



Global Alliance
for Genomics & Health

How to light a beacon

Miro Cupak

Senior software engineer, DNASTack

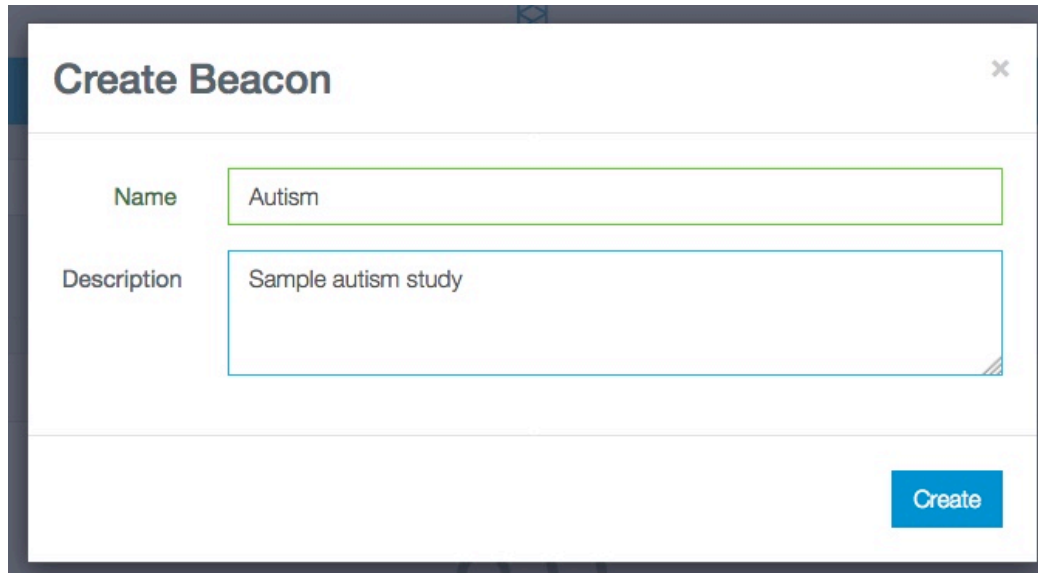
28/02/2017

Options

	Ease of implementation	Flexibility
#1 Managed solutions	***	*
#2 Wrapping existing APIs	**	**
#3 Implementing beacon from scratch	*	***

#1 Managed solutions

- Light a beacon by clicking a button.
- Easiest, fully automated, no development effort needed.
- Data hosted by a 3rd party.
- E.g. DNASTack (<https://dnastack.com/>).



The screenshot shows a web interface titled "Create Beacon" with a close button (X) in the top right corner. It contains two input fields: "Name" with the value "Autism" and "Description" with the value "Sample autism study". A blue "Create" button is located at the bottom right of the form.

Field	Value
Name	Autism
Description	Sample autism study

#1 DNASTack



Global Alliance
for Genomics & Health



Miro Cupak ▾

Autism Sequenci... » Apps » Uploader

Uploader

📁 3 files queued for upload 📁 128 MB

▶ Start upload ⏹ Cancel upload 🗑 Clear list

+ Choose Files... ➡ Import from Partner

HG00121.vcf

VCF ▾ 📁 2 MB

Cancel

⚙ Done ▴

Variant Call Information

Reference

GRCh38 ▴

ERR000025_1.filt.fastq.gz, ERR000025_2.filt.fastq.gz

Unpair

FASTQ ▾ 📁 126 MB

Cancel

⚙ Configure ▾

#1 DNASTack

Access Levels



Public Access



Registered Access



Metadata to include in response

Allele Exists Allele Frequency ×



Controlled Access



Metadata to include in response

Allele Exists Allele Frequency ×
Number of observations × Info ×

Allowed users

Type username...

Variant Datasets

Add Variant Dataset

HG00121.vcf

Remove

#2 API wrappers

- Beacon on top of your existing system (API).
- E.g. GA4GH Core API (<https://github.com/ga4gh/schemas/>).
- There are open-source projects to help you get started.
 - Apache 2.0 license.
- Host yourself or contact us.

