

Internet Engineering Task Force  
Internet-Draft  
Intended status: Informational  
Expires: October 3, 2013

L. Zhou  
N. Kong  
S. Shen  
CNNIC  
S. Sheng  
ICANN  
Apr 2013

Domain Name Registration Data Access Protocol Object Inventory Analysis  
[draft-zhou-weirds-dnrd-ap-object-inventory-01](#)

## Abstract

WHOIS output elements from 124 TLDs were collected and analyzed. This document describes the statistical analysis process and result of WHOIS information. The purpose of this document is to build an object inventory to facilitate discussions of domain name data objects in WHOIS response.

## Status of this Memo

This Internet-Draft is submitted in full conformance with the provisions of [BCP 78](#) and [BCP 79](#).

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://datatracker.ietf.org/drafts/current/>.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on October 3, 2013.

## Copyright Notice

Copyright (c) 2013 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to [BCP 78](#) and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must

include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

This document may contain material from IETF Documents or IETF Contributions published or made publicly available before November 10, 2008. The person(s) controlling the copyright in some of this material may not have granted the IETF Trust the right to allow modifications of such material outside the IETF Standards Process. Without obtaining an adequate license from the person(s) controlling the copyright in such materials, this document may not be modified outside the IETF Standards Process, and derivative works of it may not be created outside the IETF Standards Process, except to format it for publication as an RFC or to translate it into languages other than English.

## Table of Contents

<a href="#">1.</a>	Introduction . . . . .	<a href="#">3</a>
<a href="#">2.</a>	Methodology . . . . .	<a href="#">3</a>
<a href="#">3.</a>	Terminology . . . . .	<a href="#">4</a>
<a href="#">4.</a>	Analysis . . . . .	<a href="#">4</a>
<a href="#">4.1.</a>	Overview . . . . .	<a href="#">5</a>
<a href="#">4.2.</a>	Public Objects . . . . .	<a href="#">5</a>
<a href="#">4.2.1.</a>	WHOIS Data of Domain . . . . .	<a href="#">5</a>
<a href="#">4.2.2.</a>	WHOIS Data of Contact . . . . .	<a href="#">6</a>
<a href="#">4.2.2.1.</a>	Registrant . . . . .	<a href="#">6</a>
<a href="#">4.2.2.2.</a>	Admin Contact . . . . .	<a href="#">6</a>
<a href="#">4.2.2.3.</a>	Tech Contact . . . . .	<a href="#">7</a>
<a href="#">4.2.2.4.</a>	Billing Contact . . . . .	<a href="#">8</a>
<a href="#">4.2.3.</a>	WHOIS Data of Nameserver . . . . .	<a href="#">8</a>
<a href="#">4.2.4.</a>	WHOIS Data of Registrar . . . . .	<a href="#">9</a>
<a href="#">4.3.</a>	Other Objects . . . . .	<a href="#">9</a>
<a href="#">5.</a>	Conclusion . . . . .	<a href="#">11</a>
<a href="#">5.1.</a>	Preliminary Statistics . . . . .	<a href="#">11</a>
<a href="#">5.2.</a>	Data Elements Analysis . . . . .	<a href="#">13</a>
<a href="#">5.3.</a>	Labels Analysis . . . . .	<a href="#">14</a>
<a href="#">5.4.</a>	Other Objects Analysis . . . . .	<a href="#">15</a>
<a href="#">6.</a>	Limitation . . . . .	<a href="#">15</a>
<a href="#">7.</a>	IANA Considerations . . . . .	<a href="#">16</a>
<a href="#">8.</a>	Security considerations . . . . .	<a href="#">16</a>
<a href="#">9.</a>	Acknowledgements . . . . .	<a href="#">16</a>
<a href="#">10.</a>	Normative References . . . . .	<a href="#">16</a>
	Authors' Addresses . . . . .	<a href="#">17</a>



## **1. Introduction**

Internet registries for both number resources and names have historically maintained a lookup service to permit public access to some portion of the registry database. Most registries offer the service via the WHOIS protocol [[RFC3912](#)], with additional services being offered via world wide web pages, bulk downloads, and other services, such as RPSL [[RFC2622](#)].

Although the WHOIS protocol specified in [[RFC3912](#)] is widely adopted and supported, it has several shortcomings that limits its usefulness to the evolving needs of the Internet community. For example, the WHOIS protocol has not been Internationalized, it does not consistently support Internationalized Domain Name (IDN, described in [[RFC5890](#)]); WHOIS has no query and response format; and WHOIS protocol does not support user authentication, access control for differentiated access.

