Using macOS Spotlight and Osquery to Prevent Data Breaches

KOLIDE

Fritz Ifert-Miller

UX Designer @ Kolide

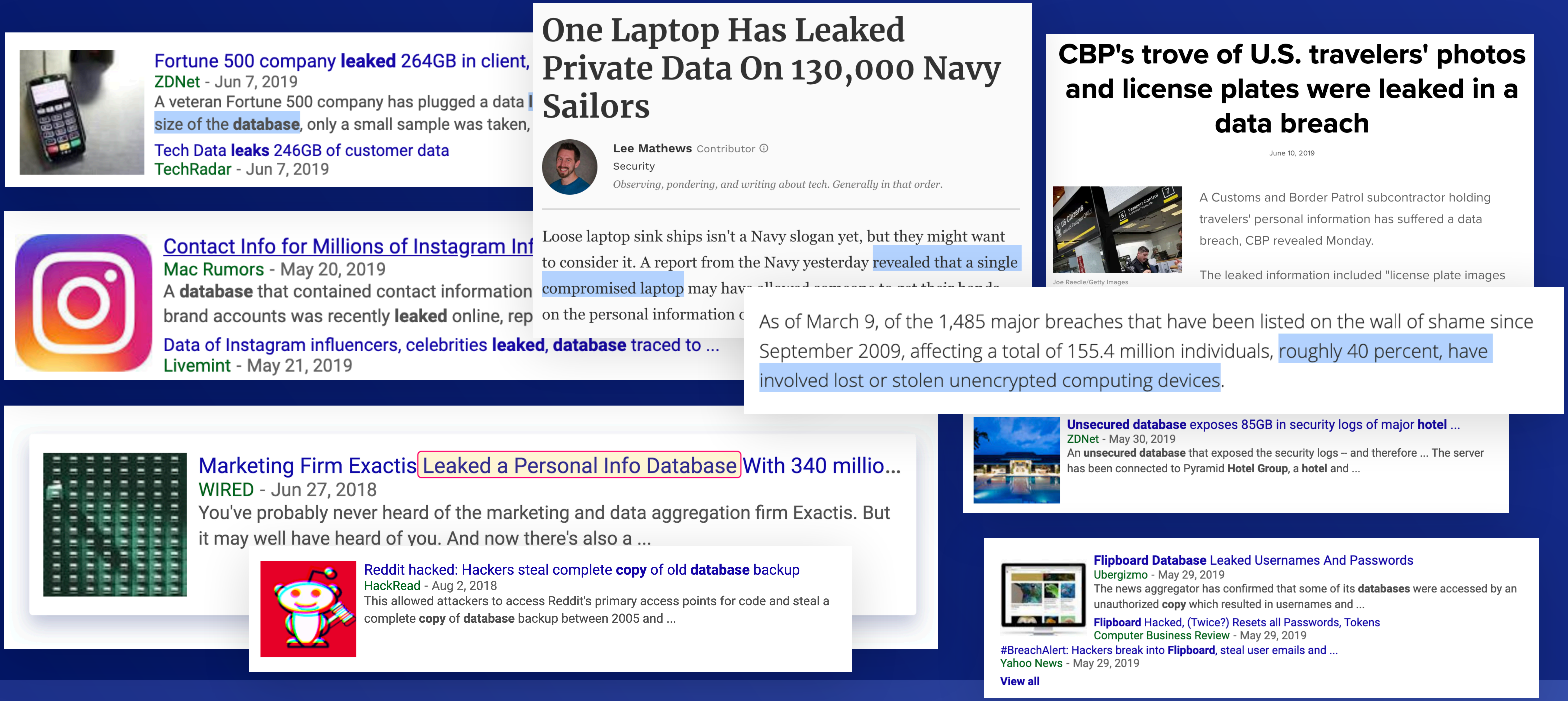
✉ **fritz** @ kolide.com

github.com / **fritzx6**

fritz @ osquery Slack

“There’s no such thing as bad publicity”

~ Exploitative racist jerk: P.T. Barnum



Fortune 500 company **leaked** 264GB in client,
ZDNet - Jun 7, 2019
A veteran Fortune 500 company has plugged a data l
size of the **database**, only a small sample was taken,
Tech Data **leaks** 246GB of customer data
TechRadar - Jun 7, 2019

One Laptop Has Leaked Private Data On 130,000 Navy Sailors

Lee Mathews Contributor
Security
Observing, pondering, and writing about tech. Generally in that order.

Loose laptop sink ships isn't a Navy slogan yet, but they might want to consider it. A report from the Navy yesterday revealed that a single compromised laptop may hav
on the personal information c

CBP's trove of U.S. travelers' photos and license plates were leaked in a data breach

June 10, 2019

Joe Raedle/Getty Images

A Customs and Border Patrol subcontractor holding travelers' personal information has suffered a data breach, CBP revealed Monday.

The leaked information included "license plate images

Contact Info for Millions of Instagram Inf
Mac Rumors - May 20, 2019
A **database** that contained contact information
brand accounts was recently **leaked** online, rep
Data of Instagram influencers, celebrities **leaked**, **database** traced to ...
Livemint - May 21, 2019

As of March 9, of the 1,485 major breaches that have been listed on the wall of shame since September 2009, affecting a total of 155.4 million individuals, roughly 40 percent, have involved lost or stolen unencrypted computing devices.

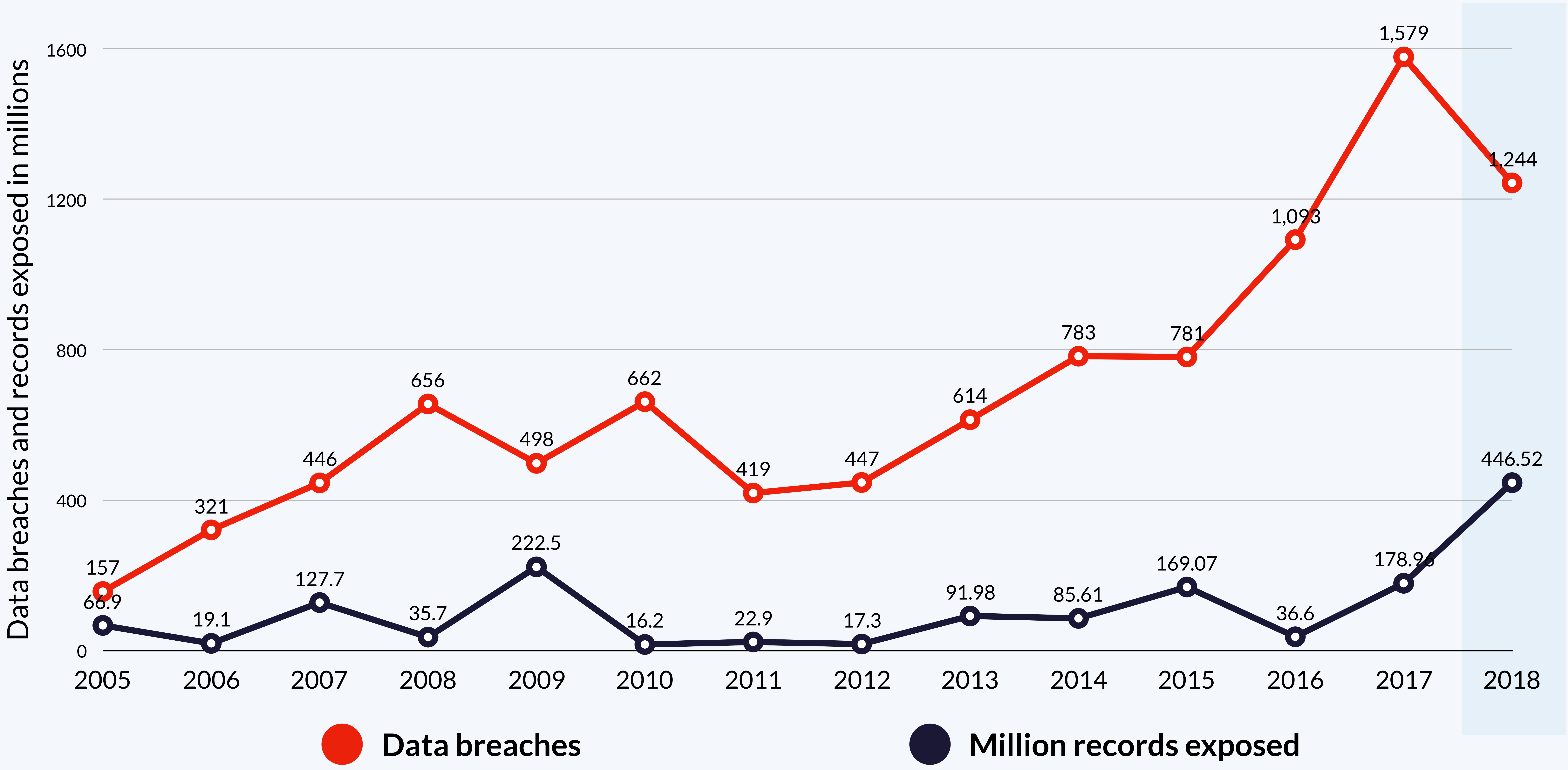
Unsecured database exposes 85GB in security logs of major **hotel** ...
ZDNet - May 30, 2019
An **unsecured database** that exposed the security logs -- and therefore ... The server has been connected to Pyramid **Hotel Group**, a **hotel** and ...

Marketing Firm Exactis Leaked a Personal Info Database With 340 millio...
WIRED - Jun 27, 2018
You've probably never heard of the marketing and data aggregation firm Exactis. But it may well have heard of you. And now there's also a ...

Reddit hacked: Hackers steal complete **copy** of old **database** backup
HackRead - Aug 2, 2018
This allowed attackers to access Reddit's primary access points for code and steal a complete **copy** of **database** backup between 2005 and ...

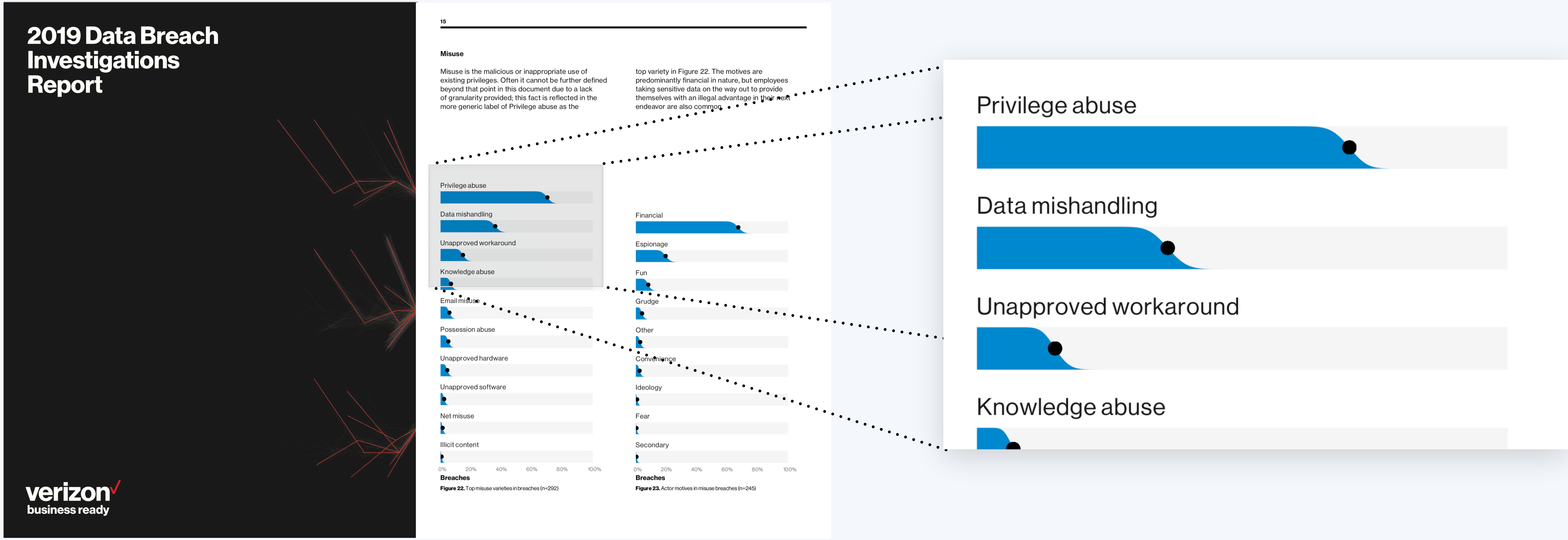
Flipboard Database Leaked Usernames And Passwords
Ubergizmo - May 29, 2019
The news aggregator has confirmed that some of its **databases** were accessed by an unauthorized **copy** which resulted in usernames and ...
Flipboard Hacked, (Twice?) Resets all Passwords, Tokens
Computer Business Review - May 29, 2019
#BreachAlert: Hackers break into **Flipboard**, steal user emails and ...
Yahoo News - May 29, 2019
View all

Data Exfiltration in Numbers



https://www.statista.com/statistics/273550/data-breaches-recorded-in-the-united-states-by-number-of-breaches-and-records-exposed/

How do databases get compromised / breached?



2019 Data Breach Investigations Report

verizon✓
business ready

15

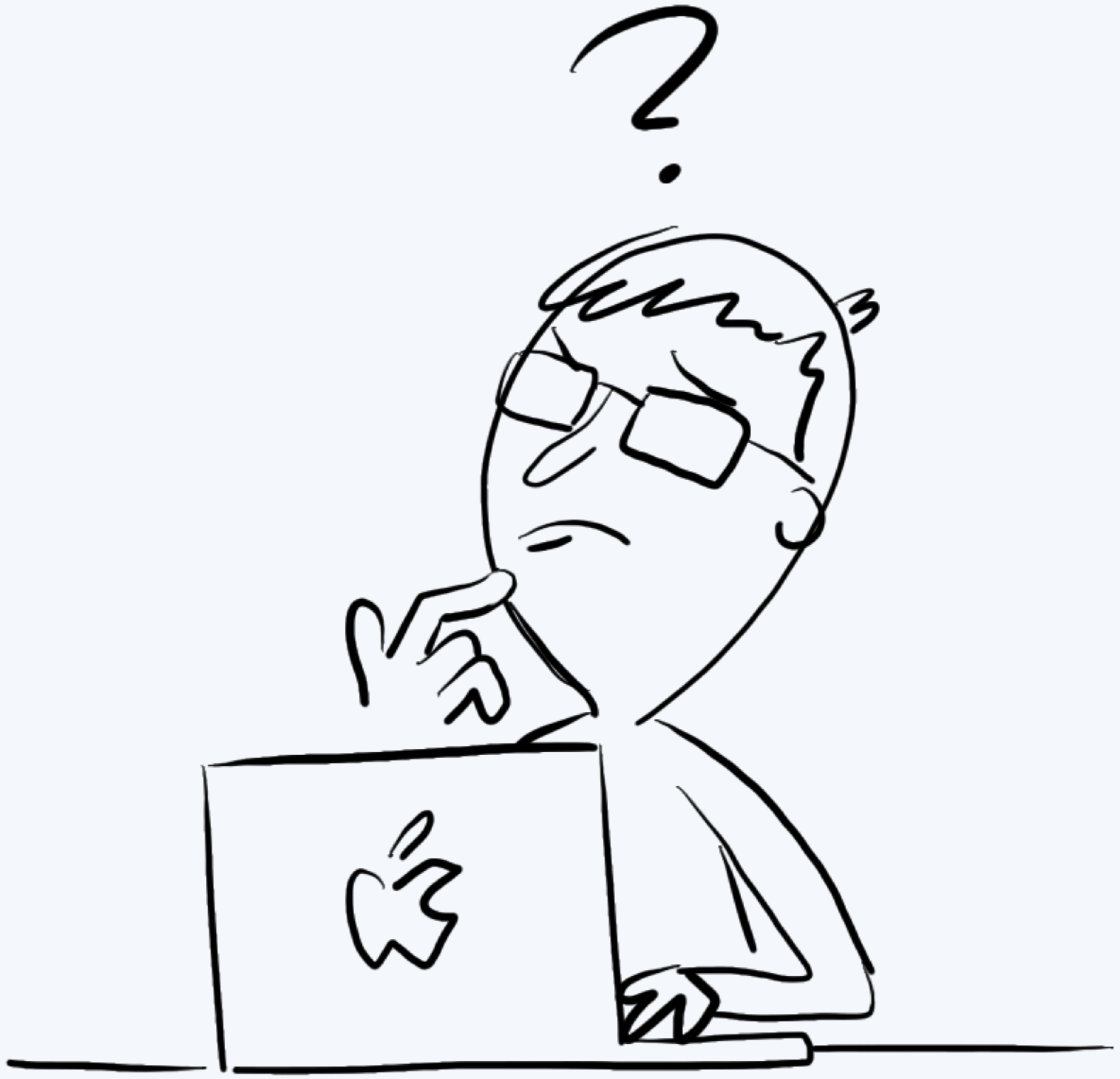
Misuse

Misuse is the malicious or inappropriate use of existing privileges. Often it cannot be further defined beyond that point in this document due to a lack of granularity provided; this fact is reflected in the more generic label of Privilege abuse as the top variety in Figure 22. The motives are predominantly financial in nature, but employees taking sensitive data on the way out to provide themselves with an illegal advantage in their next endeavor are also common.

