Internet Engineering Task Force Internet-Draft Intended status: Informational Expires: January 24, 2015 L. Zhou N. Kong S. Shen CNNIC S. Sheng ICANN A. Servin LACNIC July 23, 2014

Registration Data Access Protocol Object Inventory Analysis draft-ietf-weirds-object-inventory-03

Abstract

WHOIS output objects from registries (including both Regional Internet Registries (RIRs) and Domain Name Registries (DNRs)) were collected and analyzed. This document describes the statistical analysis process and result of existing WHOIS information. The purpose of this document is to build an object inventory to facilitate discussions of data objects included in Registration Data Access Protocol (RDAP) responses.

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of <u>BCP 78</u> and <u>BCP 79</u>.

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 January 24, 2015.

Copyright Notice

Copyright (c) 2014 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to <u>BCP 78</u> and the IETF Trust's Legal Provisions Relating to IETF Documents

Zhou, et al. Expires January 24, 2015 [Page 1]

(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

$\underline{1}$. Introduction		
<u>2</u> . Terminology		<u>4</u>
<u>3</u> . Methodology		<u>4</u>
4. RIR Objects Analysis		<u>5</u>
4.1. WHOIS Data of Organizations	Holding a Resource	· · <u>4</u> · · <u>5</u> · · <u>5</u>
4.2. WHOIS Data of Contact		
4.3. WHOIS Data of IP		· · <u>6</u> · · <u>7</u> · · <u>9</u> · · 9
4.4. WHOIS Data of ASN		9
4.5. Conclusion		9
5. DNR Objects Analysis		
<u>5.1</u> . Overview		
5.2. Public Objects		
5.2.1. WHOIS Data of Domain .		
5.2.2. WHOIS Data of Contact .		
<u>5.2.2.1</u> . Registrant		
5.2.2.2. Admin Contact		
5.2.2.3. Tech Contact		
5.2.2.4. Billing Contact		
5.2.3. WHOIS Data of Nameserve		
5.2.4. WHOIS Data of Registrar		
<u>5.3</u> . Other Objects		
5.4. Conclusion		
5.4.1. Preliminary Statistics		
5.4.2. Data Elements Analysis		
<u>5.4.3</u> . Labels Analysis		

<u>5.5</u> . Limitation
<u>6</u> . Reference Extension Objects
6.1. RIR Reference Extension Objects
6.2. DNR Reference Extension Objects
7. IANA Considerations
8. Security considerations
9. Acknowledgements
<u>10</u> . Normative References
Appendix A. Change Log
Authors' Addresses

1. Introduction

RIRs and DNRs 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 registration data objects. Based on this statistics result, it may help to form and specify the RDAP response format.

In number space, there are altogether 5 RIRs. All RIRs provide information about IP addresses, Autonomous System Number (ASNs) and contacts, the data model used is different for each RIR. 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 all these data formats is critical.

This document describes the WHOIS data collection procedures and gives a data object inventory analysis based on the collected data from 5 RIRs and 106 ccTLDs and 18 gTLDs from DNRs. The RIR data objects are classified into IP address, ASN, person or contact and the organization that held the resource. The DNR data objects are classified into domain, contact, nameserver and registrar related objects. Other objects that do not belong to above categories are viewed as private designed objects. In this document, we are not

intended to analyze all the query and response types existed in RIRs and DNRs. The most common query objects will be discussed, other objects such as RPSL data structure used by Internet Routing Registries (IRRs) will be added at a later time if the community feels it is necessary.

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

2. 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.

3. Methodology

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

RIR objects collection process:

- (1) The process of RIR data collection is relatively easy. There are altogether 5 RIRs which are AFRINIC, APNIC, ARIN, LACNIC and RIPE NCC. All the RIRs provide information of IPs, ASNs and contacts. Find the 5 RIR WHOIS servers firstly.
- (2) Query the corresponding IPs, ASNs, contacts and organizations registered in 5 RIRs and make a comparative analysis of the responses data.
- (3) Data elements with the same meaning, but using different labels, are grouped together.

DNR objects collections process:

(1) A programming script is applied to collect port 43 response data from 294 ccTLDs. "nic.ccTLD" is used as the guery 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.

4. RIR Objects Analysis

4.1. WHOIS Data of Organizations Holding a Resource

The following table shows the organization objects of 5 RIRs.