This document is aimed to build an object inventory to facilitate discussions of domain name data objects in WHOIS response. Based on this statistics result, it may help to form and specify the response format.

In domain name space, there are now over 200 ccTLDs and 21 gTLDs. Different domain name registries may have different WHOIS response objects and formats. A common understanding of domain name data format is critical.

This document describes the WHOIS data collection procedures and gives a data object inventory analysis based on the collected data from 106 ccTLDs and 18 gTLDs. The statistics result includes port 43 response data and web response data. All the data objects are classified into domain, contact, nameserver and registrar related objects. Other objects that do not belong to above four categories are viewed as private designed objects.

Since this is a document with statistics analysis and there are no protocol specifications, the [[RFC2119](#)] language does not apply.

## **2. Methodology**

WHOIS information, including port 43 response and web response data, is collected following the procedures descibed below.



- (1) A programming script is applied to collect port 43 response data from 294 ccTLDs. "nic.ccTLD" is used as the query string which is usually registered in a domain registry. Responses of 106 ccTLDs were received. 18 gTLDs' port 43 response data is collected from their contracts with ICANN. So the sample size of port 43 WHOIS response data is 124 registries in total.
- (2) WHOIS data from web is collected manually from the 124 registries that have port 43 WHOIS responses.
- (3) Some of the responses which are collected by program may not seem to be correct. So data of top 10 ccTLD registries, like .de, .eu and .uk etc., was re-verified by querying domain names other than "nic.ccTLD".
- (4) In accordance with the specification 4 of new gTLD applicant guide book, [[RFC5730](#)], [[RFC5731](#)], [[RFC5732](#)] and [[RFC5733](#)], the response data objects are classified into public and other data objects. Public data objects are those which are defined in the above two documents. Other objects are those which are self designed data elements or objects in different registries.
- (5) Data elements with the same meaning, but using different labels, are grouped together. The numbers of registries that support the data elements is calculated in the total count column.

### **3. Terminology**

- o Data element -- The name of specific response object.
- o Label -- Different registries may have different naming ways for the same data element. So there may be several labels with the same meaning and belong to one group of data element.
- o Most popular label -- The label which is most supported by the registries.
- o Number of labels -- The number of different labels.
- o Total count -- The number of registries that support a certain data element.

### **4. Analysis**



#### 4.1. Overview

WHOIS data is collected from 124 registries, including 106 ccTLDs and 18 gTLDs. All the 124 registries support domain query. Among 124 registries, 8 ccTLDs and 15 gTLDs support contact query. 10 ccTLDs and 18 gTLDs support name server query. 4 ccTLDs and 18 gTLDs support registrar query. Domain WHOIS data contains 68 data elements that use a total of 550 labels. There are total 392 other objects for Domain WHOIS data. The raw data can be accessed with the following link: WHOIS Statistics Data File [[Stat-Data-File](#)]

#### 4.2. Public Objects

As mentioned above, public objects are those data elements selected according to new gTLD application guide book, [[RFC5730](#)], [[RFC5731](#)], [[RFC5732](#)] and [[RFC5733](#)]. They are generally classified into four categories: domain, contact, nameserver and registrar related information.

##### 4.2.1. WHOIS Data of Domain

WHOIS data of domain includes "Domain Name", "Creation Date", "Domain Status", "Expiration Date", "Updated Date", "Domain ID", "DNSSEC" and "Last Transferred Date". The following table gives the element name, most popular label and corresponding numbers of TLDs and labels.

Data Element	Most Popular Label	No. of TLDs	No. of Labels
Domain Name	Domain Name	118	6
Creation Date	Created	106	24
Domain Status	Status	95	8
Expiration Date	Expiration Date	81	21
Updated Date	Modified	70	20
Domain ID	Domain ID	34	5
DNSSEC	DNSSEC	14	4
Last Transferred Date	Last Transferred Date	4	3

WHOIS Data of Domain

Analyzing the above data, about 95.16% of the 124 registries support "Domain Name" data element; 85.48% of the 124 registries support "Creation Date" data element; 76.61% of the 124 registries support "Domain Status" data element. On the other hand, some elements such as "DNSSEC" and "Last Transferred Date" are only supported by 11.29%





and 3.23% of all the registries seperatedly.

#### **4.2.2. WHOIS Data of Contact**

In domain name space, contacts are typically devided into registrant, administrative contact, technical contact and billing contact.

##### **4.2.2.1. Registrant**

The following table shows all the contact information of registrant. 14 data elements are listed below.