Figure 22. Top misuse varieties in breaches (n=292)

Figure 23. Actor motives in misuse breaches (n=245)

How do local copies of your production database end up on devices?



"Oh No! I can't reproduce this customer problem in development? What can I do?"

"I Know! I'll download A COPY of the production database and Test with that!"

Making a copy isn't evil, but forgetting to delete it may hurt your company

What are characteristics of database copies and data exports?

Large File Size

Filetype like .db or .sql .gzip

Downloaded from a consistent source

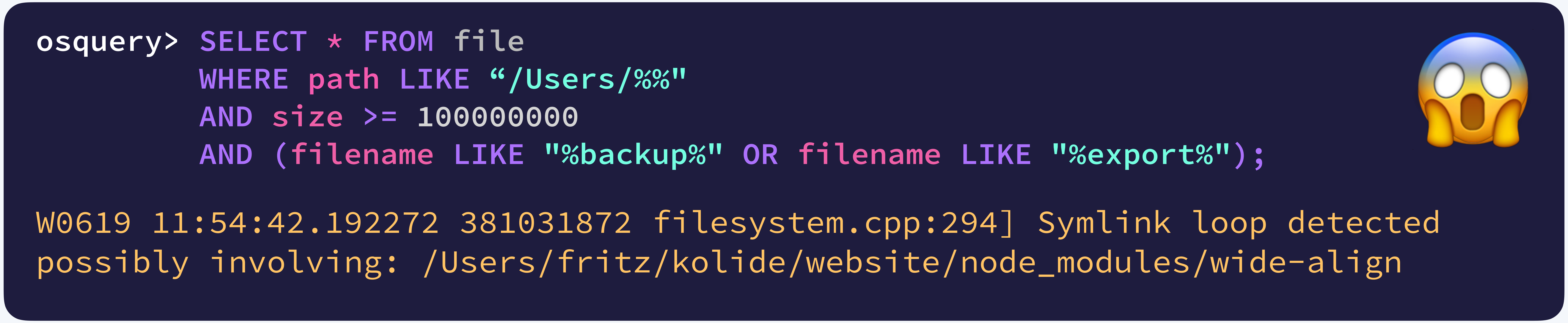
Named ‘backup’… I hope?

Contains customer data

Finding files with the file table

(Don't do this)

The file table in osquery is amazing, but it has limits

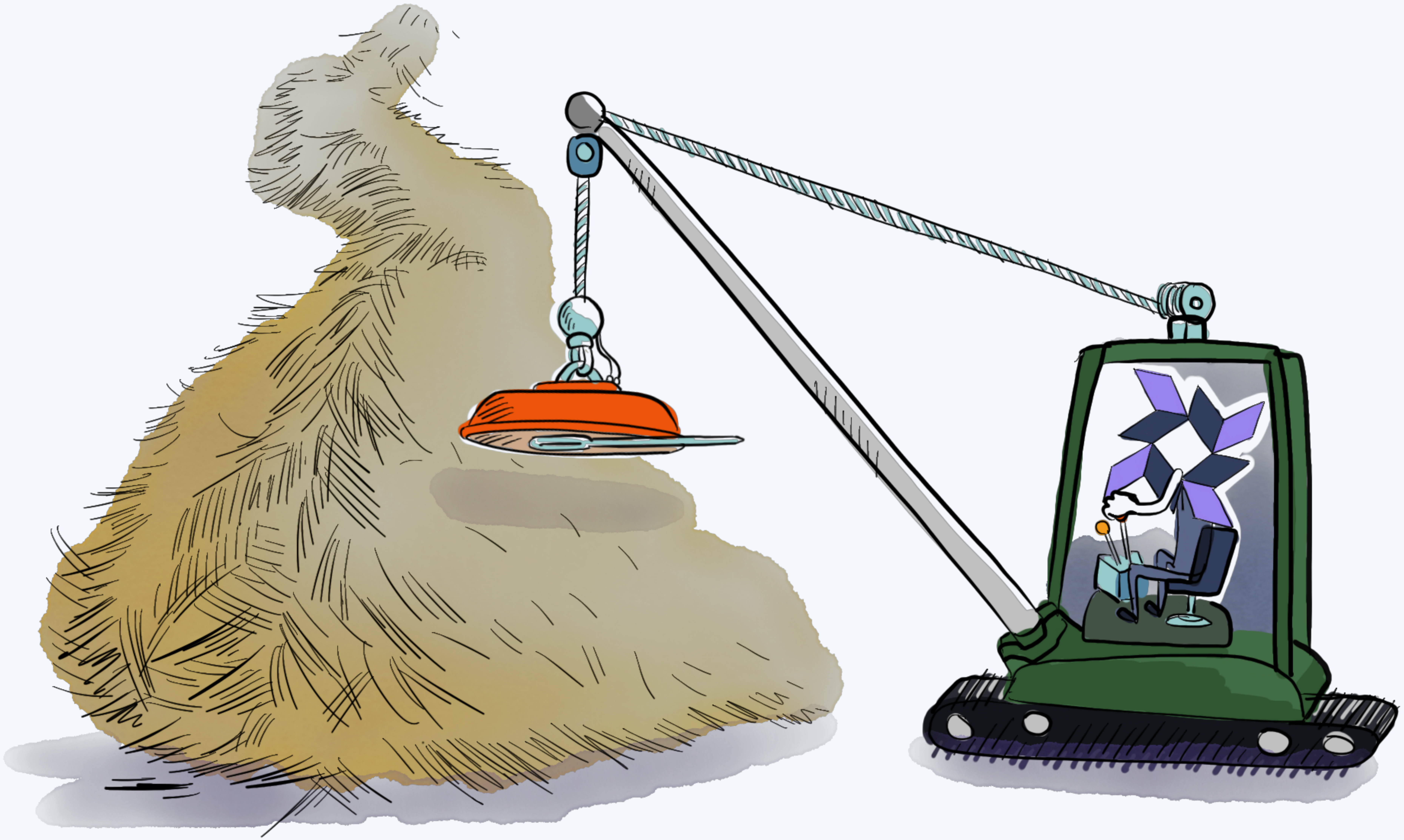


```
osquery> SELECT * FROM file
         WHERE path LIKE “/Users/%%"
         AND size >= 100000000
         AND (filename LIKE "%backup%" OR filename LIKE "%export%");

W0619 11:54:42.192272 381031872 filesystem.cpp:294] Symlink loop detected
possibly involving: /Users/fritz/kolide/website/node_modules/wide-align
```

- You generally need to know where a file is located ahead of time
- Wildcards permitted, but broadly scoped recursive searching generally sucks (symlinks = sadness)
- The file table's introspection capabilities are ineffectually shallow

There must be another way!

mdfinding a needle in a haystack



KOLIDE

What is mdfind?

- Powers the macOS Spotlight omni-search feature
- Introduced in OS X 10.4 Tiger (2005)
- A highly performant, fully indexed Search (across 200+ metadata attributes)
- Supports boolean operators and other basic functions
- Indexes file text contents of many common filetypes

How can we interact with mdfind?

Spotlight *(GUI)*

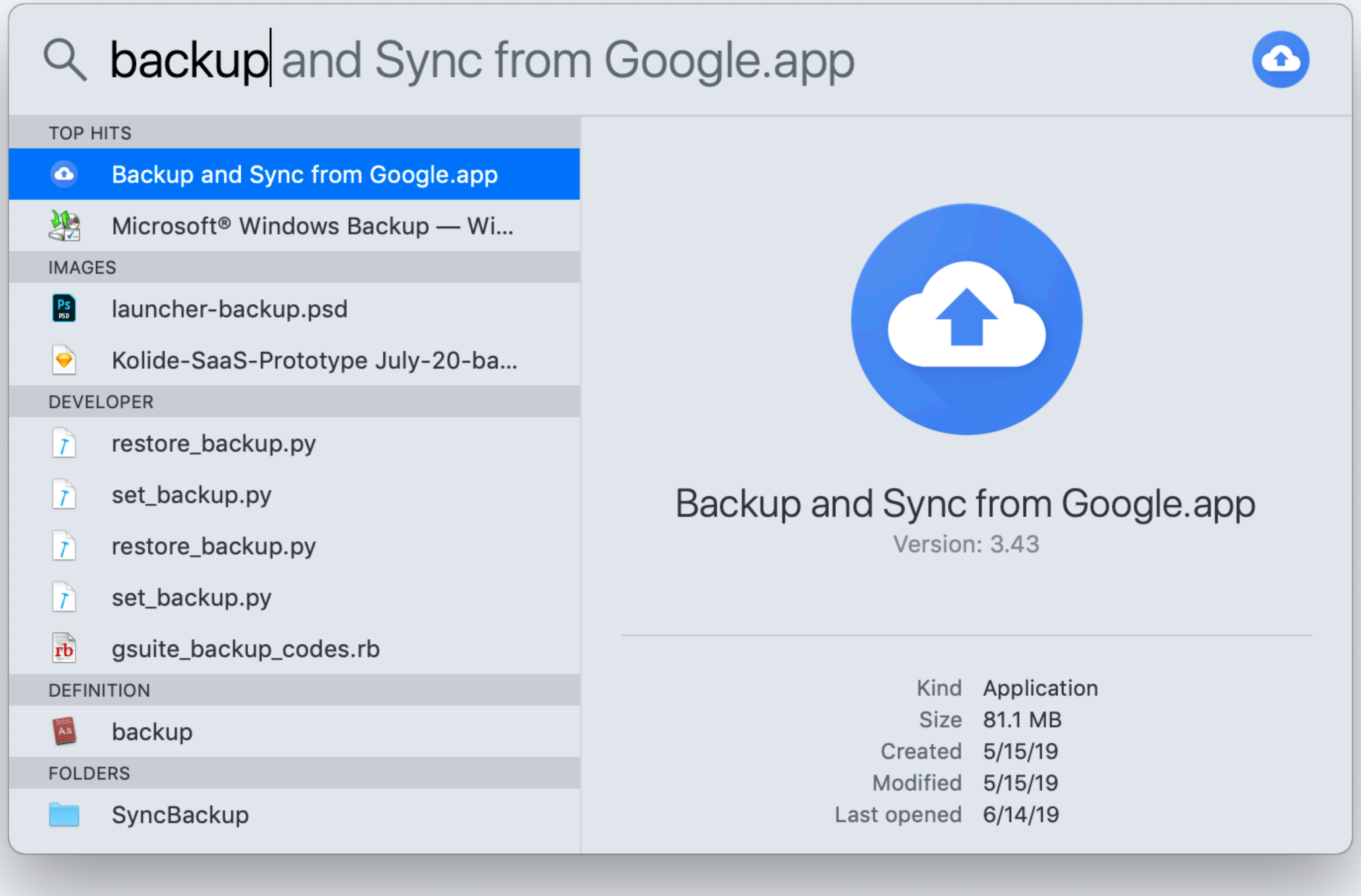
mdfind *(CLI)*

mdls *(CLI)*

mdutil *(CLI)*

mdimport *(CLI)*

osquery mdfind table



backup and Sync from Google.app

TOP HITS
- Backup and Sync from Google.app
- Microsoft® Windows Backup — Wi...

IMAGES
- launcher-backup.psd
- Kolide-SaaS-Prototype July-20-ba...

DEVELOPER
- restore_backup.py
- set_backup.py
- restore_backup.py
- set_backup.py
- gsuite_backup_codes.rb

DEFINITION
- backup

FOLDERS
- SyncBackup

Backup and Sync from Google.app
Version: 3.43

Kind	Application
Size	81.1 MB
Created	5/15/19
Modified	5/15/19
Last opened	6/14/19

How can we interact with mdfind?