+	+	+	+	+	++
RIR	AFRINIC	APNI	ARIN	LACNIC	RIPE NCC
0bjects		C			
++	++	+	+	++	++
Organizati	organisat	NA	Name	0wner	org-name
on name	ion				
Organizati	org-name	NA	Handle	owner-id	organisati
on ID					on
Company	NA	NA	Company	NA	NA
Name of	NA	NA	NA	responsib	NA
person res				le	
ponsible					
Type of or	org-type	NA	NA	NA	org-type
ganization					
Country	country	NA	country	country	country
Postal	address	NA	address	address	address
Address					
City	NA	NA	city	NA	address

State	NA	NA	StatePro	NA NA	address
Postal	NA	I I NA	v PostalCo	I NA	 address
Code	NA		de		
Phone	nhana	I I NA	i na	 nhana	 nhana
1	phone	1	1	phone	phone
Fax Number	fax-no	NA	NA	NA	fax-no
ID of admi	admin-c	NA	Admin	owner-c	admin-c
nistrative			POC		(multiple)
contact					
ID of	tech-c	NA	Tech POC	tech-c	tech-c
technical					(multiple)
contact					
Reference	mnt-ref	NA	NOC POC	NA NA	mnt-ref
of					
maintainer					
Reference	mnt-by	NA	Abuse	NA NA	mnt-by
of			P0C		
maintainer					
Remarks	remarks	NA	NA NA	NA	remarks
Date of	Changed	NA	RegDate	created	Changed
record		ĺ			
creation		İ	Ì	ĺ	
Date of	changed	I NA	Updated	changed	changed
record	5	i	i .	ĺ	
changed		i	i	İ	i
List of	NA	I NA	I NA	 list of	NA I
resources		İ		resources	
Source	source	I NA	I NA	NA NA	source
Reference	NA	I NA	Ref	I NA	NA I
кетегепсе ++	NA 	NA +	кет +	NA +	NA +

WHOIS Data of Organizations Holding a Resource

4.2. WHOIS Data of Contact

The following table shows the contact objects of 5 RIRs.

+	+	+	+	+	+
RIR Objects 	AFRINIC 	APNIC	ARIN	LACNIC	RIPE NCC
Name	person	person	Name	person	person
Company	NA	NA	Company	NA NA	NA
Postal	address	address	Address	address	address
Address					
City	NA	NA	City	NA	address
State	NA NA	NA	StateProv	NA NA	address
Postal Code	NA NA	NA	PostalCode	NA NA	address
Country	NA NA	country	Country	country	NA
Phone	phone	phone	Mobile	phone	phone
Fax Number	fax-no	fax-no	Fax	NA NA	fax-no
Email	e-mail	e-mail	Email	e-mail	NA
ID	nic-hdl	nic-hdl	Handle	nic-hdl	nic-hdl
Remarks	remarks	remarks	Remarks	NA NA	remarks
Notify	notify	notify	NA	NA NA	notify
ID of	mnt-by	mnt-by	NA	NA NA	mnt-by
maintainer					
Registration	changed	NA	RegDate	created	changed
Registration update	 changed 	 changed 	Updated	 changed 	changed
Source	source	source	NA	NA	source
Reference	NA	NA NA	Ref	I NA	NA
+	+	+	+	+	

WHOIS Data of Contact

4.3. WHOIS Data of IP

The following table shows the IP address objects of 5 RIRs.

+	+		+	+	F+
RIR Objects	AFRIN IC	APNIC	ARIN	LACNIC	RIPE NCC
+	+		+	+	++
IP	inetn	inetnum	NetRange	NA	inetnum
address	um l		ĺ		i i
range					
IPV6	inet6	inet6num/i	CIDR	inetnu	inet6num/in
address	num	netnum		m	etnum
range	İ				İ İ
Descripti	descr	descr	NetName	NA	descr
on					İ İ
Remarks	remar	remarks	NA	NA	remarks
	ks				

