Investigating surface websites' ownership and history

Analysts collecting publicly available information (PAI) encounter various sites and services with valuable information. While this information is of intelligence value, there are biases, agendas, and different reasons for the dissemination of such information.

To identify these reasons, analysts have to find information on the individuals/organizations behind the site/service which hosted, maintained, and funded them.

This information is commonly obfuscated, but accessible with proper research tools and tradecraft.

Resources used for site ownership research

Analysts can leverage the following sites and services:

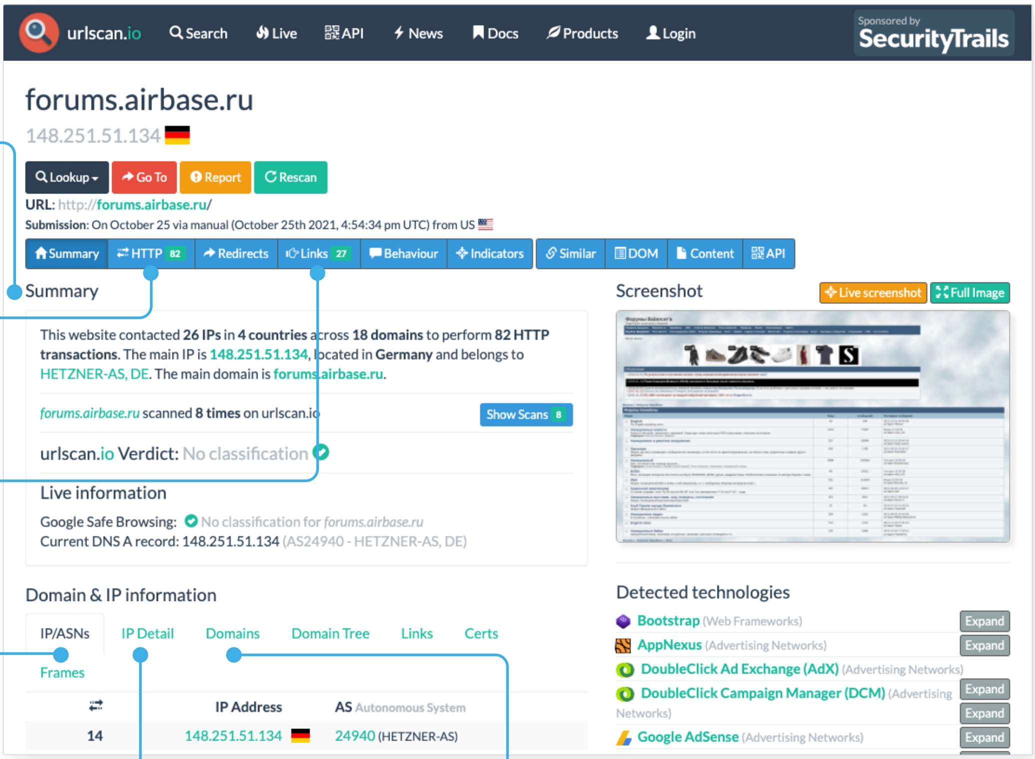
- **WHOIS records:** WHOIS records provide top level domain (e.g., russianmilitaryblog.com) information such as exact dates of registration, addresses, names, and phone numbers associated with the domain. In addition, it provides web host information.
  - URL Scan: https://urlscan.io
- **Advanced search engine use:** Using advanced search engines and search engine parameters on uniquely identifying information found on the site or WHOIS records (i.e., emails, names, mail servers, other IP addresses, etc.) can provide additional information on the site or service administrator/s.
  - Carbon Date: http://carbondate.cs.odu.edu
  - Google Dorking: https://www.google.com

On the following pages we describe how to use these tools and give examples of information that can be gleaned from them.

For more information please contact osint@authentic8.com.