Spotlight *(GUI)*

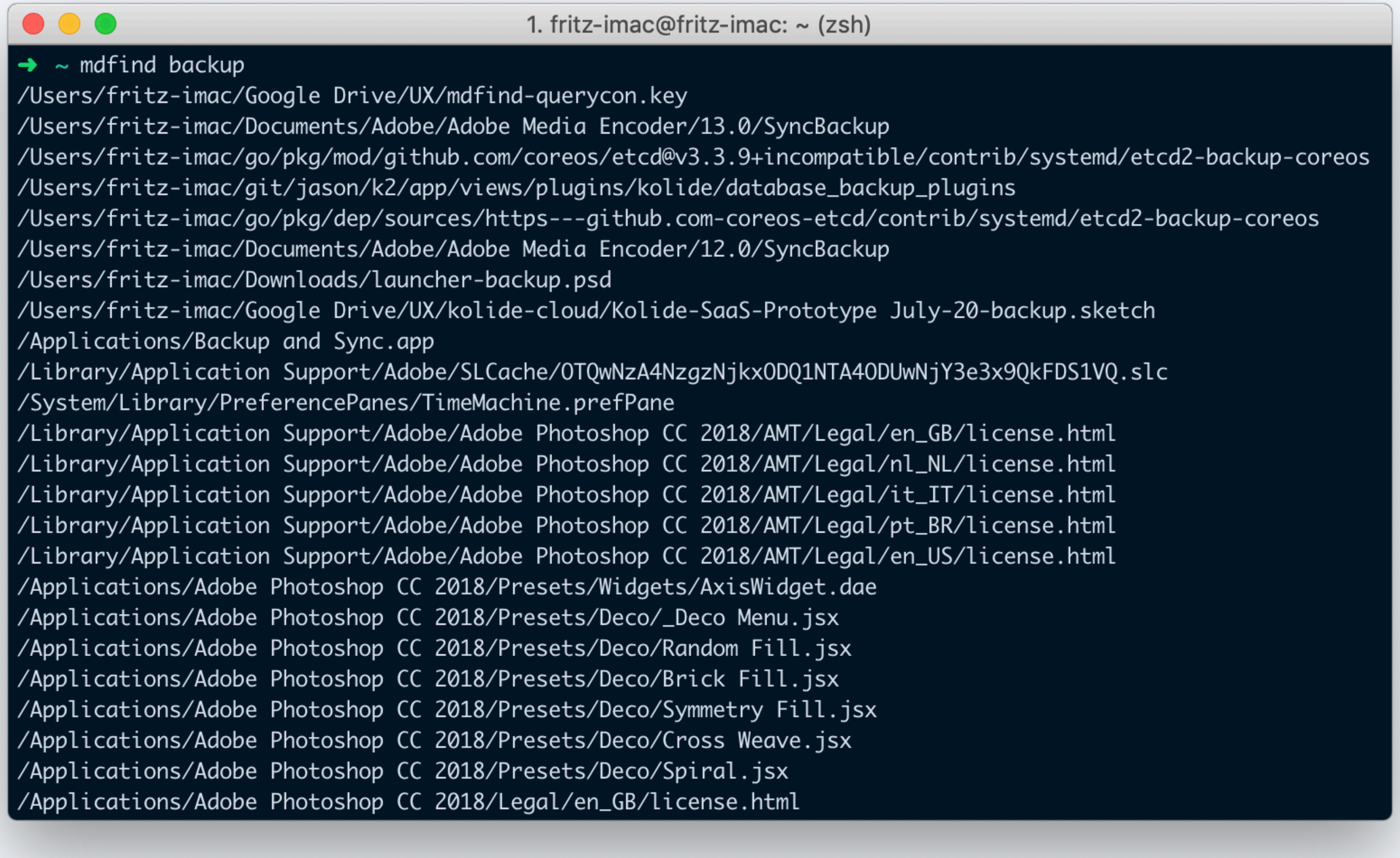
mdfind *(CLI)*

mdls *(CLI)*

mdutil *(CLI)*

mdimport *(CLI)*

osquery mdfind table



1. fritz-imac@fritz-imac: ~ (zsh)

```
➜ ~ mdfind backup
/Users/fritz-imac/Google Drive/UX/mdfind-querycon.key
/Users/fritz-imac/Documents/Adobe/Adobe Media Encoder/13.0/SyncBackup
/Users/fritz-imac/go/pkg/mod/github.com/coreos/etcd@v3.3.9+incompatible/contrib/systemd/etcd2-backup-coreos
/Users/fritz-imac/git/jason/k2/app/views/plugins/kolide/database_backup_plugins
/Users/fritz-imac/go/pkg/dep/sources/https---github.com-coreos-etcd/contrib/systemd/etcd2-backup-coreos
/Users/fritz-imac/Documents/Adobe/Adobe Media Encoder/12.0/SyncBackup
/Users/fritz-imac/Downloads/launcher-backup.psd
/Users/fritz-imac/Google Drive/UX/kolide-cloud/Kolide-SaaS-Prototype July-20-backup.sketch
/Applications/Backup and Sync.app
/Library/Application Support/Adobe/SLCache/OTQwNzA4NzgzNjkxODQ1NTA4ODUwNjY3e3x9QkFDS1VQ.slc
/System/Library/PreferencePanes/TimeMachine.prefPane
/Library/Application Support/Adobe/Adobe Photoshop CC 2018/AMT/Legal/en_GB/license.html
/Library/Application Support/Adobe/Adobe Photoshop CC 2018/AMT/Legal/nl_NL/license.html
/Library/Application Support/Adobe/Adobe Photoshop CC 2018/AMT/Legal/it_IT/license.html
/Library/Application Support/Adobe/Adobe Photoshop CC 2018/AMT/Legal/pt_BR/license.html
/Library/Application Support/Adobe/Adobe Photoshop CC 2018/AMT/Legal/en_US/license.html
/Applications/Adobe Photoshop CC 2018/Presets/Widgets/AxisWidget.dae
/Applications/Adobe Photoshop CC 2018/Presets/Deco/_Deco Menu.jsx
/Applications/Adobe Photoshop CC 2018/Presets/Deco/Random Fill.jsx
/Applications/Adobe Photoshop CC 2018/Presets/Deco/Brick Fill.jsx
/Applications/Adobe Photoshop CC 2018/Presets/Deco/Symmetry Fill.jsx
/Applications/Adobe Photoshop CC 2018/Presets/Deco/Cross Weave.jsx
/Applications/Adobe Photoshop CC 2018/Presets/Deco/Spiral.jsx
/Applications/Adobe Photoshop CC 2018/Legal/en_GB/license.html
```

How can we interact with mdfind?

Spotlight *(GUI)*

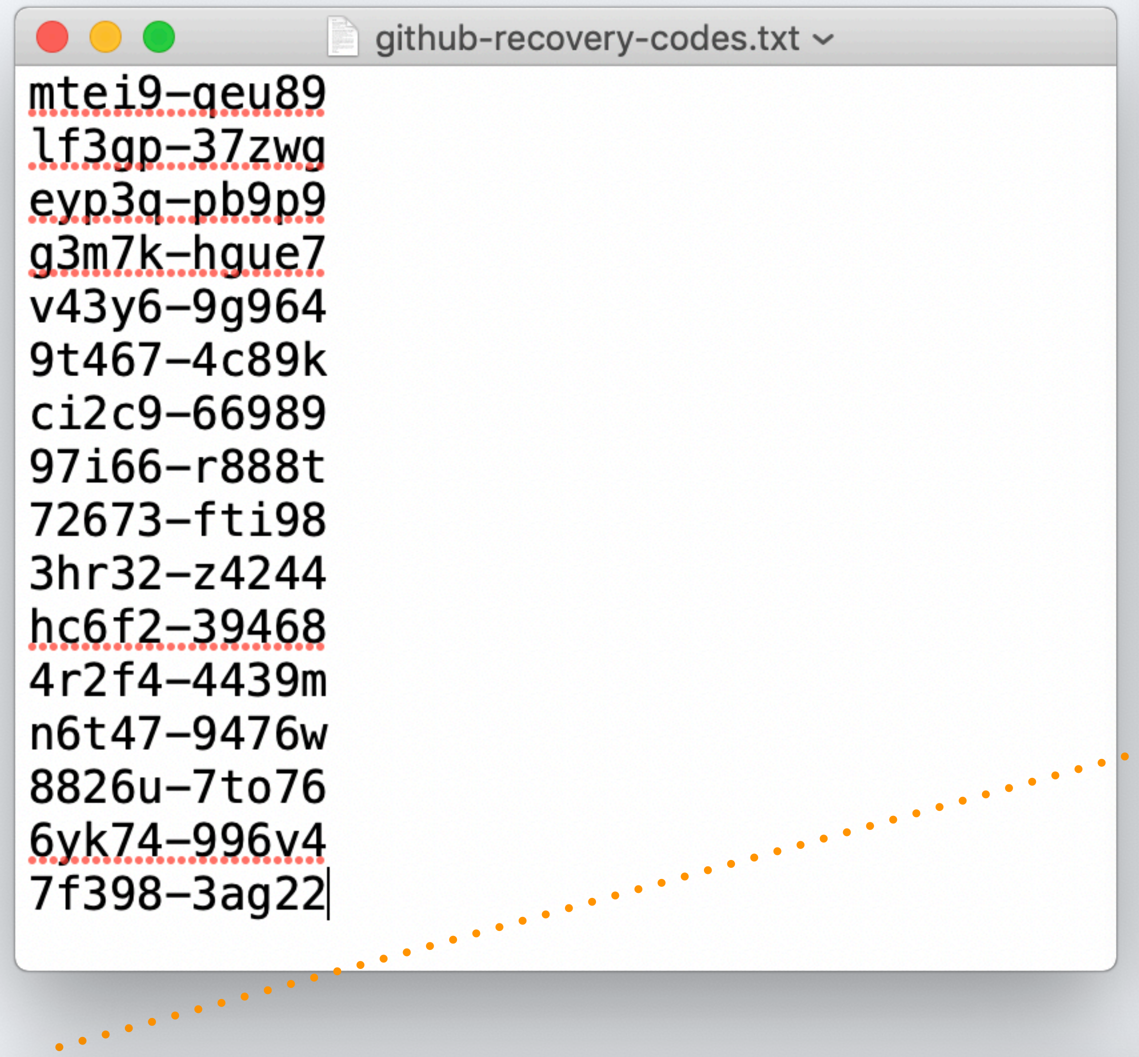
mdfind *(CLI)*

mdls *(CLI)*

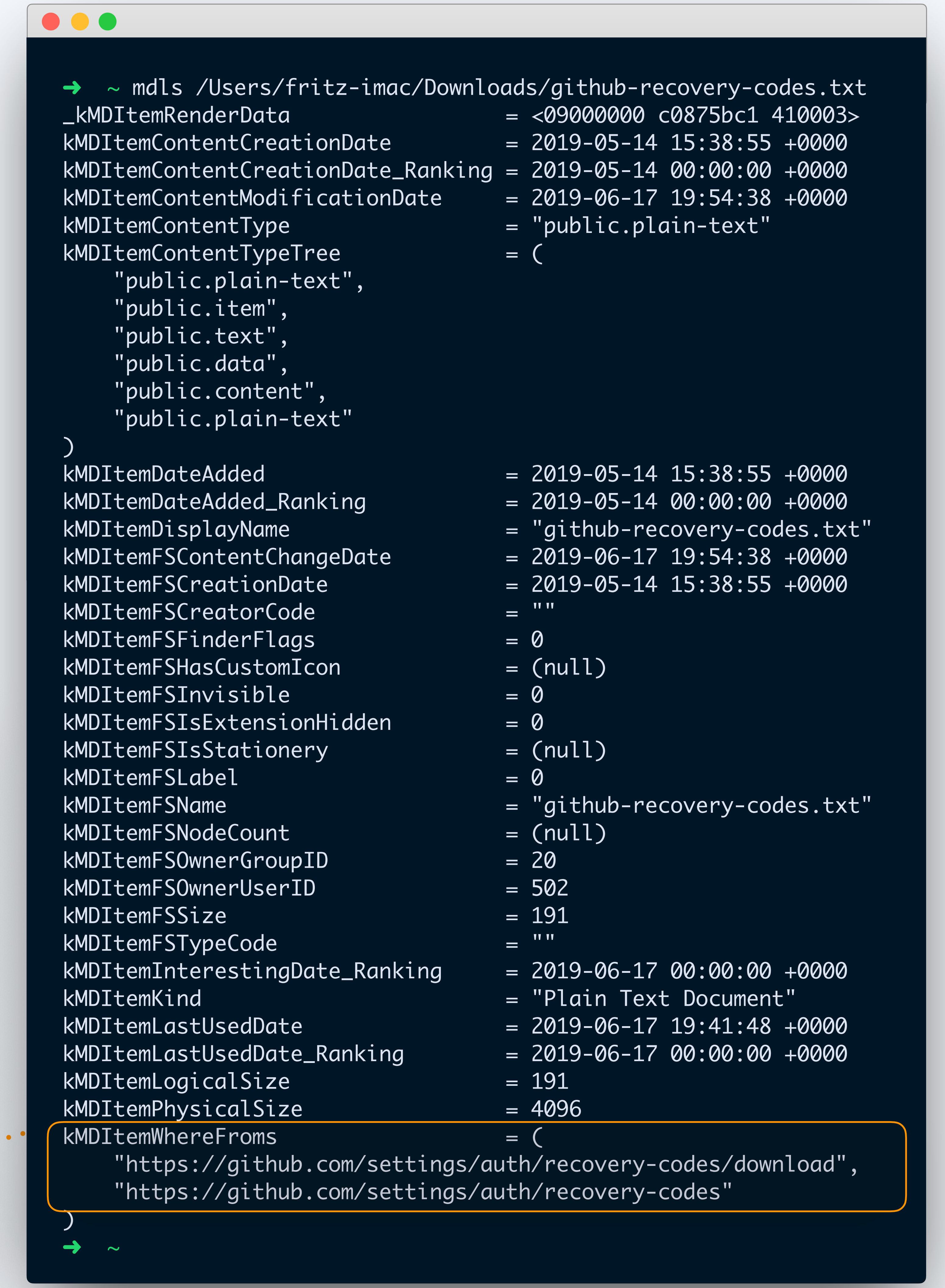
mdutil *(CLI)*

mdimport *(CLI)*

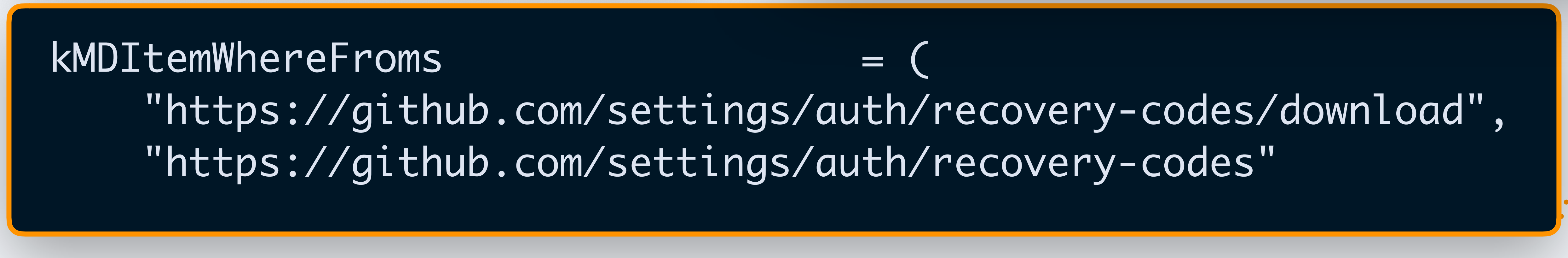
osquery mdfind table



```
mtei9-qeu89
lf3gp-37zwg
eyp3q-pb9p9
g3m7k-hgue7
v43y6-9g964
9t467-4c89k
ci2c9-66989
97i66-r888t
72673-fti98
3hr32-z4244
hc6f2-39468
4r2f4-4439m
n6t47-9476w
8826u-7to76
6yk74-996v4
7f398-3ag22
```



```
➜  ~ mdls /Users/fritz-imac/Downloads/github-recovery-codes.txt
_kMDItemRenderData                 = <09000000 c0875bc1 410003>
kMDItemContentCreationDate         = 2019-05-14 15:38:55 +0000
kMDItemContentCreationDate_Ranking = 2019-05-14 00:00:00 +0000
kMDItemContentModificationDate     = 2019-06-17 19:54:38 +0000
kMDItemContentType                 = "public.plain-text"
kMDItemContentTypeTree             = (
    "public.plain-text",
    "public.item",
    "public.text",
    "public.data",
    "public.content",
    "public.plain-text"
)
kMDItemDateAdded                   = 2019-05-14 15:38:55 +0000
kMDItemDateAdded_Ranking           = 2019-05-14 00:00:00 +0000
kMDItemDisplayName                 = "github-recovery-codes.txt"
kMDItemFSContentChangeDate         = 2019-06-17 19:54:38 +0000
kMDItemFSCreationDate              = 2019-05-14 15:38:55 +0000
kMDItemFSCreatorCode               = ""
kMDItemFSFinderFlags               = 0
kMDItemFSHasCustomIcon             = (null)
kMDItemFSInvisible                 = 0
kMDItemFSIsExtensionHidden         = 0
kMDItemFSIsStationery              = (null)
kMDItemFSLabel                     = 0
kMDItemFSName                      = "github-recovery-codes.txt"
kMDItemFSNodeCount                 = (null)
kMDItemFSOwnerGroupID              = 20
kMDItemFSOwnerUserID               = 502
kMDItemFSSize                      = 191
kMDItemFSTypeCode                  = ""
kMDItemInterestingDate_Ranking     = 2019-06-17 00:00:00 +0000
kMDItemKind                        = "Plain Text Document"
kMDItemLastUsedDate                = 2019-06-17 19:41:48 +0000
kMDItemLastUsedDate_Ranking        = 2019-06-17 00:00:00 +0000
kMDItemLogicalSize                 = 191
kMDItemPhysicalSize                = 4096
kMDItemWhereFroms                  = (
    "https://github.com/settings/auth/recovery-codes/download",
    "https://github.com/settings/auth/recovery-codes"
)
➜  ~
```



```
kMDItemWhereFroms                  = (
    "https://github.com/settings/auth/recovery-codes/download",
    "https://github.com/settings/auth/recovery-codes"
```

KOLIDE

How can we interact with mdfind?