Data Element	Most Popular Label	No. of TLDs	No. of Labels
Registrant Name	Name	65	7
Registrant Email	Registrant Email	59	7
Registrant ID	Registrant ID	50	12
Registrant Phone	Registrant Phone	48	6
Registrant Fax	Registrant Fax	44	6
Registrant Organization	Registrant Organization	42	4
Registrant Country Code	Country	42	6
Registrant City	Registrant City	38	4
Registrant Postal Code	Registrant Postal Code	37	5
Registrant State/Province	Registrant State/Province	32	4
Registrant Street	Registrant Street1	31	16
Registrant Country	Registrant Country	19	4
Registrant Phone Ext.	Registrant Phone Ext.	18	2
Registrant Fax Ext	Registrant Fax Ext	17	2

#### **Registrant**

Among all the data elements, "Registrant Name", "Registrant Email" and "Registrant ID" are the top 3 data elements.

##### **4.2.2.2. Admin Contact**

The following table shows all the contact information of administrative contact. 14 data elements are listed below.



Data Element	Most Popular Label	No. of TLDs	No. of Labels
Admin Street	Address	64	19
Admin Name	Admin Name	60	9
Admin Email	Admin Email	54	12
Admin ID	Admin ID	52	16
Admin Fax	Admin Fax	44	8
Admin Phone	Admin Phone	43	9
Admin Organization	Admin Organization	42	9
Admin Country Code	Country	42	7
Admin City	Admin City	35	5
Admin Postal Code	Admin Postal Code	35	7
Admin	Admin	28	5
State/Province	State/Province		
Admin Country	Admin Country	17	5
Admin Phone Ext.	Admin Phone Ext.	17	3
Admin Fax Ext.	Admin Fax Ext.	17	3

#### Admin Contact

Among all the data elements, "Admin Street", "Admin Name" and "Admin Email" are the top 3 data elements.

#### [4.2.2.3.](#) Tech Contact

The following table shows all the information about domain name technical contact. 14 data elements are listed below.

Data Element	Most Popular Label	No. of TLDs	No. of Labels
Tech Email	Tech Email	59	9
Tech ID	Tech ID	55	16
Tech Name	Tech Name	47	6
Tech Fax	Tech Fax	45	9
Tech Phone	Tech Phone	45	10
Tech Country Code	Country	43	9
Tech Organization	Tech Organization	39	7
Tech City	Tech City	36	4
Tech Postal Code	Tech Postal Code	36	7
Tech	Tech	30	4
State/Province	State/Province		
Tech Street	Tech Street1	27	16
Tech Country	Tech Country	18	5



Tech Fax Ext	Tech Fax Ext	18	3	
Tech Phone Ext.	Tech Phone Ext.	13	3	
+-----+	+-----+	+-----+	+-----+	+-----+

#### Tech Contact

Among all the data elements, "Tech Email", "Tech ID" and "Tech Name" are the top 3 data elements.

#### [4.2.2.4.](#) Billing Contact

The following table shows all the information about domain name billing contact. 14 data elements are listed below.

Data Element	Most Popular Label	No. of TLDs	No. of Labels
Billing Name	Name	47	5
Billing Fax	Fax	43	6
Billing Email	Email Address	42	7
Billing Country Code	Country	38	4
Billing Phone	Phone Number	34	6
Billing ID	Billing ID	28	9
Billing City	Billing City	28	4
Billing Organization	Billing	28	5
Billing Postal Code	Organization		
Billing State/Province	Billing Postal Code	27	4
Billing Street	Billing	21	4
Billing Country	State/Province		
Billing Phone Ext.	Billing Street1	19	13
Billing Fax Ext	Billing Country	13	5
	Billing Phone Ext.	10	2
	Billing Fax Ext	10	2

#### Billing Contact

Among all the data elements, "Billing Name", "Tech Fax" and "Billing Email" are the top 3 data elements.

#### [4.2.3.](#) WHOIS Data of Nameserver

114 registries (about 92% of all the 124 registries) have the nameserver data element in their WHOIS response. But there are 63 different labels for this element. Top 3 labels for this element are



Name Server which is supported by 25% of all the registries, Name Servers which is supported by 16% of all the registries and nserver which is supported by 12% of all the registries.

Data Element	Most Popular Label	No. of TLDs	No. of Labels
NameServer	NameServer	114	63

#### WHOIS Data of Nameserver

Some registries have nameserver elements such like "nameserver 1", "nameserver 2" till "nameserver n". So there are more labels than of other data elements.

#### [4.2.4.](#) WHOIS Data of Registrar

There are three data elements about registrar information.