#2 Beacon Development Kits (BDK)

- Open-source quickstarts.
- Compliant with the latest beacon specification.
- Proven to work well.
- Available in:
 - Java
 - <https://github.com/mcupak/beacon-java>
 - Python
 - <https://github.com/mcupak/beacon-python>
 - JavaScript
 - <https://github.com/mcupak/beacon-javascript>

```
117  # TODO: override with the details of your beacon
118  @app.route('/beacon-python/', methods=['GET'])
119  def info():
120      return jsonify(Beacon)
```

#2 Java BDK

- Most advanced of the BDKs.
- Provides:
 - Compatible beacon API.
 - Sample beacon implementation.
 - Integration of Beacon data models.
 - Various utilities: output serialization, input normalization, error handling and status code mapping...
 - Test suite.
 - Security (prototype).

#2 BDK adapter system

- Adapter API:
 - <https://github.com/mcupak/beacon-adapter-api>
- Adapter implementations for common data stores/APIs:
 - Variants API
 - <https://github.com/mcupak/beacon-adapter-variants>
 - variant set → beacon dataset
 - Annotation API
 - <https://github.com/mcupak/beacon-adapter-annotations>
 - annotation set → beacon dataset
 - VCF
 - <https://github.com/mcupak/beacon-adapter-vcf>
 - file → beacon dataset
- Easy to implement your own and plug into Java BDK.

#2 Beaconizer

- Extension of Java BDK for more complex *beaconizing*.
- Support for multiple adapters and multiple beacons.
- Open-source: <https://github.com/mcupak/beaconizer>

```
{  
  "name": "beacon_id",  
  "adapterClass": "com.dnastack.beacon.core.adapter.impl.SampleBeaconAdapterImpl",  
  "configValues": [  
    {  
      "name": "sampleName",  
      "value": "33-03-03"  
    }  
  ]  
}
```



```
GET /beacons  
GET /beacons/:id  
GET /beacons/:id/query
```

#3 Implementing beacon

- DIY.
- Most flexibility, control and software development effort.
- Not hard.
 - Beacon API is simple.
 - Many institutions have done this before.
 - Many open-source implementations to help you.

#3 How to start?

- Check out the API specification.
 - <https://github.com/ga4gh/beacon-team/>
 - Very simple and easy to understand.
 - Data model with support for various programming languages.
- Implement 2 REST endpoints returning JSON.
 - Information about the beacon and its datasets.
 - Query for an allele.

#3 Query API

Request:

"Do you have information about this mutation?"

```
{  
  "referenceName": "13",  
  "start": 32936731,  
  "referenceBases": "G",  
  "alternateBases": "C",  
  "assemblyId": "GRCh37"  
}
```

Response:

"yes"/"no"
(+additional info about the mutation)

```
{  
  "exists": true,  
  "datasetAlleleResponses": [  
    {  
      "exists": true,  
      "info": {  
        ...  
      }  
    }  
  ]  
}
```

#3 Status of the Beacon API

- Latest: 0.3
 - <https://github.com/ga4gh/beacon-team/releases/tag/v0.3.0>
 - Improved support for datasets and cross-dataset queries.
 - Modular and extensible.
 - Data versioning.
 - Improved data model.
 - Tooling.
- Future: 1.0
 - Help us define what's needed.
 - <https://github.com/ga4gh/beacon-team/issues>
- WIP: 0.4
 - Complex variants.
 - Improved data use conditions.
 - Documentation.
 - Developer experience.
 - Various minor fixes.

#3 Beacon implementations

- Full list: <https://github.com/ga4gh/beacon-team/wiki/Resources>
- Java
 - <https://github.com/mcupak/beacon-java>
 - <https://github.com/elixirhub/human-data-beacon>
 - <https://github.com/prota/cnag-beacon>
- Python
 - <https://github.com/Geneccloud/simplebeacon>
 - <https://github.com/maximilianh/ucscBeacon>
 - <https://github.com/kozbo/Beacon-on-GA4GH-API>
- JavaScript
 - <https://github.com/ElementoLab/beacon>
- Go
 - <https://github.com/googlegenomics/beacon-go>



#3 Beacon implementations

- Full list: <https://github.com/ga4gh/beacon-team/wiki/Resources>

- Java

- <https://github.com/mcupak/beacon-java>
- <https://github.com/elixirhub/human-data-beacon>
- <https://github.com/prota/cnag-beacon>