Spotlight *(GUI)*

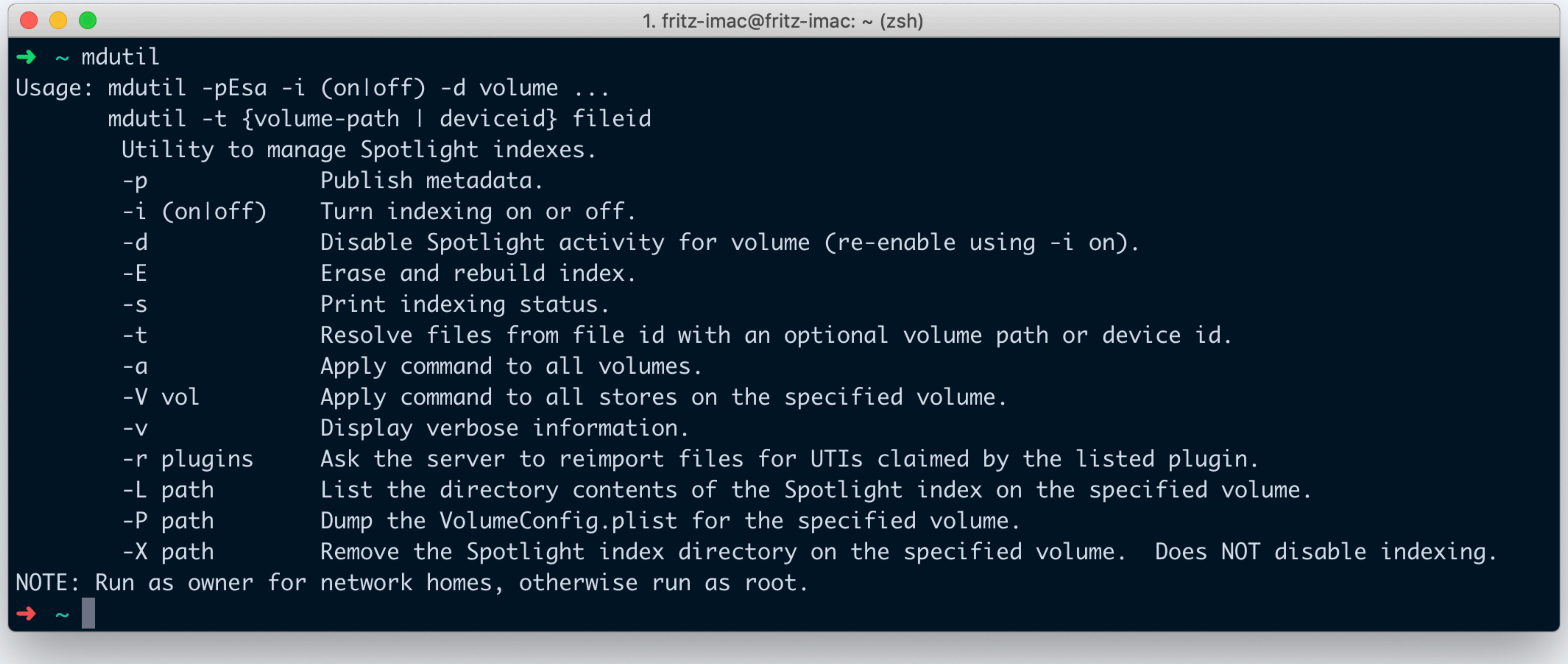
mdfind *(CLI)*

mdls *(CLI)*

mdutil *(CLI)*

mdimport *(CLI)*

osquery mdfind table



```
➜ ~ mdutil
Usage: mdutil -pEsa -i (on|off) -d volume ...
       mdutil -t {volume-path | deviceid} fileid
        Utility to manage Spotlight indexes.
        -p              Publish metadata.
        -i (on|off)     Turn indexing on or off.
        -d              Disable Spotlight activity for volume (re-enable using -i on).
        -E              Erase and rebuild index.
        -s              Print indexing status.
        -t              Resolve files from file id with an optional volume path or device id.
        -a              Apply command to all volumes.
        -V vol          Apply command to all stores on the specified volume.
        -v              Display verbose information.
        -r plugins      Ask the server to reimport files for UTIs claimed by the listed plugin.
        -L path         List the directory contents of the Spotlight index on the specified volume.
        -P path         Dump the VolumeConfig.plist for the specified volume.
        -X path         Remove the Spotlight index directory on the specified volume.  Does NOT disable indexing.
NOTE: Run as owner for network homes, otherwise run as root.
➜ ~
```

How can we interact with mdfind?

Spotlight *(GUI)*

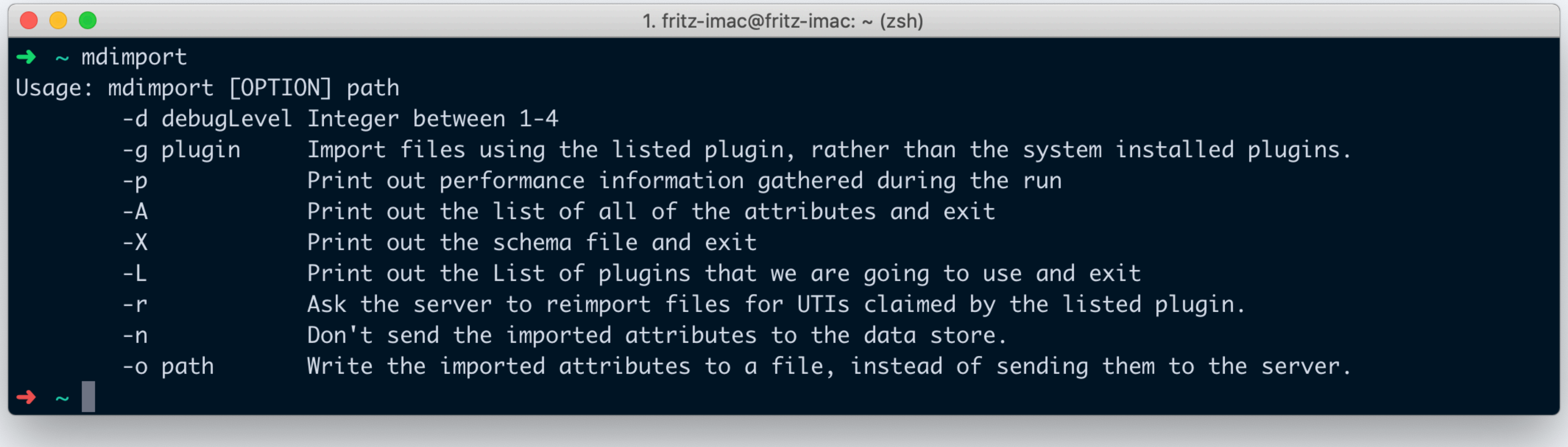
mdfind *(CLI)*

mdls *(CLI)*

mdutil *(CLI)*

mdimport *(CLI)*

osquery mdfind table



1. fritz-imac@fritz-imac: ~ (zsh)

```
➜ ~ mdimport
Usage: mdimport [OPTION] path
        -d debugLevel Integer between 1-4
        -g plugin     Import files using the listed plugin, rather than the system installed plugins.
        -p            Print out performance information gathered during the run
        -A            Print out the list of all of the attributes and exit
        -X            Print out the schema file and exit
        -L            Print out the List of plugins that we are going to use and exit
        -r            Ask the server to reimport files for UTIs claimed by the listed plugin.
        -n            Don't send the imported attributes to the data store.
        -o path       Write the imported attributes to a file, instead of sending them to the server.
➜ ~
```

KOLIDE

How can we interact with mdfind?

Spotlight *(GUI)*

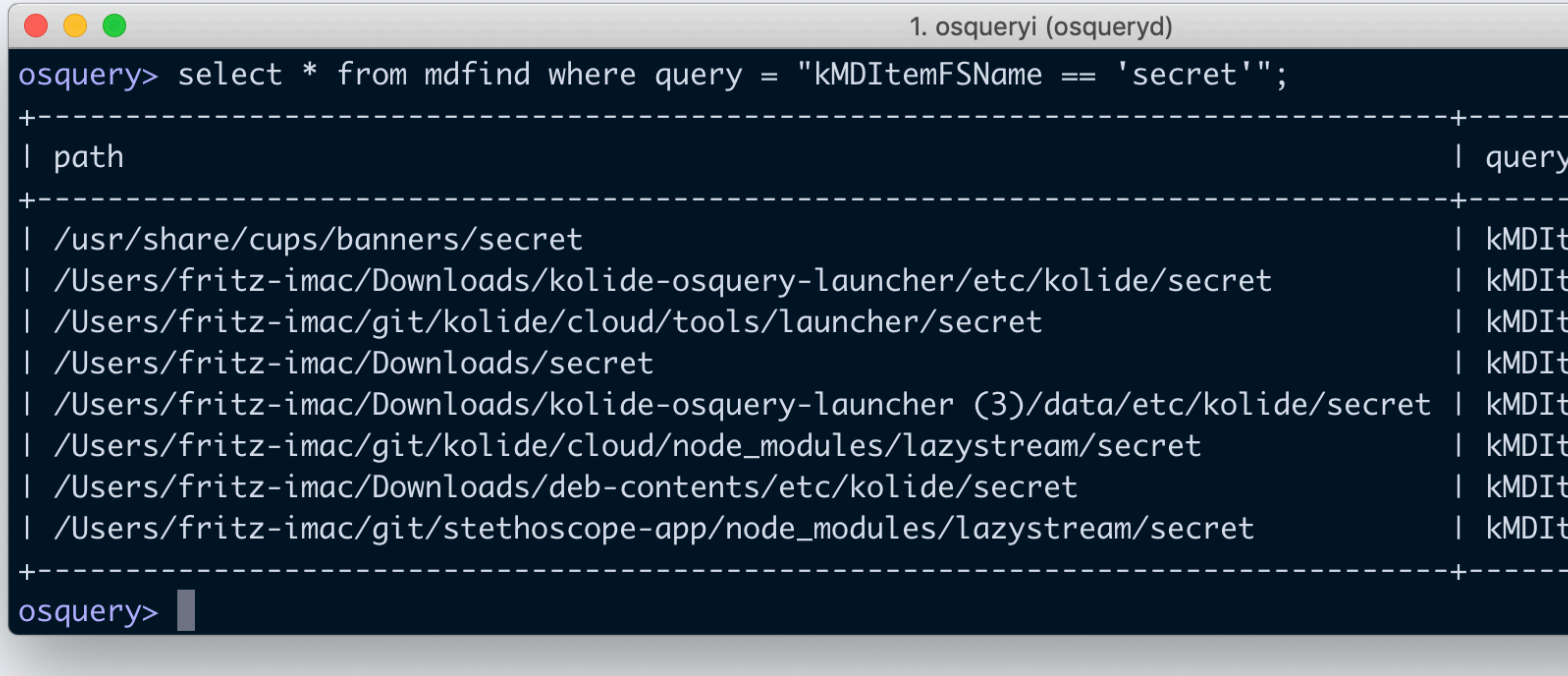
mdfind *(CLI)*

mdls *(CLI)*

mdutil *(CLI)*

mdimport *(CLI)*

osquery mdfind table



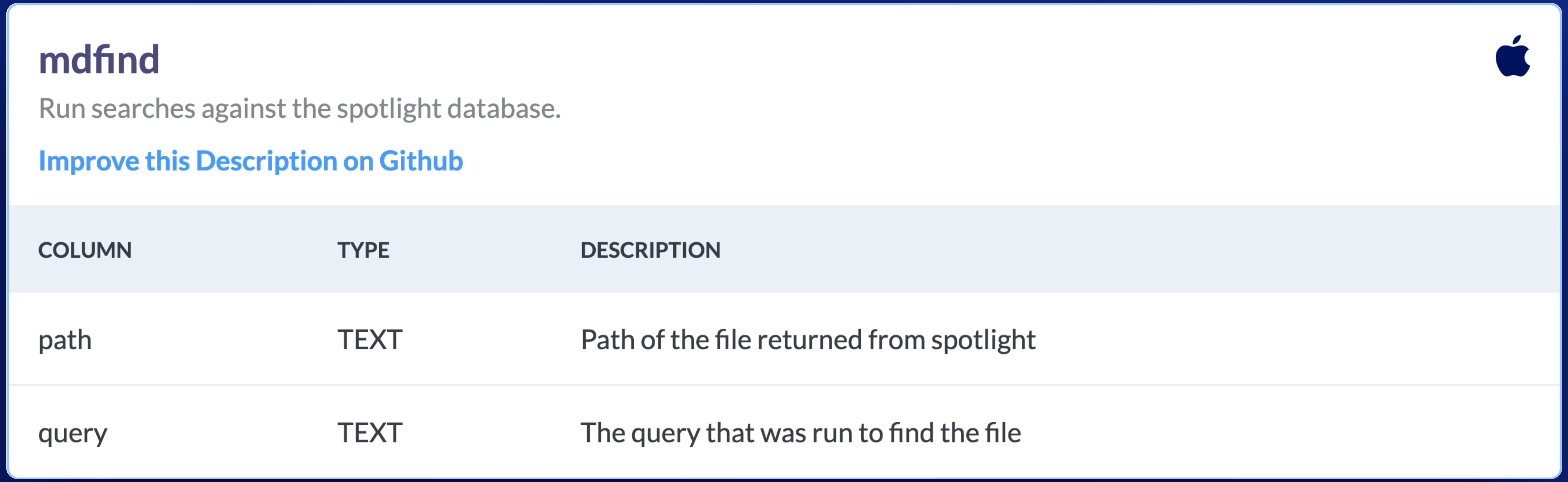
1. osqueryi (osqueryd)

```
osquery> select * from mdfind where query = "kMDItemFSName == 'secret'";
+---------------------------------------------------------------------------+------
| path                                                                      | query
+---------------------------------------------------------------------------+------
| /usr/share/cups/banners/secret                                            | kMDIt
| /Users/fritz-imac/Downloads/kolide-osquery-launcher/etc/kolide/secret     | kMDIt
| /Users/fritz-imac/git/kolide/cloud/tools/launcher/secret                  | kMDIt
| /Users/fritz-imac/Downloads/secret                                        | kMDIt
| /Users/fritz-imac/Downloads/kolide-osquery-launcher (3)/data/etc/kolide/secret | kMDIt
| /Users/fritz-imac/git/kolide/cloud/node_modules/lazystream/secret         | kMDIt
| /Users/fritz-imac/Downloads/deb-contents/etc/kolide/secret                | kMDIt
| /Users/fritz-imac/git/stethoscope-app/node_modules/lazystream/secret      | kMDIt
+---------------------------------------------------------------------------+------
osquery>
```

Groob

Obelisk

Using mdfind in osquery



mdfind

Run searches against the spotlight database.

