

Human or Machine?

Identifying, analyzing and filtering non-human website traffic

A practical guide for web analytics users

Table of Contents

Executive summary	1
1. Why non-human traffic changes the picture	2
2. Identifying traffic by organisation type	3
3. Cloud traffic is a signal - not a verdict	3
4. When a bot creates a conversion	4
5. Analyse human and machine traffic separately	4
6. A practical filtering checklist	4
7. Conclusion	5
8. How do I proceed?	5

Executive summary

Web analytics is normally used to understand people: who visited, what they viewed, where they came from and whether they converted.

But an increasing share of website activity is generated by software - search crawlers, AI scrapers, monitoring tools, security scanners and other automated systems.

The goal is not simply to remove bots, Snoobi Analytics already removes most standard bots from your analytics reports.

It is to separate traffic according to the question you are trying to answer. Human-focused analytics should show actual visitor behavior, while machine traffic should remain available for separate analysis because it can reveal how search engines, AI systems and other automated services interact with your content.

The key principle

Filtering should **improve meaning**, not just reduce numbers.

Keep human and (most) non-human traffic visible but analyze them in the right context.

1. Why non-human traffic changes the picture

Non-human traffic is website activity generated by software rather than by a person actively browsing a website.

Common sources include search-engine crawlers, AI crawlers and content scrapers, uptime monitoring services, SEO tools, automated testing, security scanners and malicious bots.

And this non-human traffic has increased more than 100% since AI training crawlers have exploded the internet traffic. Reports indicate that over half of website traffic is not from actual human visitors.

When this activity is mixed with normal visitor statistics, familiar metrics can become harder to interpret:

- Page views will be inflated when crawlers request hundreds of pages in a short period.
- Visitor and session counts rise even though no additional people visited your site.
- Popular-page reports will overvalue URLs that are easy for automated systems to discover.
- Direct traffic and geographic reports will reflect cloud datacenters rather than real audience locations.
- Engagement and time-based metrics will be distorted because machines do not browse like people.

Important distinction

The traffic itself is real. The problem is **interpreting automated activity** as if it represents human behavior.

Separate traffic by purpose - not by volume alone

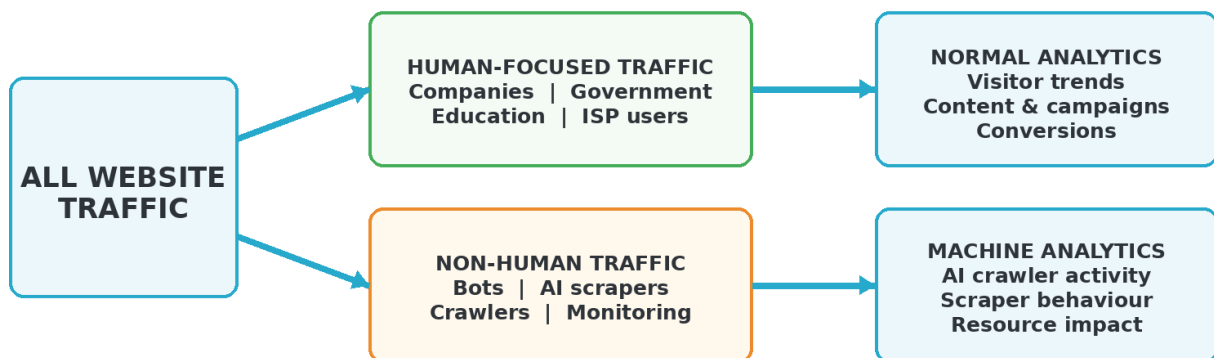


Figure 1. Separate traffic into human-focused analytics and machine analytics.

2. Identifying traffic by organization type

One useful signal is the type of network or organization from which a visit originates.

Snoobi Analytics can include or exclude visitor analytics based on organization categories such as Company, Government, Educational, Internet Service Provider (ISP) and Cloud/Hosting.

Cloud Provider (CLD) is the latest addition to the types of visiting organizations that Snoobi Analytics identifies. Where identification is available, Snoobi can also show the organization name, good for sales leads follow-up.

Organization type	Typical interpretation	Analytics use
Company	Employee or business-network traffic	Useful for B2B audience analysis
Government	Public-sector organization	Useful for public-sector engagement
Educational	University, school or research network	Useful for education/research audiences
ISP (Internet Service Provider)	Typical consumer or business internet access	Often represents ordinary human visitors
Cloud Provider	Datacenter or hosted infrastructure	Strong signal for possible automation; needs context

3. Cloud Provider traffic is a signal - not a verdict

A request from a cloud provider does **not** automatically mean the visitor is a bot.

Cloud infrastructure is also used for remote desktops, VPN services, browser testing, business applications and corporate systems.

Some automated traffic can originate from ordinary ISP networks as bots and scrapers may 'hide' their footprint in order for their visit not to be rejected.

Starting with filtering our Cloud Providers is a good start however as that covers 80-85% of the actual bot/scrapper traffic.