- Python

- <https://github.com/Genecloud/simplebeacon>
- <https://github.com/maximilianh/ucscBeacon>
- <https://github.com/kozbo/Beacon-on-GA4GH-API>

- JavaScript

- <https://github.com/ElementoLab/beacon>

- Go

- <https://github.com/googlegenomics/beacon-go>



#3 Other tools: Compliance suite

- Test beacon compatibility with 0.3 specification.
- Groovy-based prototype.
- Open source: <https://github.com/mcupak/beacon-compliance>

Test results

1 suite, 8 failed tests

All suites

Surefire suite

Info

- [unset file name]
- 1 test
- 0 groups
- Times
- Reporter output
- Ignored methods
- Chronological view

Results

- 13 methods, 8 failed, 5 passed
- Failed methods (hide)
 - ✗ testGetAllele
 - ✗ testGetAlleleWithDataSets
 - ✗ testGetAlleleWithMissingOptionalParams
 - ✗ testGetAlleleWithMissingRequiredParams
 - ✗ testGetAlleleWithoutDatasets
 - ✗ testGetBeacon
 - ✗ testPostAllele
 - ✗ testPostInvalidRequest
- Passed methods (show)

com.dnastack.beacon.compliance.ComplianceSuiteTest

```
testGetAllele
  java.lang.NullPointerException: Cannot invoke method getSampleAlleleRequests
    at org.codehaus.groovy.runtime.NullObject.invokeMethod(NullObject.java:41)
    at org.codehaus.groovy.runtime.callsite.PogoMetaClassSite.call(PogoMetaClassSite.java:46)
    at org.codehaus.groovy.runtime.callsite.CallSiteArray.defaultCall(CallSiteArray.java:49)
    at org.codehaus.groovy.runtime.callsite.NullCallSite.call(NullCallSite.java:43)
    at org.codehaus.groovy.runtime.callsite.CallSiteArray.defaultCall(CallSiteArray.java:49)
    at org.codehaus.groovy.runtime.callsite.AbstractCallSite.call(AbstractCallSite.java:127)
    at org.codehaus.groovy.runtime.callsite.AbstractCallSite.call(AbstractCallSite.java:133)
    at com.dnastack.beacon.compliance.ComplianceSuiteTest.testGetAllele(...)
    ... Removed 28 stack frames

testGetAlleleWithDataSets
  java.lang.NullPointerException: Cannot invoke method getSampleAlleleRequests
    at org.codehaus.groovy.runtime.NullObject.invokeMethod(NullObject.java:41)
    at org.codehaus.groovy.runtime.callsite.PogoMetaClassSite.call(PogoMetaClassSite.java:46)
    at org.codehaus.groovy.runtime.callsite.CallSiteArray.defaultCall(CallSiteArray.java:49)
    at org.codehaus.groovy.runtime.callsite.NullCallSite.call(NullCallSite.java:43)
    at org.codehaus.groovy.runtime.callsite.CallSiteArray.defaultCall(CallSiteArray.java:49)
    at org.codehaus.groovy.runtime.callsite.AbstractCallSite.call(AbstractCallSite.java:127)
    at org.codehaus.groovy.runtime.callsite.AbstractCallSite.call(AbstractCallSite.java:133)
    at com.dnastack.beacon.compliance.ComplianceSuiteTest.getSampleAlleleRequests(...)
    at com.dnastack.beacon.compliance.ComplianceSuiteTest.testGetAlleleWithDataSets(...)
    ... Removed 28 stack frames

testGetAlleleWithMissingOptionalParams
  java.lang.NullPointerException: Cannot invoke method getSampleAlleleRequests
    at org.codehaus.groovy.runtime.NullObject.invokeMethod(NullObject.java:41)
```

#3 Other tools: Utilities

- Coordinate conversion library.
 - Wraps Liftover and HGVS.
 - Extensible, with built-in converters for standard assemblies.
 - Open-source: <https://github.com/mcupak/genome-coordinates-converter>

```
LiftOver intervalLiftOver = UCSCliftOver.hg19ToHg38();  
Interval newInterval = intervalLiftOver.liftOver("chr1",743267,743268);
```

```
String hgvs = "NM_182763.2:c.688+403C>T"  
GenomeInterval interval = HGVSConverter.hgvsToGenomic(hgvs);
```