Improve this Description on Github

COLUMN	TYPE	DESCRIPTION
path	TEXT	Path of the file returned from spotlight
query	TEXT	The query that was run to find the file

```
SELECT * FROM mdfind
WHERE query = "kMDItemFSName == 'foobar'";
```

kMD Attributes are *the key* to mdfind

```
kMDItemFSName

k  - (Hungarian Notation for konstant)
MD - (metadata)
Item
AttributeName
```

```
$> mdimport -A
```

```
kMDItemAccountIdentifier
kMDItemAcquisitionMake
kMDItemAcquisitionModel
kMDItemAdditionalRecipientEmailAddresses
kMDItemAlbum
kMDItemAlternateNames
kMDItemAltitude
kMDItemAperture
kMDItemAppStoreCategory
kMDItemAppStoreCategoryType
kMDItemAppleLoopDescriptors
kMDItemAppleLoopsKeyFilterType
kMDItemAppleLoopsLoopMode
kMDItemAppleLoopsRootKey
kMDItemApplicationCategories
kMDItemAttributeChangeDate
kMDItemAudiences
kMDItemAudioBitRate
kMDItemAudioChannelCount
kMDItemAudioEncodingApplication
kMDItemAudioSampleRate
kMDItemAudioTrackNumber
kMDItemAuthorAddresses
kMDItemAuthorContactIdentifiers
kMDItemAuthorEmailAddresses
kMDItemAuthors
kMDItemBitsPerSample
kMDItemBundleIdentifier
kMDItemCFBundleIdentifier
kMDItemCalendarHolidayIdentifier
kMDItemCity
kMDItemCodecs
kMDItemColorSpace
kMDItemComment
kMDItemComposer
kMDItemContactKeywords
kMDItemContentCreationDate
kMDItemContentModificationDate
kMDItemContentType
kMDItemContentTypeTree
kMDItemContributors
kMDItemCopyright
kMDItemCountry
kMDItemCoverage
kMDItemCreator
kMDItemDateAdded
kMDItemDeliveryType
kMDItemDescription
kMDItemDestinationRecipients
kMDItemDirector
kMDItemDisplayName
kMDItemDocumentContainer
kMDItemDocumentLineage
kMDItemDueDate
kMDItemDurationSeconds
kMDItemEXIFGPSVersion
kMDItemEXIFVersion
kMDItemEditors
kMDItemEmailAddresses
kMDItemEmailCategory
kMDItemEmailConversationID
kMDItemEncodingApplications
kMDItemExecutableArchitectures
kMDItemExecutablePlatform
kMDItemExposureMode
kMDItemExposureProgram
kMDItemExposureTimeSeconds
kMDItemFNumber
kMDItemFSContentChangeDate
kMDItemFSCreationDate
kMDItemFSExists
kMDItemFSHasCustomIcon
kMDItemFSInvisible
kMDItemFSIsExtensionHidden
kMDItemFSIsReadable
kMDItemFSIsStationery
kMDItemFSIsWriteable
kMDItemFSLabel
kMDItemFSName
kMDItemFSNodeCount
kMDItemFSOwnerGroupID
kMDItemFSOwnerUserID
kMDItemFSSize
kMDItemFinderComment
kMDItemFinderOpenDate
kMDItemFlashOnOff
kMDItemFocalLength
kMDItemFonts
kMDItemGenre
kMDItemHasAlphaChannel
kMDItemHeadline
kMDItemHiddenAdditionalRecipientEmailAddresses
kMDItemISOSpeed
kMDItemIdentifier
kMDItemInformation
kMDItemInstantMessageAddresses
kMDItemInstructions
kMDItemIsApplicationManaged
kMDItemIsGeneralMIDISequence
kMDItemIsLikelyJunk
kMDItemIsQuarantined
kMDItemIsScreenCapture
kMDItemKeySignature
kMDItemKeywords
kMDItemKind
kMDItemLanguages
kMDItemLastUsedDate
kMDItemLatitude
kMDItemLayerNames
kMDItemLensModel
kMDItemLogicSongAlternatives
kMDItemLogicSongUsedAudioFiles
kMDItemLogicSongUsedEXSInstruments
kMDItemLogicSongUsedImpulseResponses
kMDItemLogicSongUsedUltrabeatFiles
kMDItemLogicSongUsedVideoFiles
kMDItemLogicalSize
kMDItemLongitude
kMDItemLyricist
kMDItemMailboxes
kMDItemMaxAperture
kMDItemMediaTypes
kMDItemMeteringMode
kMDItemMusicalGenre
kMDItemMusicalInstrumentCategory
kMDItemMusicalIns
kMDItemNamedLoca
kMDItemNumberOfPa
kMDItemOrganizat
kMDItemOrientati
kMDItemOriginApp
kMDItemOriginMes
kMDItemOriginSen
kMDItemOriginSen
kMDItemOriginSub
kMDItemOriginalF
kMDItemOriginalS
kMDItemPageHeigh
kMDItemPageWidth
kMDItemParticipa
kMDItemPath
kMDItemPerformer
kMDItemPhoneNumb
kMDItemPhysicalS
kMDItemPixelCoun
kMDItemPixelHeig
kMDItemPixelWidth
kMDItemPrimaryRecipientEmailAddresses
kMDItemProducer
kMDItemProfileName
kMDItemProjects
kMDItemPublishers
kMDItemPurchaseDate
kMDItemRecipientAddresses
kMDItemRecipientContactIdentifiers
kMDItemRecipientEmailAddresses
kMDItemRecipients
kMDItemRecordingDate
kMDItemRecordingYear
kMDItemRedEyeOnOff
kMDItemRelatedUniqueIdentifier
kMDItemResolutionHeightDPI
kMDItemResolutionWidthDPI
kMDItemRights
kMDItemScreenCaptureType
kMDItemSecurityMethod
kMDItemStarRating
kMDItemStateOrProvince
kMDItemStreamable
kMDItemSubject
kMDItemSupportFileType
kMDItemTempo
kMDItemTextContent
kMDItemTheme
kMDItemTimeSignature
kMDItemTitle
kMDItemTotalBitRate
kMDItemUserTags
kMDItemVersion
kMDItemVideoBitRate
kMDItemWeakRelatedUniqueIdentifier
kMDItemWhereFroms
kMDItemWhiteBalance
```

```
kMDItemFSContentChangeDate
kMDItemFSCreationDate
kMDItemFSName
kMDItemFSSize
kMDItemTextContent
kMDItemWhereFroms
```

KOLIDE

Mdfind Query Syntax

== equals

```
"kMDItemFSName == 'foo'"
```

!= doesn't equal

```
"kMDItemFSName != 'foobar'"
```

< less than

```
"kMDItemFSSize < 24"
```

> greater than

```
"kMDItemFSSize > 10737418240"
```

<= less than or equal to

```
"kMDItemFSName <= 'foobar'"
```

>= greater than or equal to

```
"kMDItemFSContentChangeDate >= $time.this_month(-2)"
```

c makes string case insensitive

```
"kMDItemFSName = 'FoO'c"
```

d ignores diacritical marks (such as à, ê, ñ, ß, etc.)

```
"kMDItemFSName = 'föo'd"
```

*** string wildcard (can trail or lead, Cannot infix)**

```
"kMDItemFSName == ‘*foo*’"
```

&& AND condition

```
"kMDItemFSName = 'foo' && kMDItemTextContent = 'bar'"
```

|| OR condition

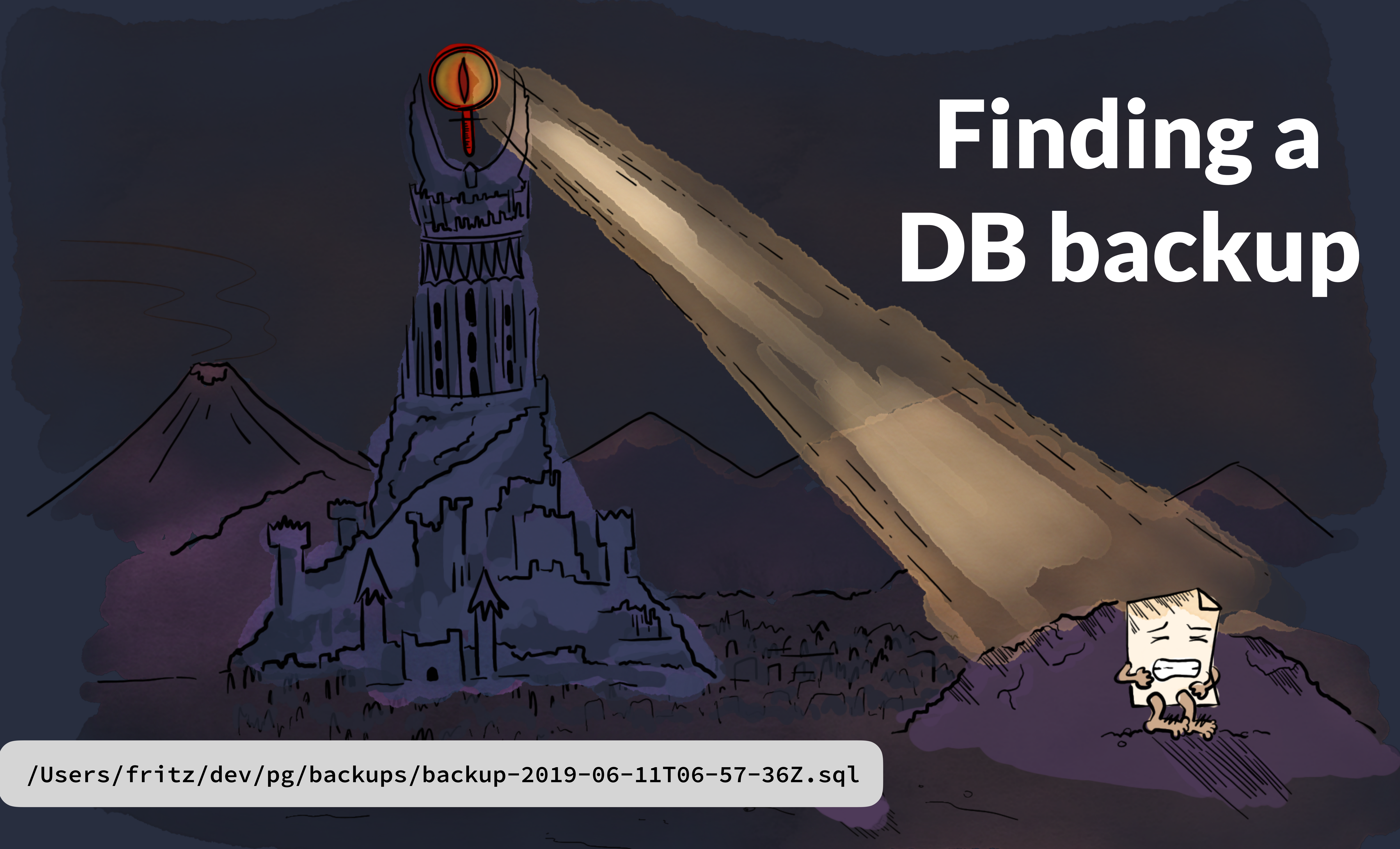
```
"kMDItemFSName = 'foo' || kMDItemFSName = 'bar'"
```

() Separating and nesting groups of conditions

```
"(kMDItemFSName = 'foo' || kMDItemFSName = 'bar') &&…
```

KOLIDE

Finding a DB backup



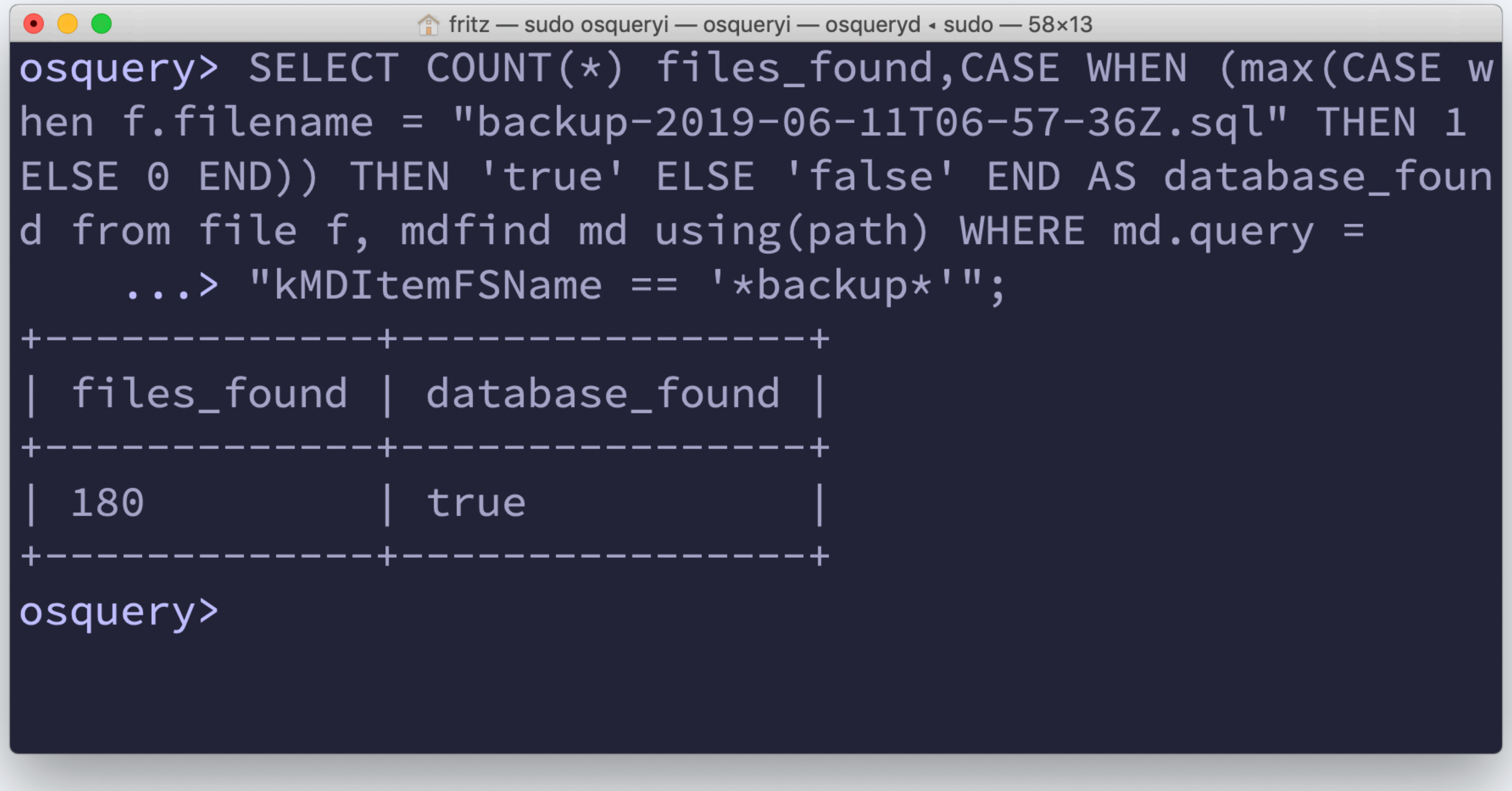
```
/Users/fritz/dev/pg/backups/backup-2019-06-11T06-57-36Z.sql
```