WHOIS record analysis: URLscan.io

URLscan.io conducts analysis of a domain, providing the end user with information on all HTTP connections made during the site's retrieval, outbound links from the page, as well as detailed IP address information.



forums.airbase.ru

148.251.51.134

URL: http://forums.airbase.ru/

Submission: On October 25 via manual (October 25th 2021, 4:54:34 pm UTC) from US

Summary

This website contacted **26 IPs** in **4 countries** across **18 domains** to perform **82 HTTP transactions**. The main IP is 148.251.51.134, located in **Germany** and belongs to HETZNER-AS, DE. The main domain is forums.airbase.ru.

forums.airbase.ru scanned **8 times** on urlscan.io

urlscan.io Verdict: No classification

Live information

Google Safe Browsing: No classification for *forums.airbase.ru*

Current DNS A record: 148.251.51.134 (AS24940 - HETZNER-AS, DE)

Domain & IP information

IP/ASNs | IP Detail | Domains | Domain Tree | Links | Certs | Frames

⇄	IP Address	AS Autonomous System
14	148.251.51.134	24940 (HETZNER-AS)

Screenshot

Detected technologies

- Bootstrap (Web Frameworks)
- AppNexus (Advertising Networks)
- DoubleClick Ad Exchange (AdX) (Advertising Networks)
- DoubleClick Campaign Manager (DCM) (Advertising Networks)
- Google AdSense (Advertising Networks)

"Summary" provides a top level summary of what country the site is hosted in.

"HTTP" details how many HTTP connections are made during initial load.

"Links" details what other sites are linked to on the main page.

"IP/ASN" details the IPs of everything used upon initial load and the geographic location as well as ASN.

"IP Detail" contains the exact city/state/country an IP address is assigned to, and redirects.

"Domains" identifies how many subdomains a top level domain contains.

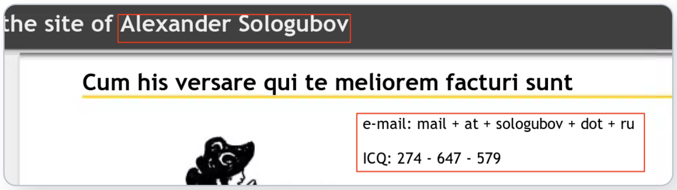
Example analysis of result panels

Forums.airbase.ru, a russian military forum, uses hosting primarily in Germany, which is likely due to Germany's strict data privacy laws. From the HTTP panel, the site uses Google Analytics for user tracking and also uses Yandex.ru for email. From the Links panel, a live "Telegram" chat is also available for users.

Example analysis of the result panel

Only one other IP aside from the current German IP has been used for hosting forums.airbase.ru.

This IP is 95.31.43.16, which also is used by a range of other domains — one of which, sologubov.ru has personal information on the individual behind forums.airbase.ru. This reveals the web host's full name, email, and ICQ number for further targeting.



the site of Alexander Sologubov

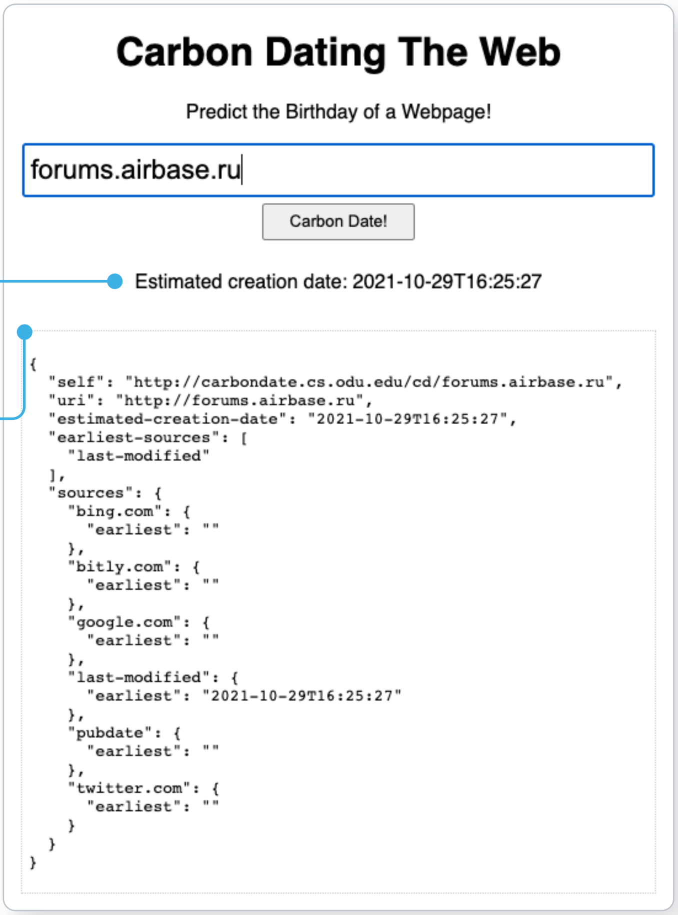
Cum his versare qui te meliorem facturi sunt