#3 Other tools: Client and CLI

- Client library.
 - Query beacons from Java.
 - Integration with Beacon 0.3 schema.
 - Open-source: <https://github.com/mcupak/beacon-client>

```
BeaconNetworkClient bnc = new BeaconNetworkClientImpl();  
bnc.getResponse(CHR1, 1L, A, HG19, "brca-exchange");
```

- CLI.
 - Query beacons from your shell.
 - Open-source: <https://github.com/mcupak/beacon-cli>

```
$ java -jar beacon-cli.jar -u http://localhost:5000/test/ response \  
-a GRCh37 -ab A -r 1 -rb T -s 1000
```

```
{  
  "beaconId": "test",  
  "exists": true,  
  "alleleRequest": ...  
}
```

Beacon Network



Global Alliance
for Genomics & Health

[Search Beacons](#)

Search [all beacons](#) for allele

GRCh37 ▾

13 : 32936732 G > C

Search

Response **All** **None**

☒ Found 14

☒ Not Found 37

☐ Not Applicable 14



Organization **All** **None**

☒ AMPLab, UC Berkel...

☒ BGI

☒ BioReference Labor...

☒ Brazilian Initiative o...

☒ BRCA Exchange

☒ Broad Institute

☒ Centre for Genomic...

☒ Centro Nacional de...

☒ Curoverse

☒ EMBL European Bio...

☒ Global Alliance for ...

☒ Google

☒ Institute for System...

☒ Instituto Nacional d...



AMPLab - 1000 Genomes Project

Hosted by AMPLab, UC Berkeley

Not Found



Biobase - HGMD

Hosted by University of California, Santa Cruz

Found



BioReference

Hosted by BioReference Laboratories

Not Found



BRCA Exchange

Hosted by BRCA Exchange

[Hide Metadata](#)

Found

Clinical significance citations [PMID:23108138](#)