Finding a DB Backup

Attempt #1: Naive Single Condition (filename only)

The basic mdfind query: Single condition, zero complexity

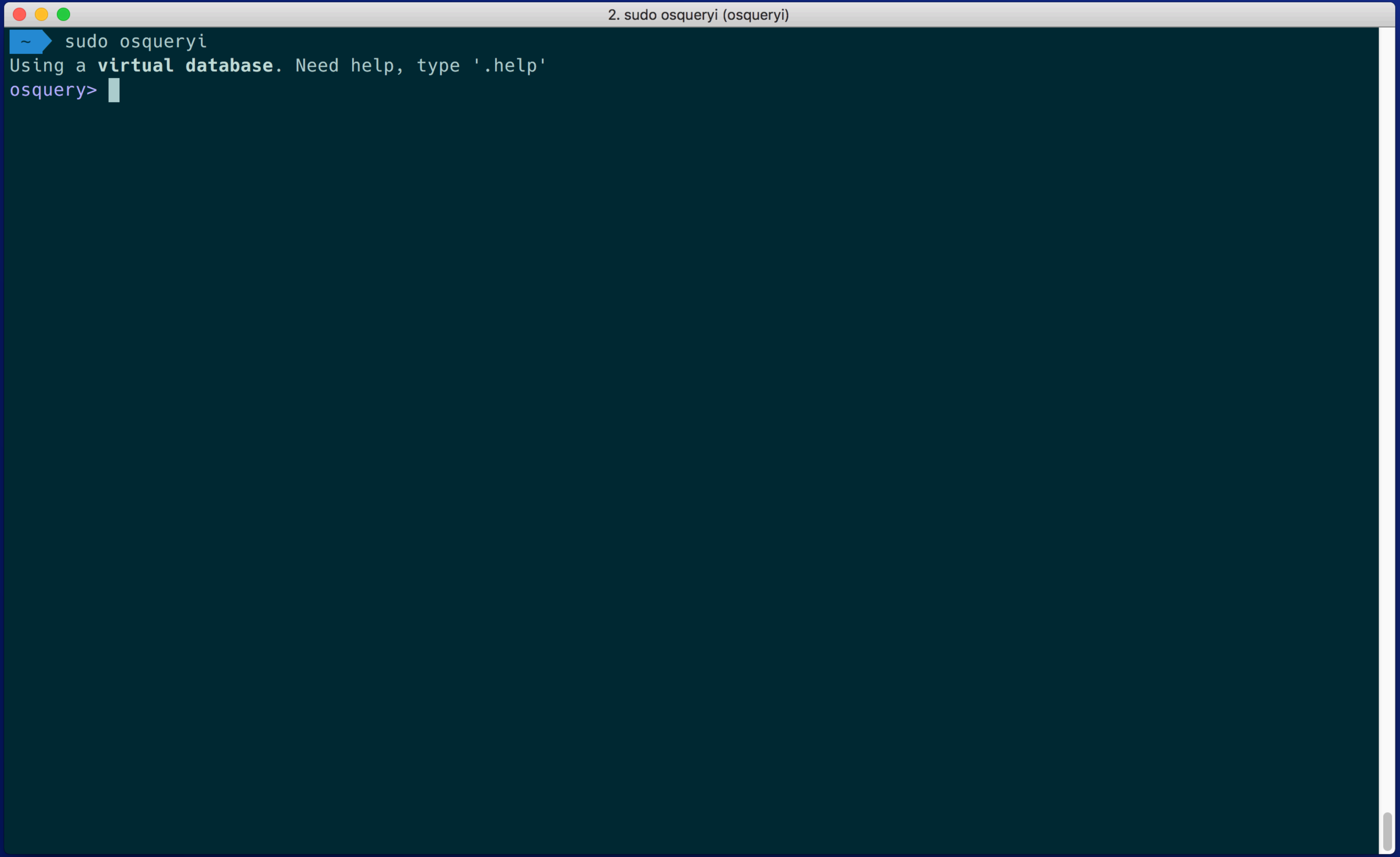
- Looks for a file with the string backup in the title.
- Similar output to what a perfect recursive file table search would return



```
osquery> SELECT COUNT(*) files_found,CASE WHEN (max(CASE w
hen f.filename = "backup-2019-06-11T06-57-36Z.sql" THEN 1
ELSE 0 END)) THEN 'true' ELSE 'false' END AS database_foun
d from file f, mdfind md using(path) WHERE md.query =
    ...> "kMDItemFSName == '*backup*'";
+-------------+----------------+
| files_found | database_found |
+-------------+----------------+
| 180         | true           |
+-------------+----------------+
osquery>
```

```
osquery> SELECT f.path
         FROM file f, mdfind md USING (path)
         WHERE md.query = "kMDItemFSName == '*backup*'"
```

Live demo



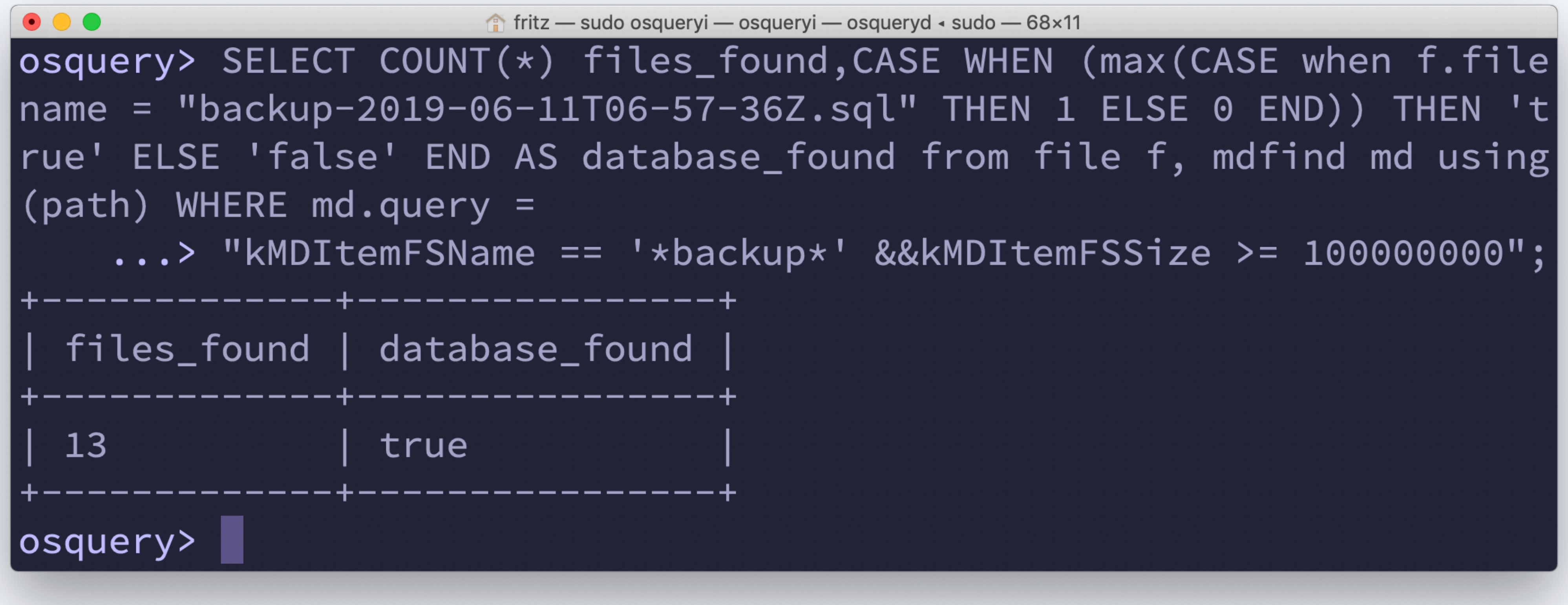
```
~  sudo osqueryi
Using a virtual database. Need help, type '.help'
osquery>
```

Finding a DB Backup

Attempt #2: Two conditions (scope based on size)

Two Conditions:

- Looks for a file with the string 'backup' in the title.
- Omits any results that are smaller than 100 MB in size



```
osquery> SELECT COUNT(*) files_found,CASE WHEN (max(CASE when f.file
name = "backup-2019-06-11T06-57-36Z.sql" THEN 1 ELSE 0 END)) THEN 't
rue' ELSE 'false' END AS database_found from file f, mdfind md using
(path) WHERE md.query =
    ...> "kMDItemFSName == '*backup*' &&kMDItemFSSize >= 100000000";
+-------------+----------------+
| files_found | database_found |
+-------------+----------------+
| 13          | true           |
+-------------+----------------+
osquery>
```

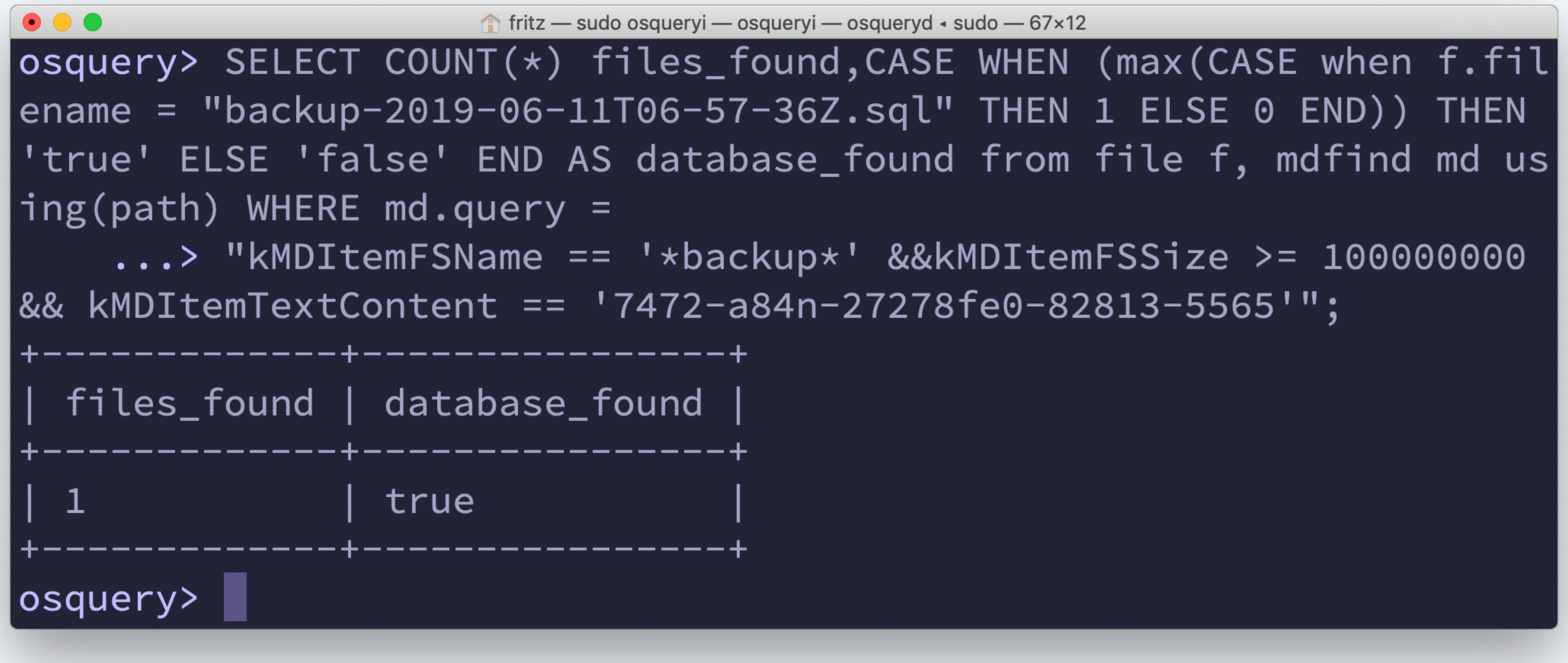
```
osquery> SELECT f.path
         FROM file f, mdfind md USING (path)
         WHERE md.query =
         "kMDItemFSName == '*backup*' && kMDItemFSSize >= 100000000";
```

Finding a DB Backup

Attempt #3: Three conditions (Check Item Content Index)

Three Conditions:

- Looks for a file with the string 'backup' in the title.
- Omits any results that are smaller than 100 MB in size
- Looks for any file whose contents contains a known customer UUID



```
osquery> SELECT COUNT(*) files_found,CASE WHEN (max(CASE when f.fil
ename = "backup-2019-06-11T06-57-36Z.sql" THEN 1 ELSE 0 END)) THEN
'true' ELSE 'false' END AS database_found from file f, mdfind md us
ing(path) WHERE md.query =
    ...> "kMDItemFSName == '*backup*' &&kMDItemFSSize >= 100000000
&& kMDItemTextContent == '7472-a84n-27278fe0-82813-5565'";
+-------------+-----------------+
| files_found | database_found |
+-------------+-----------------+
| 1           | true            |
+-------------+-----------------+
osquery> 
```

```
osquery> SELECT f.path
         FROM file f, mdfind md USING (path)
         WHERE md.query =
         "kMDItemFSName == '*backup*' && kMDItemFSSize >= 100000000 &&
kMDItemTextContent == '7472-a84n-2727f8e0-82813-5565';
```