e-mail: mail + at + sologubov + dot + ru

ICQ: 274 - 647 - 579

Advanced search engine: Carbon Date

This advanced search engine automates advanced searches against web.archive.org, archive.md, Bing, bit.ly, Google, and Twitter to identify the earliest scrape/index or mention of a website on the web.



Carbon Dating The Web

Predict the Birthday of a Webpage!

forums.airbase.ru

Estimated creation date: 2021-10-29T16:25:27

```
{
  "self": "http://carbondate.cs.odu.edu/cd/forums.airbase.ru",
  "uri": "http://forums.airbase.ru",
  "estimated-creation-date": "2021-10-29T16:25:27",
  "earliest-sources": [
    "last-modified"
  ],
  "sources": {
    "bing.com": {
      "earliest": ""
    },
    "bitly.com": {
      "earliest": ""
    },
    "google.com": {
      "earliest": ""
    },
    "last-modified": {
      "earliest": "2021-10-29T16:25:27"
    },
    "pubdate": {
      "earliest": ""
    },
    "twitter.com": {
      "earliest": ""
    }
  }
}
```

"Estimated creation date" pulls the earliest date from the result set.

The result set shows the results from each source searched, and when available, a URL to the direct source itself.

Example analysis of the results panel

The earliest mention of forums.airbase.ru was in October of 2003. To view the first ever scrape of this site by web.archive.org, use the URL in the "uri-m" field.

Advanced search engine: Google Dorking

Advanced Google search parameters and features are used in a technique called "Google Dorking."

Users must combine various search parameters to effectively search and filter down results of interest to them.

The most commonly used Google Dorks are:

- **Intitle:** identifies any mention of search text in the web page title
- **Allintitle:** only identifies pages with all of the search text in the web page title
- **Inurl:** identifies any mention of search text in the web page URL
- **Intext:** only identifies pages with all of the search text in the web page URL
- **Site:** limits results to the specified file type
- **Filetype:** limits results to only the specified file type
- **Cache:** shows the most recent cache of a site specified
- **Around (X):** searches for two different words within X words of one another

The most commonly used Boolean logic search operators are:

- **AND:** searches for content mentioning two phrases anywhere
- **OR:** used in multi-part search, searches for content mentioning any combination of the first search term and two unique second search variables
- *** :** the asterisk acts as a wildcard and searches for any word or phrase
- **– :** the dash excludes any specific word or phrase (if using brackets or quotation marks)
- **( ) :** the parenthesis group specific terms or search operators together

Example analysis using advanced Google Search parameters

```
site:sologubov.ru ICQ OR email
```

This search will find mentions of ICQ or email on a site of interest, resulting in an ICQ number and email previously unknown to an analyst.

```
site:forums.airbase.ru contact OR admin OR mod OR moderator OR donation
```

This search will find uniquely identifying information that can be linked to a person, such as mentions of a moderator, a contact page, or a donation page (such as Paypal, Bitcoin, etc), resulting in multiple pages with mentions of the moderator and a donation page for their health bills.

```
“95.31.43.16”
```

This search will find exact mentions of forums.airbase.ru, resulting in mentions on another forum of Russian censorship of the servers IP address.

Conclusion

This workflow covers how to investigate the ownership and hosting information related to a site/service of interest. Results from the analysis include key identifiers such as server IPs, other related domains, and the webhost’s email address/name/ICQ number that can then be incorporated further into a finished intelligence product.

For more information please contact osint@authentic8.com.