HGVS cDNA c.7878G>C

Condition ID type OMIM

Date last evaluated 8/10/15

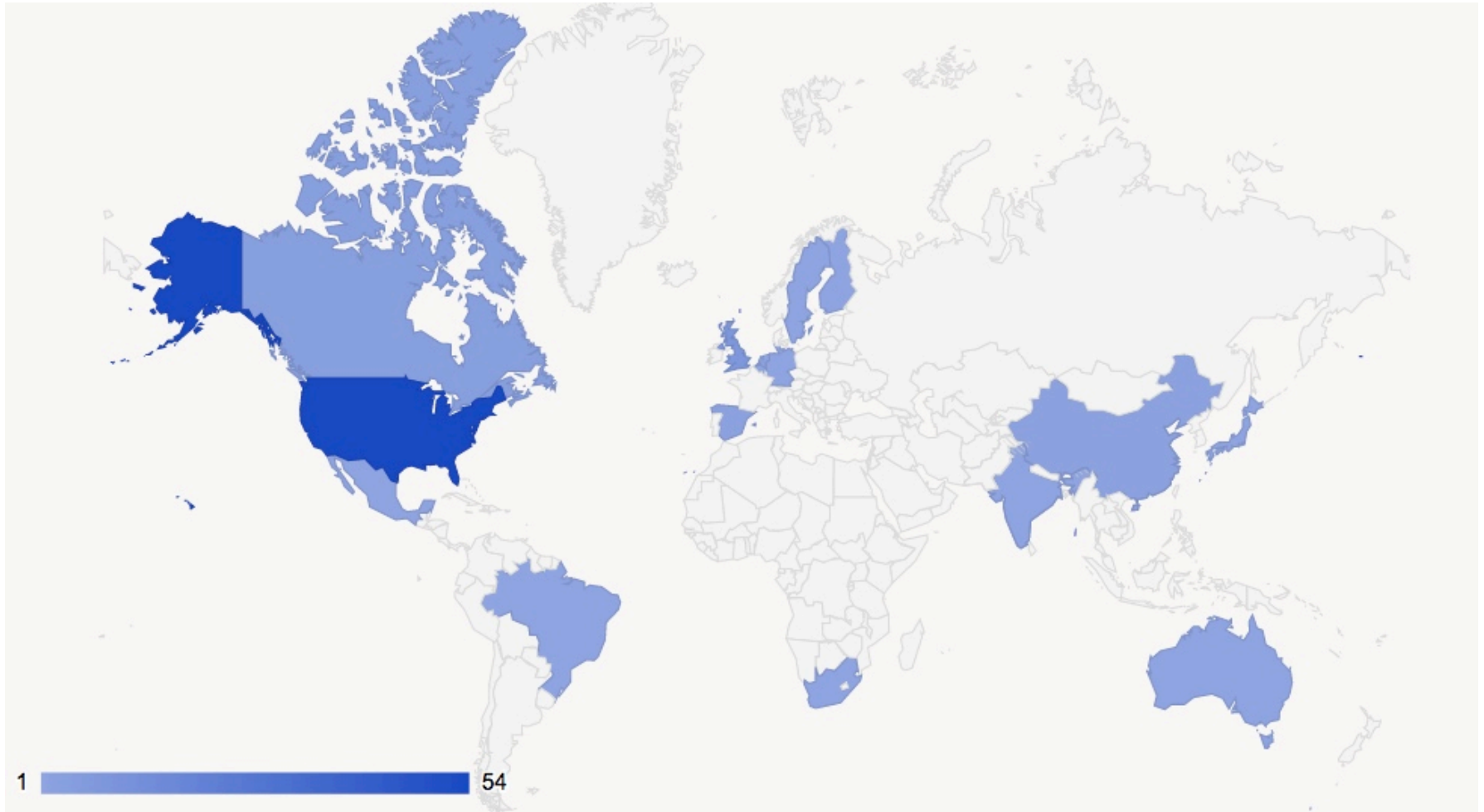
<https://beacon-network.org/>

Beacon Network

- Tooling:
 - REST API: <https://beacon-network.org/api/>
 - Client library: <https://github.com/mcupak/beacon-network-client>
 - CLI: <https://github.com/mcupak/beacon-network-cli>
 - Widget: <https://beacon-network.org/#/developers/embed>
- Registration:
 - Automated for managed solutions.
 - Manual if you implement yourself.
 - Email  API.