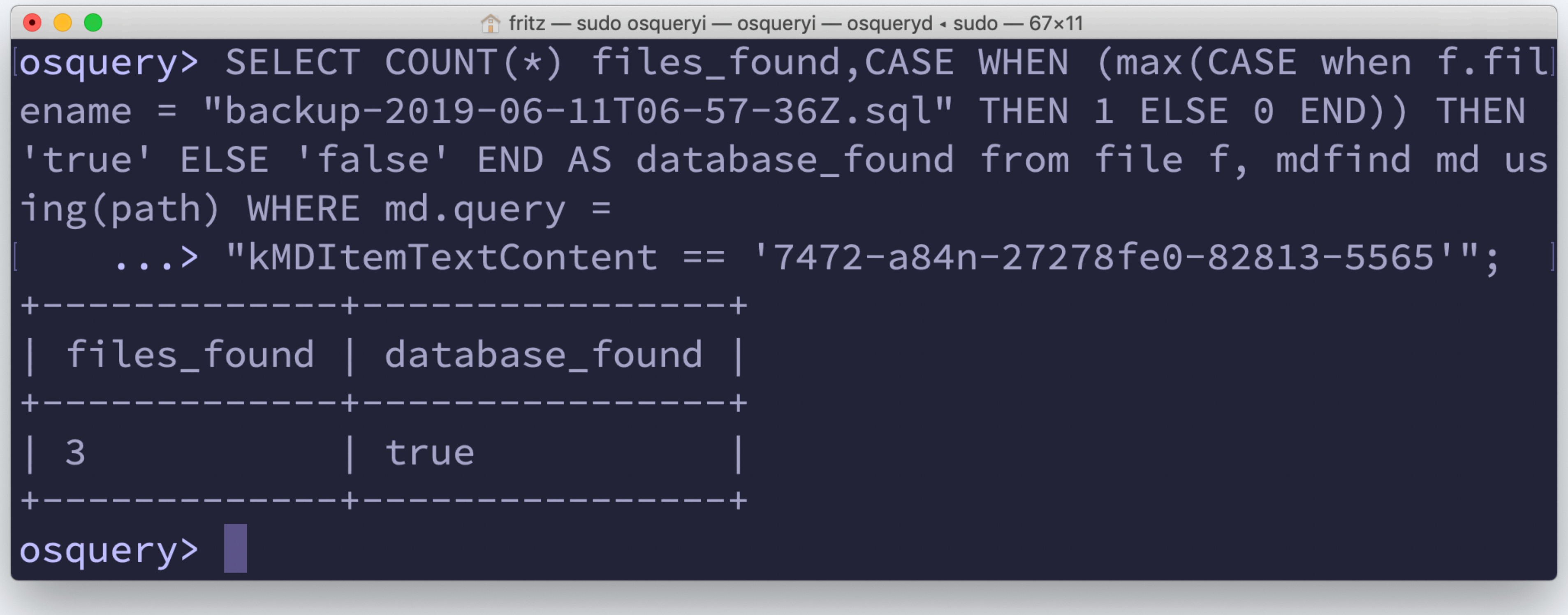
Finding a DB Backup

Attempt #4: Back to one condition (Less complexity can be a good thing)

One Condition:

- Looks for any file whose contents contains the unique UUID string:

'7472-a84n-2727f8e0-82813-5565'



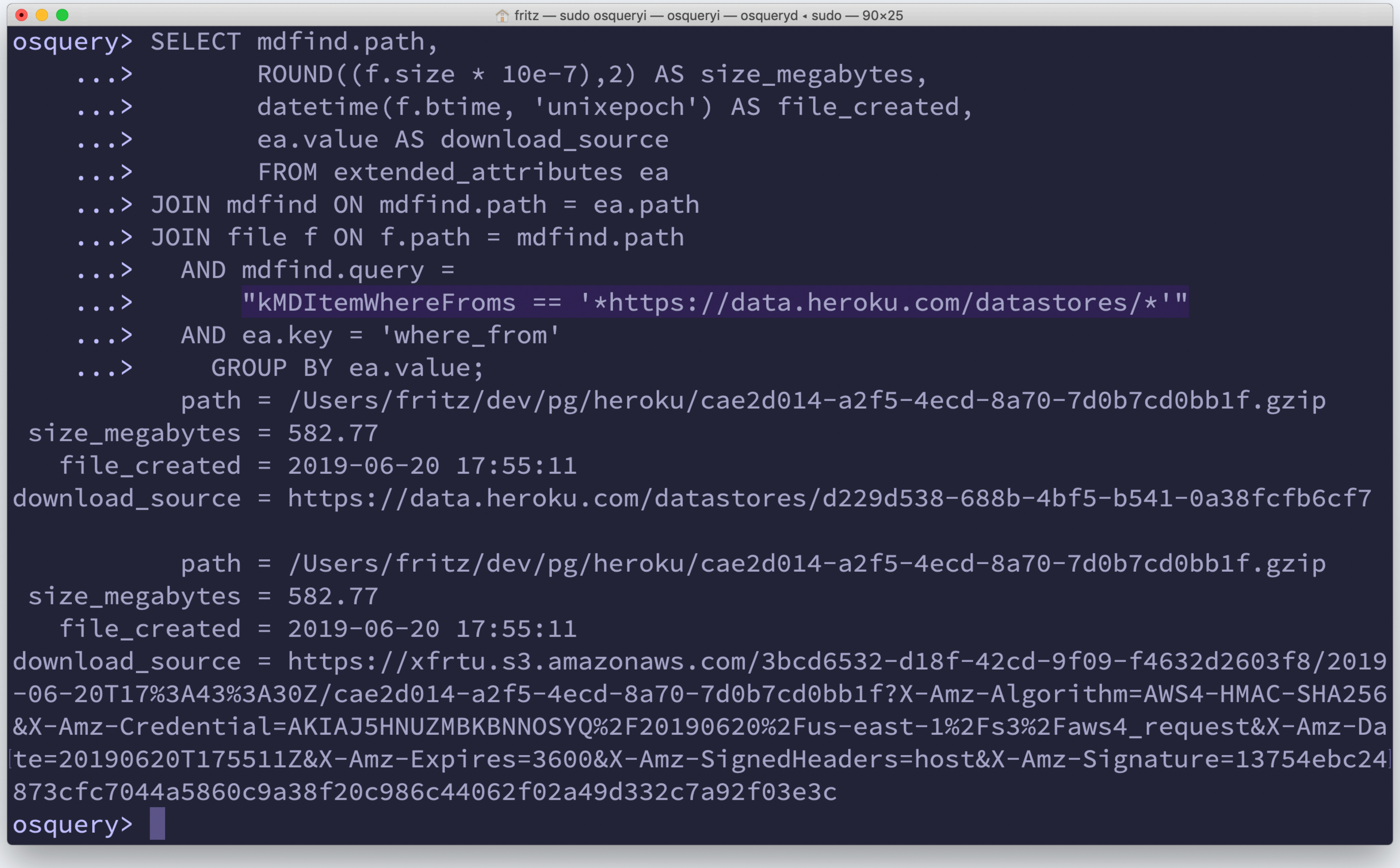
```
osquery> SELECT COUNT(*) files_found,CASE WHEN (max(CASE when f.fil
ename = "backup-2019-06-11T06-57-36Z.sql" THEN 1 ELSE 0 END)) THEN
'true' ELSE 'false' END AS database_found from file f, mdfind md us
ing(path) WHERE md.query =
    ...> "kMDItemTextContent == '7472-a84n-27278fe0-82813-5565'";
+-------------+----------------+
| files_found | database_found |
+-------------+----------------+
| 3           | true           |
+-------------+----------------+
osquery>
```

```
osquery> SELECT f.path
         FROM file f, mdfind md USING (path)
         WHERE md.query =
         "kMDItemTextContent == '7472-a84n-2727f8e0-82813-5565'";
```

A different approach

Using wherefrom

- What about files that are gzipped or tarred where file contents aren’t cached?
- kMDItemWhereFroms



```
osquery> SELECT mdfind.path,
    ...>        ROUND((f.size * 10e-7),2) AS size_megabytes,
    ...>        datetime(f.btime, 'unixepoch') AS file_created,
    ...>        ea.value AS download_source
    ...>        FROM extended_attributes ea
    ...> JOIN mdfind ON mdfind.path = ea.path
    ...> JOIN file f ON f.path = mdfind.path
    ...>   AND mdfind.query =
    ...>       "kMDItemWhereFroms == '*https://data.heroku.com/datastores/*'"
    ...>   AND ea.key = 'where_from'
    ...>     GROUP BY ea.value;
           path = /Users/fritz/dev/pg/heroku/cae2d014-a2f5-4ecd-8a70-7d0b7cd0bb1f.gzip
 size_megabytes = 582.77
   file_created = 2019-06-20 17:55:11
download_source = https://data.heroku.com/datastores/d229d538-688b-4bf5-b541-0a38fcfb6cf7

           path = /Users/fritz/dev/pg/heroku/cae2d014-a2f5-4ecd-8a70-7d0b7cd0bb1f.gzip
 size_megabytes = 582.77
   file_created = 2019-06-20 17:55:11
download_source = https://xfrtu.s3.amazonaws.com/3bcd6532-d18f-42cd-9f09-f4632d2603f8/2019
-06-20T17%3A43%3A30Z/cae2d014-a2f5-4ecd-8a70-7d0b7cd0bb1f?X-Amz-Algorithm=AWS4-HMAC-SHA256
&X-Amz-Credential=AKIAJ5HNUZMBKBNNOSYQ%2F20190620%2Fus-east-1%2Fs3%2Faws4_request&X-Amz-Da
te=20190620T175511Z&X-Amz-Expires=3600&X-Amz-SignedHeaders=host&X-Amz-Signature=13754ebc24
873cfc7044a5860c9a38f20c986c44062f02a49d332c7a92f03e3c
osquery>
```

```
osquery> SELECT f.path
         FROM file f, mdfind md USING (path)
         WHERE md.query =
         "kMDItemWhereFroms == '*https://data.heroku.com/datastores/*'"
```

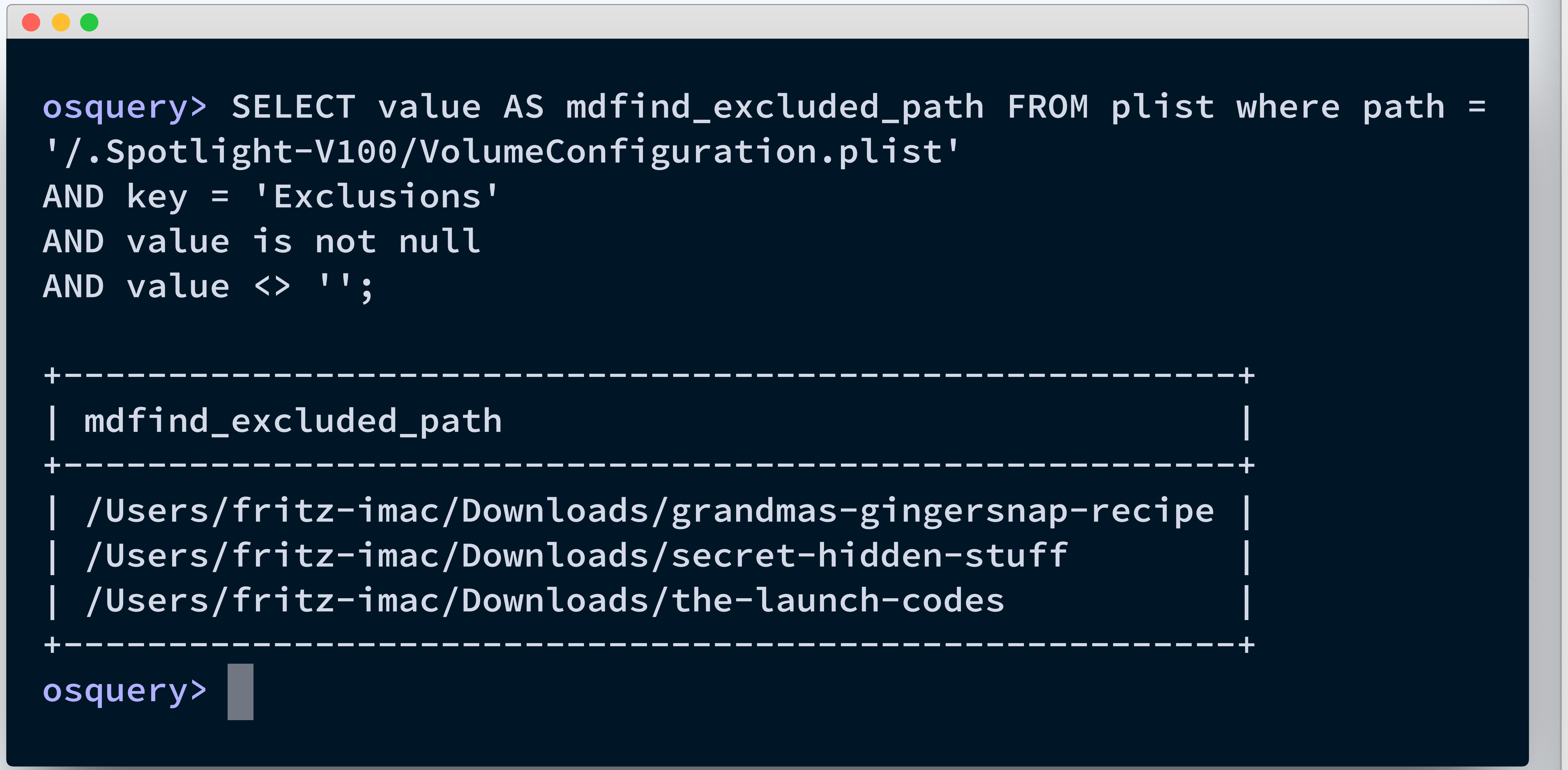
We did it! We're the best!

But sometimes mdfind won't work…

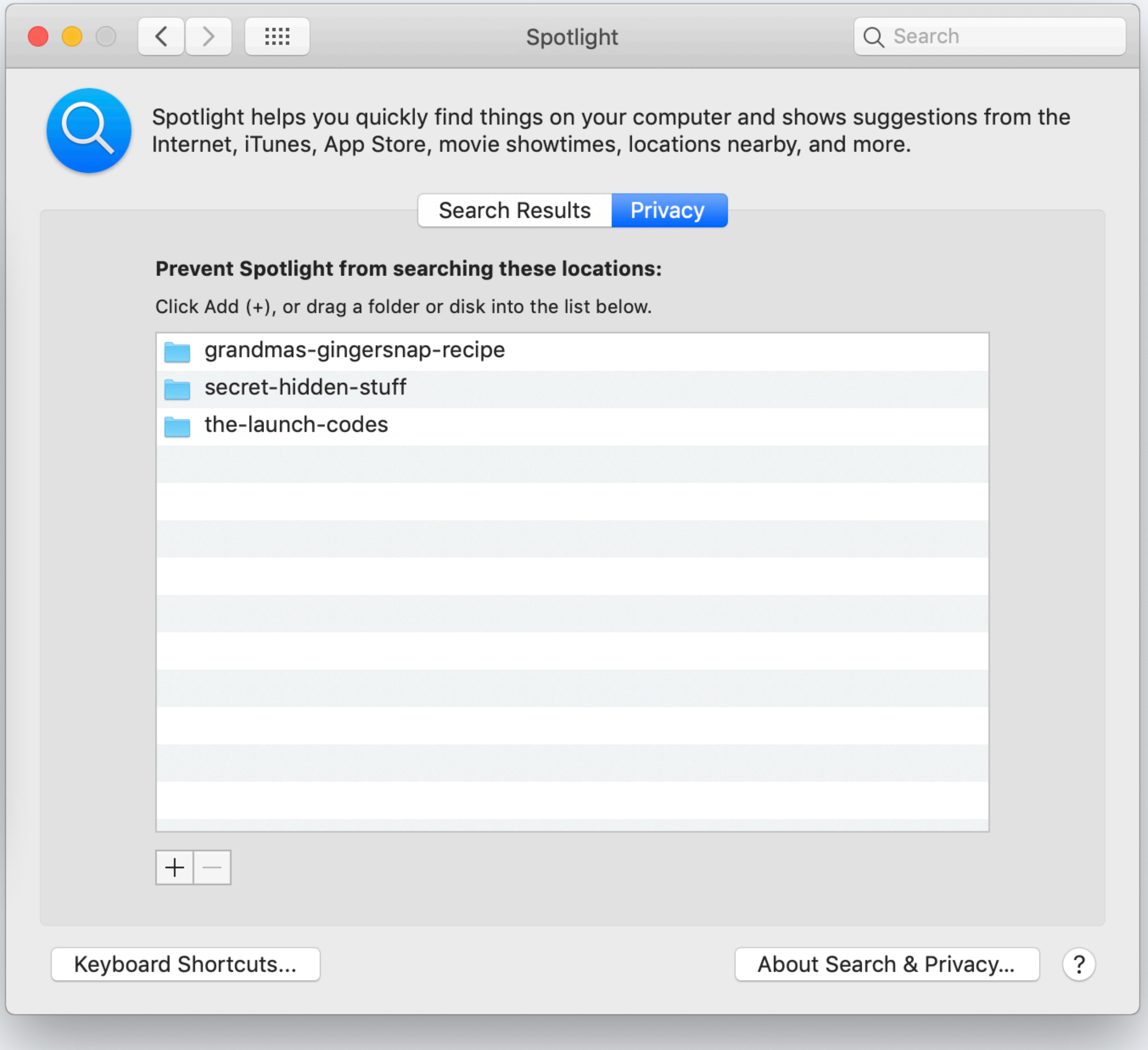
Caveats

- macOS only
- Can be disabled by the end user
- Limited to User-Visible directories (with optional omissions)
- Attempts to not cache files that appear to contain key material
- Uses a different query syntax than osquery