Origin AS 	NA 	origin (on route/6)	OriginAS 	Origin AS (fu ture)	origin (on route/6)
Network name/ID	netna me	netname	 NetHandle 	inetre v	netname
Maintaine r	mnt- by	NA	NA I	NA	mnt-by
Maintaine r	mnt- lower	NA	NA NA	NA I	NA
Administr ative contact	admin -c	admin-c	OrgId 	owneri d	admin-c
Parent range	paren t	NA	Parent 	NA	NA
Status	statu s	status	NetType	status 	status
 Registrat ion Date	chang ed	NA	RegDate	create d	changed
Registrat ion update	chang ed	changed (multiple)	 Updated 	change d	changed
Reference ID organi zation holding the	NA org 	NA NA	Ref OrgId 	NA owner 	NA organisatio n
resource Referral server	NA	NA	 ReferralS erver	 NA	NA
Technical contact	tech- c	tech-c	OrgTechHa ndle	tech-c 	tech-c (multiple)
Abuse contact	NA	NA	OrgAbuseH andle	abuse- c	abuse- mailbox
Referral technical contact	NA 	NA	RTechHand le	NA 	NA
Referral abuse contact	mnt- irt	mnt-irt	RAbuseHan dle	NA NA	NA
Referral NOC	NA	NA	RNOCHandl e	NA	NA
contact Name server	NA 	NA	 NA 	 nserve r	NA

WHOIS Data of IP

4.4. WHOIS Data of ASN

+	+	+	+	++	++
RIR	AFRINI	APNIC	ARIN	LACNIC	RIPE NCC
0bjects	C				i i
+	+	+	+	+	++
ID	aut-	aut-num	ASNumber	aut-num	aut-num
	num				
Descriptio	descr	descr	NA	NA	descr
n					
Organizati	org	NA	OrgId	owner	org
on on					
Comment	remark	NA	Comment	NA	remarks
	S				
Administra	admin-	admin-c	ASHandle	owner-	admin-c
tive	С			id	
contact ID		tach a	OngTachlland	nouting	
Technical	tech-c	tech-c	OrgTechHand	routing	tech-c (m
contact ID		nic-hdl	le	- C	ultiple)
Organizati on ID	NA NA	nic-nac	NA	owner-c	organisat ion
-	 potify/	notify	NA	NA	I ION I
Notify Abuse	notify NA	notify NA	NA OrgAbuseHan	abuse-c	NA NA
contact		NA	dle	abuse-c	
Maintainer	 mnt-by	mnt-by	NA	NA	mnt-by
Maintainer	mnt-	mnt-	NA	NA	mnt-lower
	llower	lower	NA NA	NA	l mirc-cower
I Maintainer		NA	NA	NA	mnt-ref
Registrati	change	NA	RegDate	created	NA I
on Date	l d		negbare	ereacea	
Registrati	change	changed	Updated	changed	I NA I
on update	l d	(multip	opuacou	enangea	
		le)			
Source	source	source	NA	NA	source
+	+		F	F	+

WHOIS Data of ASN

4.5. Conclusion

As it can be observed, for each object (Organization, Contact Person, Net-num and ASN) there are fields that are unique to only one or a set of RIRs and there are fields that have the same meaning but are referred different for each RIR. In order to construct a single data model for each object a selection of the most common and useful fields was made. That initial selection was the starting point of the previous document of [I-D.ietf-weirds-json-response].

5. DNR Objects Analysis

5.1. 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]

<u>5.2</u>. 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.

5.2.1. 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	Last Transferred	4	3
Date	Date		

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

July 2014

"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 seperately.

5.2.2. WHOIS Data of Contact

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

5.2.2.1. 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	Registrant	42	4
Organization	Organization		
Registrant Country	Country	42	6
Code			
Registrant City	Registrant City	38	4
Registrant Postal	Registrant Postal	37	5
Code	Code		
Registrant	Registrant	32	4
State/Province	State/Province		
Registrant Street	Registrant Street1	31	16
Registrant Country	Registrant Country	19	4
Registrant Phone	Registrant Phone	18	2
Ext.	Ext.	l	
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.

5.2.2.2. 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.

5.2.2.3. 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.

5.2.2.4. Billing Contact

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

Internet-Draft

+ 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	Country	38	4
Code			
Billing Phone	Phone Number	34	6
Billing ID	Billing ID	28	9
Billing City	Billing City	28	4
Billing	Billing	28	5
Organization	Organization		
Billing Postal	Billing Postal	27	4
Code	Code		
Billing	Billing	21	4
State/Province	State/Province		
Billing Street	Billing Street1	19	13
Billing Country	Billing Country	13	5
Billing Phone Ext.	Billing Phone Ext.	10	2
Billing Fax Ext	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.

