addresses. Prompt them for retry at least one time, if not more, to collect as many valid email addresses as possible.

While it is typically recommended to prompt for retry, there are some situations that do not warrant it:

- Any time your main objective is to collect something other than email addresses. If you are implementing a web form to collect leads and email address is a secondary field for these contacts, you may want to just validate the email address and post that information to your database.
- Any time you do not want to sacrifice data conversion for the sake of email addresses.

Again, here we recommend you collect email addresses, run real-time verification and pass the results to your database.

The overall recommendation in response to an invalid email address is to ask the customer to correct the email address one time and then continue the customer down your process. Be sure to pass the verification status code into your database, so you will know that it was invalid. You will want to exclude all of these invalid email addresses from your email sends.

Batch Processing Implementation










Batch processing is running a set of email addresses through email verification at one time. This process can be done by writing an API wrapper to manage the process yourself, or you can post / send a file to Informatica to be processed.

How often should you verify email in batch?

This is the question most often asked in regards to batch verification for email address. Because 30% of email addresses change on yearly basis, it is recommended to do monthly or quarterly verifications of your database. Verify less frequently to emails with high engagement levels and more to those that are not opening or responding to your emails. If you have legacy data (i.e. old or inactive contacts that haven't been emailed to in over six months) that you want to try to reengage within your database, we highly recommend running email verification in batch mode before any marketing campaign.

Conclusion

Email verification is a critical process that needs to be part of every email marketer’s arsenal. Its flexibility allows it to be implemented into virtually any imaginable point of collection (e.g. web forms and landing pages, ecommerce shopping carts, preference centers, POS systems, etc.). Whether you opt for real-time, batch or a combination approach, email verification will increase email campaign performance and deliverability, as well as enhances company reputation and customer satisfaction. The guidelines here will ensure you are adhering to email best practices, enabling more emails to reach the inbox. As a result, you will mitigate any risk associated with emailing to potentially “bad” addresses and you will experience additional opens, higher click through rates, increased conversions, and more revenue.

		Email Verification and Hygiene Best Practices		
STATUS NUMBER	STATUS DESCRIPTION	SEND IF ACQUIRED...	AND HAVE ACTIVITY...	SUPPRESS
 200	Email Valid	Send		
 210	Domain Confirmed	Refer to Hygiene code and prior sending data if applicable*		If email bounces 1x
 220	Analytics in Progress	Organically with user-based submission evidence		If email bounces 2x
 250	Email Valid, Potentially Dangerous	Organically with user-based submission evidence	within the last 90 days	Refer to Hygiene code and prior sending data if applicable*
 260	Domain Confirmed, Potentially Dangerous	Organically with user-based submission evidence	within the last 60 days	Refer to Hygiene code and prior sending data if applicable*
 270	Analytics in Progress, Potentially Dangerous	Organically with user-based submission evidence	within the last 60 days	Refer to Hygiene code and prior sending data if applicable*
 300	Email Not Valid	Do not send		
 310	Trap	Do not send		

* Suppress unknowns, invalids, parked domains, moles, former traps, and seeds. Send only to safe and pattern match. If you have any evidence of engagement, transactional history, or purchase behavior from specific email addresses, continue sending to those particular email addresses.

Figure 4. Sign up for an Email Verification and Hygiene Free Trial: https://now.informatica.com/en_daas-free-trial-1

About Informatica

Informatica is a leading independent software provider focused on delivering transformative innovation for the future of all things data.

Organizations around the world rely on Informatica to realize their information potential and drive top business imperatives. More than 5,800 enterprises depend on Informatica to fully leverage their information assets residing on-premise, in the Cloud and on the internet, including social networks.



Worldwide Headquarters, 2100 Seaport Blvd, Redwood City, CA 94063, USA Phone: 650.385.5000 Fax: 650.385.5500
Toll-free in the US: 1.800.653.3871 informatica.com [linkedin.com/company/informatica](https://www.linkedin.com/company/informatica) twitter.com/Informatica

© 2016 Informatica LLC. All rights reserved. Informatica® and Put potential to work™ are trademarks or registered trademarks of Informatica in the United States and in jurisdictions throughout the world. All other company and product names may be trade names or trademarks.