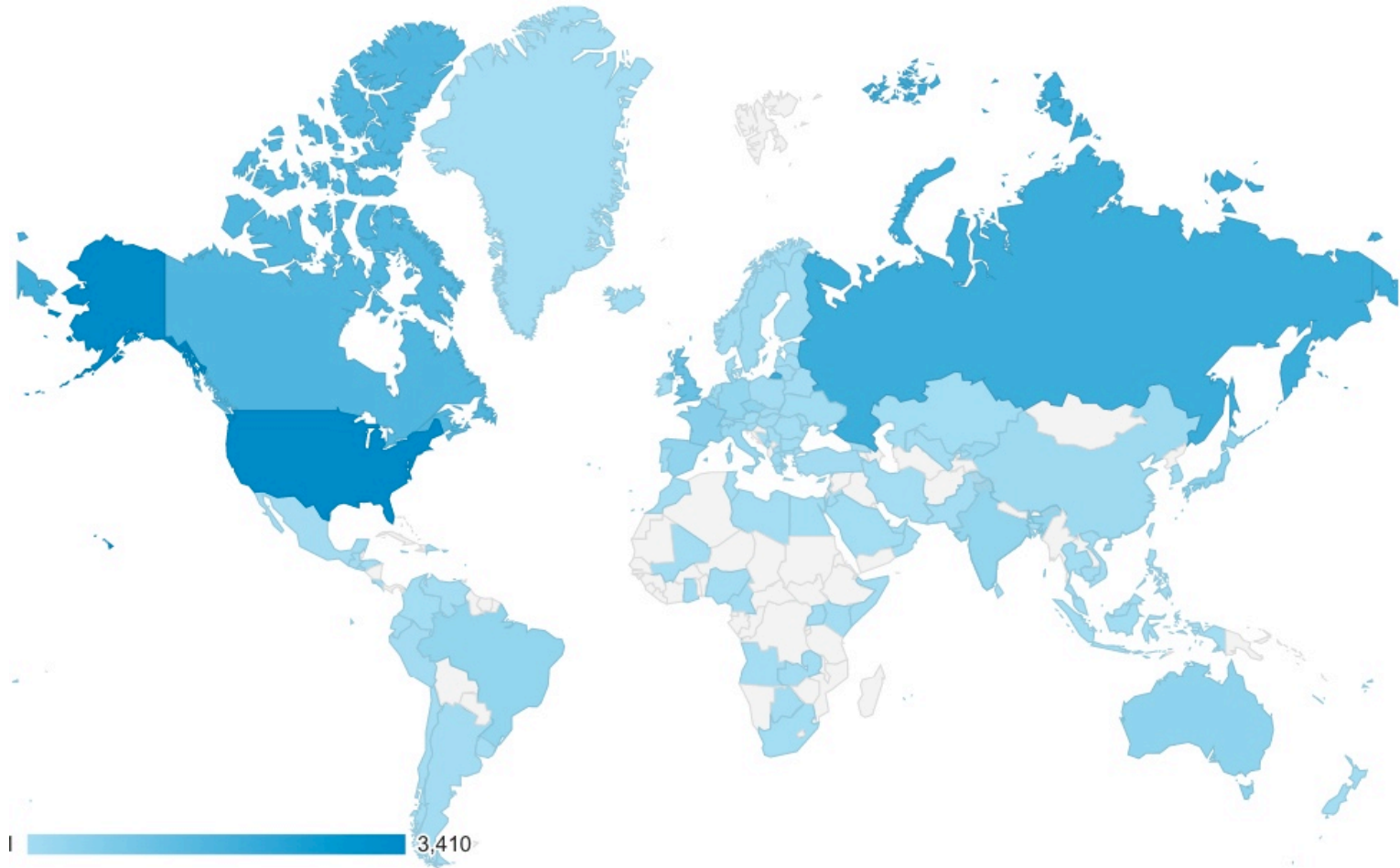
Adoption

- **38** organizations, **70** beacons, **17** countries, **6** continents.



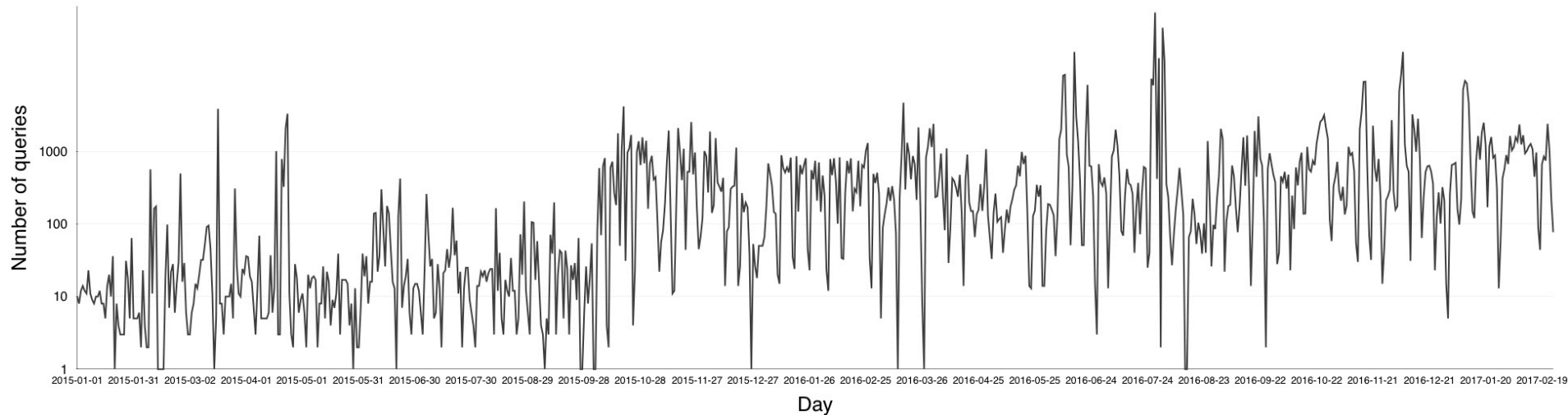
Usage

- **8.5K+** users in **118** countries.





Usage

• **600K+** queries received → **4M+** queries dispatched.






Contact

Beacon

-  <https://github.com/ga4gh/beacon-team/issues/>
-  <https://beacon-team-slackin.herokuapp.com/>
- Mailing list and bi-weekly conference calls.

Beacon Network

-  goo.gl/PsMLhr
-  info@beacon-network.org
-  @BeaconOfBeacons

Me

-  miro@dnastack.com
-  @mirocupak

Questions?