Data Element	Most Popular Label	No. of TLDs	No. of Labels
Sponsoring Registrar	Registrar	84	6
Created by Registrar	Created by	14	3
Updated by Registrar	Last Updated by Registrar	11	3

#### WHOIS Data of Registrar

67.7% of the registries have sponsoring registrar data element. Elements such as "Created by Registrar" and "Updated by Registrar" are supported by 11.3% and 8.9% of the registries.

#### [4.3.](#) Other Objects

So called other objects are those data elements that are self-designed or are difficult to be classified. There are 392 other objects altogether. The following tables lists the top 50 other objects according to the data collection result.





Data Element	No. of TLDs
Registrant	41
Phone	32
Technical contact	26
Administrative contact	15
source	14
fax-no	13
nic-hdl	13
Billing Contact	12
referral url	11
e-mail	10
WHOIS server	9
Admin Contact	9
Type	9
Website	9
zone-c	8
remarks	7
Registration URL	6
anonymous	6
anniversary	6
hold	6
nsl-id	6
obsoleted	6
Customer Service Contact	5
Customer Service Email	4
Registrar ID	4
org	4
person	4
Maintainer	4
Nombre	3
Sponsoring Registrar IANA ID	3
Trademark Number	3
Trademark Country	3
descr	3
url	3
Postal address	3
Registrar URL	3
International Name	3
International Address	3
Admin Contacts	2
Contractual Language	2
Date Trademark Registered	2
Date Trademark Applied For	2
IP Address	2
Keys	2
Language	2



	NIC handle		2	
	Record maintained by		2	
	Registration Service Provider		2	
	Registration Service Provided By		2	
	Registrar URL (registration services)		2	
+-----+-----+-----+				

### Top 50 Other Objects

Some elements like "Registrant" are difficult to be classified into any categories. A few registries have two levels of data elements, for example:

```
Registrant:
  Name:
  Email:
  ...
```

We do not think the first level of elements belong to any part. So they are put into the scope of other objects.

Some other data elements, like "Remarks", "anniversary" and "Customer service Contact" etc., are designed particularly for their own purpose by different registries.

## 5. Conclusion

### 5.1. Preliminary Statistics

Some preliminary statistics conclusion could be drawn from the raw data.

- o All of the 124 domain registries have the object names in their responses although they are in various formats.
- o Of the 118 whois services contacted, 65 registries show their registrant contact. About half of the registries (60 registries) support admin contact information. There are 47 registries, that is about one third of the total number, have technical and billing contact information. And only 7 of all the 124 registries give their abuse email in remarks part. No explicit abuse contact information is provided. The
- o There are mainly two presentation formats. One is key:value, the other is data block format. Example of key-value format:



## Domain Information

Query: na-nic.com.na

Status: Delegated

Created: 17 Apr 2004

Modified: 14 Nov 2010

Expires: 31 Dec 9999

Name Servers: oshikoko.omadhina.net

ns1.na.afrisp.net

ns2.na.afrisp.net

...

Example of data block format:



Whois database  
domain nic.vg

Domain Name nic.vg  
Registered 1998-09-02  
Expiry 2012-09-02

#### Resource Records

a 195.153.6.27  
mx 10 terpsichore.william.org  
www a 195.153.6.27

#### Contact details

Registrant,  
Technical Contact,  
Billing Contact,  
Admin. Contact AdamsNames Reserved Domains (i)  
These domains are not available for registration  
United Kingdom  
Identifier: neams048s

#### Servidor Whois de NIC-Venezuela (.VE)

Este servidor contiene informacion autoritativa exclusivamente de dominios .VE Cualquier consulta sobre este servicio, puede hacerla al correo electronico whois@nic.ve

#### Titular:

Jhonny Valera (nic.ve-dom) jhovalera@conatel.gob.ve  
Comision Nacional de Telecomunicaciones  
Av. Veracruz con calle Cali, Edif Aguila, Urb. Las Mercedes  
Caracas, Distrito Capital VE  
0212-9090493 (FAX) +582127718599

- o 11 registries give local script responses. The WHOIS information of other registries are all represented in English.

## **5.2. Data Elements Analysis**

Top 10 data elements are as follows:





Data Element	No. of TLDs
Domain Name	118
Name Server	114
Creation Date	106
Domain Status	95
Sponsoring Registrar	84
Expiration Date	81
Updated Date	70
Registrant Name	65
Admin Street	64
Admin Name	60

#### Top 10 Data Elements

Most of the domain related WHOIS information is included in the top 10 data elements. Other information like name server and registrar name are also supported by most registries.

We did a cumulative distribution analysis of all the data elements.