And if your client happens to be a Cloud Provider? Then Snoobi has the ability to label that organization in your report so they remain visible and can be analyzed.

The ultimate identification combines multiple signals, for example organization type, known cloud IP ranges, crawler identification, session information, visit frequency, page-request patterns, unusually high page counts or low page counts per visit.

The Traffic source and not the fact that bots and scrapers are accepting cookies is another indicator. Each of those can be used in Snoobi's segmentation, filtering or behavior classification.

4. When a bot creates a conversion

Automated traffic can sometimes trigger a conversion in analytics. A common example is a thank-you page that is normally reached after a form submission.

Even if an actual person cannot reach that page through normal navigation, a bot or scraper may discover the URL through a sitemap, an old link, an index or automated URL discovery.

Why page-based conversions can be misleading

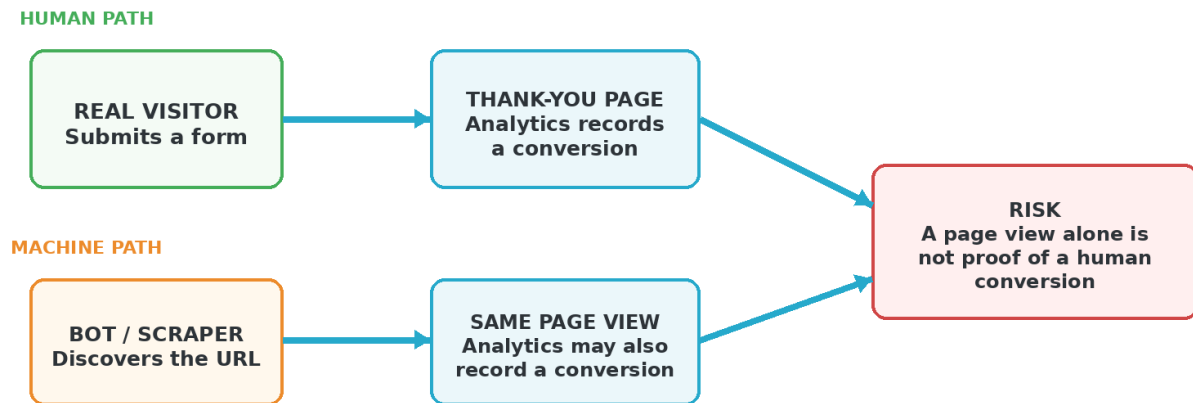


Figure 2. A page view can look like a conversion even when no person completed the action.

5. Analyze human and machine traffic separately

The best approach is not to discard non-human traffic, but to create a clean human-focused view while retaining automated activity for separate analysis.

HUMAN-FOCUSED ANALYTICS	MACHINE ANALYTICS
• Visitor trends • Content performance • Campaign analysis • Visitor journeys • Engagement • Lead generation • Conversions	• AI crawler activity • Search-engine crawling • Scraper behavior • Automated access patterns • Resource impact • Unusual machine activity • Bot trends over time

6. A practical filtering checklist

- Use the Snoobi segmentation and filtering tools to classify traffic by organization type and keep cloud/hosting networks visible as a separate category.
- Use known bot/crawler signals and engagement scoring in addition to network type; do not rely on one indicator alone.
- Exclude obvious automated activity from normal visitor and conversion reporting.
- Retain a separate report of bot, crawler and AI-scraper activity for trend and content analysis.
- Review conversions that rely only on a page view, especially thank-you or confirmation pages, use the Snoobi engagement scoring to your advantage.
- Where possible, record the successful action itself (such as a form submission) using the Snoobi event tracking option instead of relying only on the resulting page URL.

7. Conclusion

Web traffic is no longer generated only by people. Search engines, AI systems, crawlers, monitoring tools and scrapers are a normal and growing part of the web.

The most useful analytics therefore distinguishes between human behavior and machine activity rather than mixing everything into one total.

For human-focused reporting, filtering improves the reliability of visitor, content, campaign and conversion metrics.

For machine-focused reporting, the same automated traffic can reveal how AI systems, search engines and other services discover and consume website content.

A better analytics question

Do not ask only: "How much traffic did we receive?"

Ask: "What type of traffic did we receive, who or what generated it, and which part is relevant to the decision we are making?"

8. How do I proceed?

Snoobi Customer Support has extensive experience helping customers fine-tune their analytics and reports based on their actual business needs and reporting requirements.

If you would like a specialist opinion, contact us for a free, no-obligation review of your current analytics setup. We can look at the metrics and reports you are using and suggest practical improvements.

Depending on what we find, we can:

- provide simple guidance and support at no cost;
- offer training so your team can make the required changes and build the reports themselves; or
- set up a consulting project in which Snoobi creates the reports, dashboards and analysis you need, based on your specific objectives.

Whether you need a few practical tips or a complete reporting and analysis setup, we can help you choose the level of support that best fits your needs.