5.2.3. 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.

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.

5.2.4. 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.

5.3. Other Objects

So called other objects are those data elements that are selfdesigned 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

Туре	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.

5.4. Conclusion

<u>5.4.1</u>. 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.
- 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 nsl.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 195.153.6.27 а 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 Kinadom 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 Aquila, 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.

5.4.2. Data Elements Analysis

Top 10 data elements are as follows:

+	+
Data Element	No. of TLDs
++	110
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.

5.4.3. 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.

5.4.4. 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 <u>section 4.3</u>. 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.

5.5. Limitation

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

- 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.

<u>6</u>. Reference Extension Objects

There are some objects that are included in the existed WHOIS system but not mentioned in the document of [<u>I-D.ietf-weirds-json-response</u>]. This document is intended to give a list of reference extension ogjects for discussion.

<u>6.1</u>. RIR Reference Extension Objects

- o company -- the company name registered by the registrant.
- o maintainer -- authentication information that identifies who can modify the contents of this object.
- o list of resources -- include a list of all the Internet resources assigned to this organization.
- o referral NOC contact -- the Network Operation Center contact.

<u>6.2</u>. DNR Reference Extension Objects

The following objects are selected from the Top 50 other objects in section 5.3 that are supported by over 5 registries. These objects are considered as possible extension objects.

- o zone-c -- The handle of a 'role' object with authority over a zone.
- o maintainer -- authentication information that identifies who can modify the contents of this object.
- Registration URL -- it is usually the website address of a registry.

- o anonymous -- whether the registration information is anonymous or not.
- o hold -- whethe the domain is hold or not.
- nsl-id -- nameserver list ID. 0
- obsoleted -- whether a domain is obsoleted or not. 0
- o Customer Service Contact -- a kind of contact.

7. IANA Considerations

This document does not specify any IANA actions.

8. Security considerations

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

9. Acknowledgements

This document has been reviewed and improved by the Weirds working group. The authors especially thank the following individuals who gave their suggestions and contributions to this document: Guangqing Deng, Frederico A C Neves, Ray Bellis, Edward Shryane and Kaveh Ranjbar.

10. Normative References

- [I-D.ietf-weirds-json-response] Newton, A. and S. Hollenbeck, "JSON Responses for the Registration Data Access Protocol (RDAP)", draft-ietfweirds-json-response-07 (work in progress), April 2014.
- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", <u>BCP 14</u>, <u>RFC 2119</u>, March 1997.
- Alaettinoglu, C., Villamizar, C., Gerich, E., Kessens, D., [RFC2622] Meyer, D., Bates, T., Karrenberg, D., and M. Terpstra, "Routing Policy Specification Language (RPSL)", RFC 2622, June 1999.
- [RFC3912] Daigle, L., "WHOIS Protocol Specification", <u>RFC 3912</u>, September 2004.
- [RFC5730] Hollenbeck, S., "Extensible Provisioning Protocol (EPP)", STD 69, <u>RFC 5730</u>, August 2009.

- [RFC5731] Hollenbeck, S., "Extensible Provisioning Protocol (EPP) Domain Name Mapping", STD 69, <u>RFC 5731</u>, August 2009.
- [RFC5732] Hollenbeck, S., "Extensible Provisioning Protocol (EPP) Host Mapping", STD 69, <u>RFC 5732</u>, August 2009.
- [RFC5733] Hollenbeck, S., "Extensible Provisioning Protocol (EPP) Contact Mapping", STD 69, <u>RFC 5733</u>, August 2009.
- Klensin, J., "Internationalized Domain Names for [RFC5890] 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, <<u>https://docs.google.com/</u> open?id=0B96TtoK8a--MTTRuVUt3UHZMdEk>.

Appendix A. Change Log

Initial -00: Adopted as working group document.

-01:

- * Added Change Log section.
- * Added RIR data objects.
- * Exchanged section 2 and section 3.

-02:

- * Modified some object names in the section of RIR Objects Analysis.
- * Added reference extension objects.

-03:

* Updated to the keep-alive version. Changed the expiry dates and the draft number.

Authors' Addresses

Internet-Draft Registration Data Object Inventory 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 Arturo Servin LACNIC Rambla Mexico 6125 Montevideo, Montevideo 11400 Uruguay Phone: +598-2604-2222 Email: aservin@lacnic.net

July 2014