- (1) About 5% of data elements are supported by over 111 (90%) registries.
- (2) About 30% of data elements are supported by over 44 (35%) registries.
- (3) About 60% of data elements are supported by over 32 (26%) registries.
- (4) About 90% of data elements are supported by over 14 (11%) registries.

From the above result, we can conclude that only a few registries support all the public objects, most of the registries support just parts of all the objects.

### 5.3. Labels Analysis

The top 10 labels of different data elements include:



Labels	No. of Labels
Name Server	63
Creation Date	24
Expiration Date	21
Updated Date	20
Admin Street	19
Tech ID	18
Registrant Street	16
Admin ID	16
Tech Street	16
Billing Street	13

Top 10 Labels

As explained above, name server label is a unique example that many registries define the name server elements from "nameserver 1" till "nameserver n". So label numbers of name server are much more than other elements. Other elements about date, street name have more labels.

A cumulative distribution analysis of label numbers was done. About 90% of data elements have more than 2 Labels. So it is very necessary to specify a standard and unified format for object names of WHOIS response.

#### 5.4. Other Objects Analysis

According to statistics result, there are 392 other data objects in total that are not easy to be classified or privately owned by various registries. Top 50 other objects are listed in the table in [section 4.3](#). You can find that various different objects are designed for some particular purpose. So in order to ensure uniqueness of JSON names used in the Restful Whois service, establishing an IANA registry is a necessary requirement.

## 6. Limitation

- o The input "nic.ccTLD" maybe is not a good choice.
- o The classification of local script data elements may not be accurate. 11 registries give local script responses.
- o The extension data elements are used randomly by different registries. It is difficult to do statistical analysis.



- o Sample sizes of contact, name server and registrar queries are small.
- \* We only use "Whois contactID/nameserver/registrar" as the query commands to check.
- \* Some registries may not support contact, name server or registrar queries.
- \* Some may not support query contact by ID.
- \* Contact information of some registries may be protected.

## **7. IANA Considerations**

This document does not specify any IANA actions.

## **8. Security considerations**

This document does not does not provide any other security services or introduce any additional considerations.

## **9. Acknowledgements**

This document has been reviewed and improved by the design team. The authors especially thank the following individuals who gave their suggestions and contributions to this document: Guangqing Deng, Frederico A C Neves and Ray Bellis.

## **10. Normative References**

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", [BCP 14](#), [RFC 2119](#), March 1997.
- [RFC2622] Alaettinoglu, C., Villamizar, C., Gerich, E., Kessens, D., Meyer, D., Bates, T., Karrenberg, D., and M. Terpstra, "Routing Policy Specification Language (RPSL)", [RFC 2622](#), June 1999.
- [RFC3912] Daigle, L., "WHOIS Protocol Specification", [RFC 3912](#), September 2004.
- [RFC5730] Hollenbeck, S., "Extensible Provisioning Protocol (EPP)", STD 69, [RFC 5730](#), August 2009.



- [RFC5731] Hollenbeck, S., "Extensible Provisioning Protocol (EPP) Domain Name Mapping", STD 69, [RFC 5731](#), August 2009.
- [RFC5732] Hollenbeck, S., "Extensible Provisioning Protocol (EPP) Host Mapping", STD 69, [RFC 5732](#), August 2009.
- [RFC5733] Hollenbeck, S., "Extensible Provisioning Protocol (EPP) Contact Mapping", STD 69, [RFC 5733](#), August 2009.
- [RFC5890] Klensin, J., "Internationalized Domain Names for Applications (IDNA): Definitions and Document Framework", [RFC 5890](#), August 2010.
- [Stat-Data-File]  
Kong, N., Zhou, L., and G. Deng, "WHOIS Statistics Data File", July 2012, <<https://docs.google.com/open?id=0B96TtoK8a--MTTRuVUt3UHZMdEk>>.

#### Authors' Addresses

Linlin Zhou  
CNNIC  
4 South 4th Street, Zhongguancun, Haidian District  
Beijing, Beijing 100190  
China

Phone: +86 10 5881 2677  
Email: zhoulinlin@cnnic.cn

Ning Kong  
CNNIC  
4 South 4th Street, Zhongguancun, Haidian District  
Beijing, Beijing 100190  
China

Phone: +86 10 5881 3147  
Email: nkong@cnnic.cn





Sean Shen

CNNIC

4 South 4th Street, Zhongguancun, Haidian District

Beijing, Beijing 100190

China

Phone: +86 10 5881 3038

Email: shenshuo@cnnic.cn

Steve Sheng

ICANN

4676 Admiralty Way, Suite 330

Marina del Rey, CA 90292

United States of America

Phone: +1.310.823.9358

Email: steve.sheng@icann.org