Checking for mdfind exclusions on your devices



```
osquery> SELECT value AS mdfind_excluded_path FROM plist where path =
'/.Spotlight-V100/VolumeConfiguration.plist'
AND key = 'Exclusions'
AND value is not null
AND value <> '';

+------------------------------------------------------------+
| mdfind_excluded_path                                       |
+------------------------------------------------------------+
| /Users/fritz-imac/Downloads/grandmas-gingersnap-recipe |
| /Users/fritz-imac/Downloads/secret-hidden-stuff            |
| /Users/fritz-imac/Downloads/the-launch-codes               |
+------------------------------------------------------------+
osquery>
```



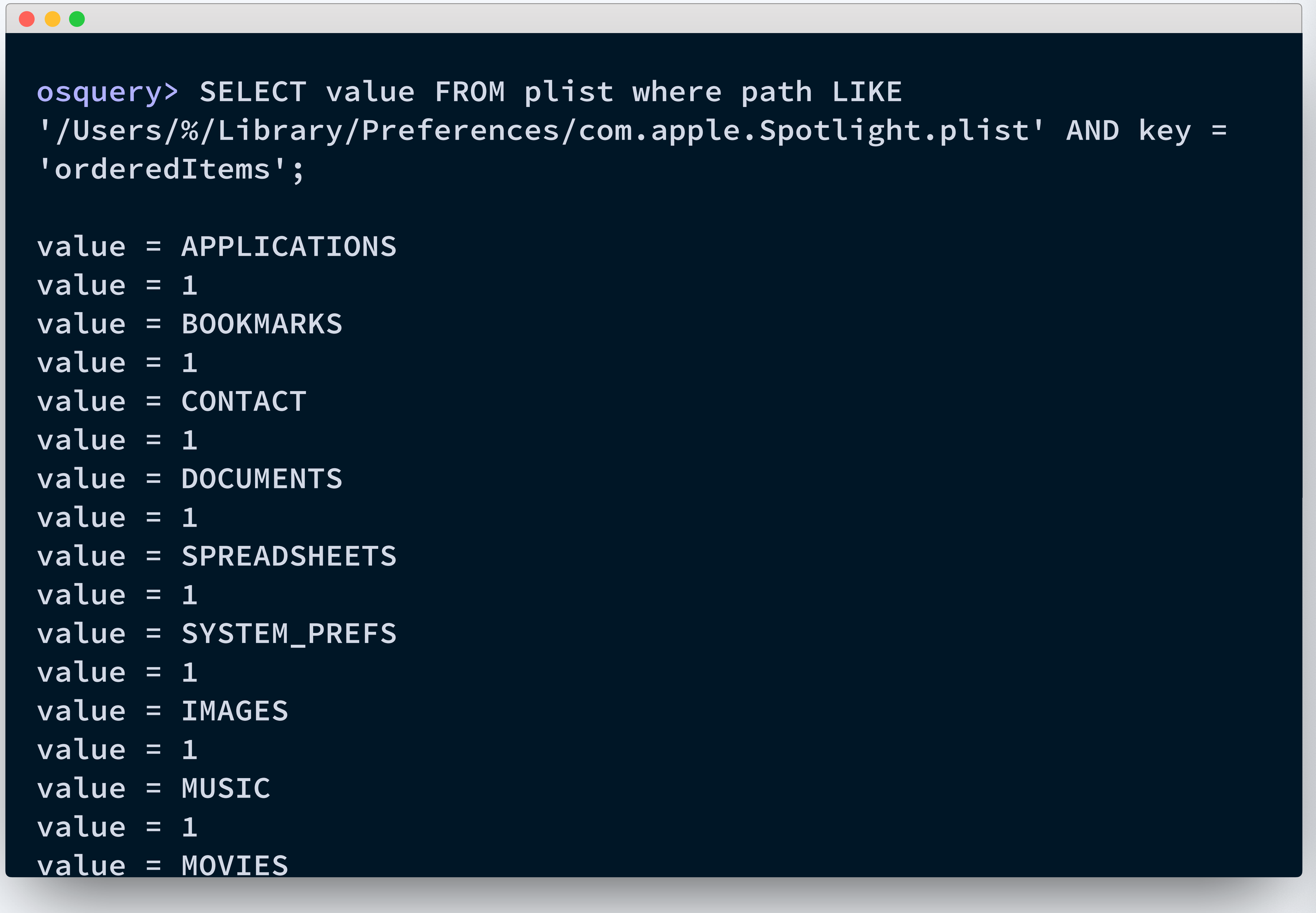
Spotlight helps you quickly find things on your computer and shows suggestions from the Internet, iTunes, App Store, movie showtimes, locations nearby, and more.

Prevent Spotlight from searching these locations:

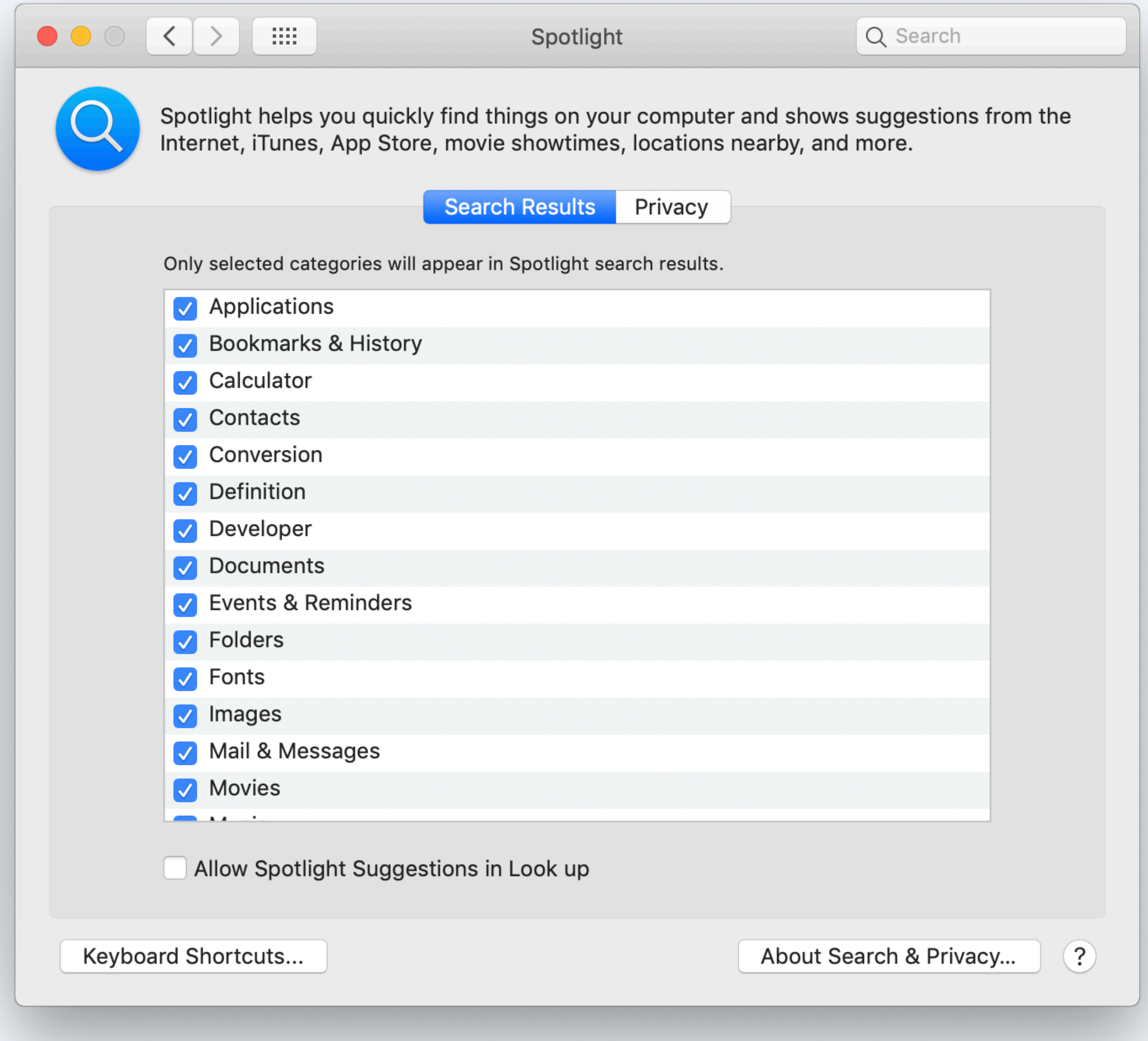
Click Add (+), or drag a folder or disk into the list below.

- grandmas-gingersnap-recipe
- secret-hidden-stuff
- the-launch-codes

Checking for mdfind exclusions on your devices



```
osquery> SELECT value FROM plist where path LIKE
'/Users/%/Library/Preferences/com.apple.Spotlight.plist' AND key =
'orderedItems';

value = APPLICATIONS
value = 1
value = BOOKMARKS
value = 1
value = CONTACT
value = 1
value = DOCUMENTS
value = 1
value = SPREADSHEETS
value = 1
value = SYSTEM_PREFS
value = 1
value = IMAGES
value = 1
value = MUSIC
value = 1
value = MOVIES
```



Spotlight helps you quickly find things on your computer and shows suggestions from the Internet, iTunes, App Store, movie showtimes, locations nearby, and more.

Only selected categories will appear in Spotlight search results.

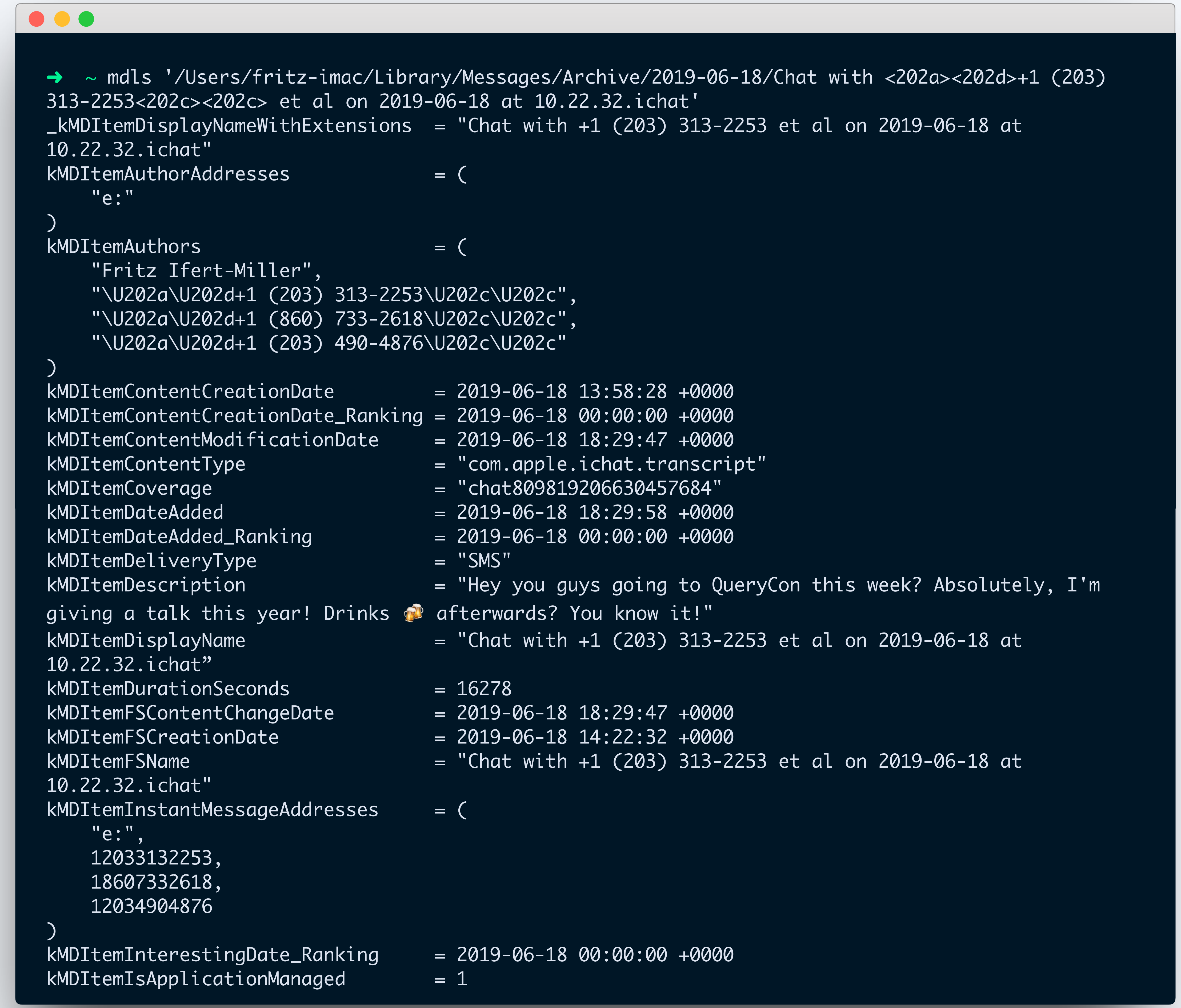
☑ Applications
☑ Bookmarks & History
☑ Calculator
☑ Contacts
☑ Conversion
☑ Definition
☑ Developer
☑ Documents
☑ Events & Reminders
☑ Folders
☑ Fonts
☑ Images
☑ Mail & Messages
☑ Movies

□ Allow Spotlight Suggestions in Look up

Privacy Implications of mdfind

kMDItemInstantMessageAddresses =
+1 (203) 313-2253
+1 (860) 733-2618
+1 (203) 490-4876

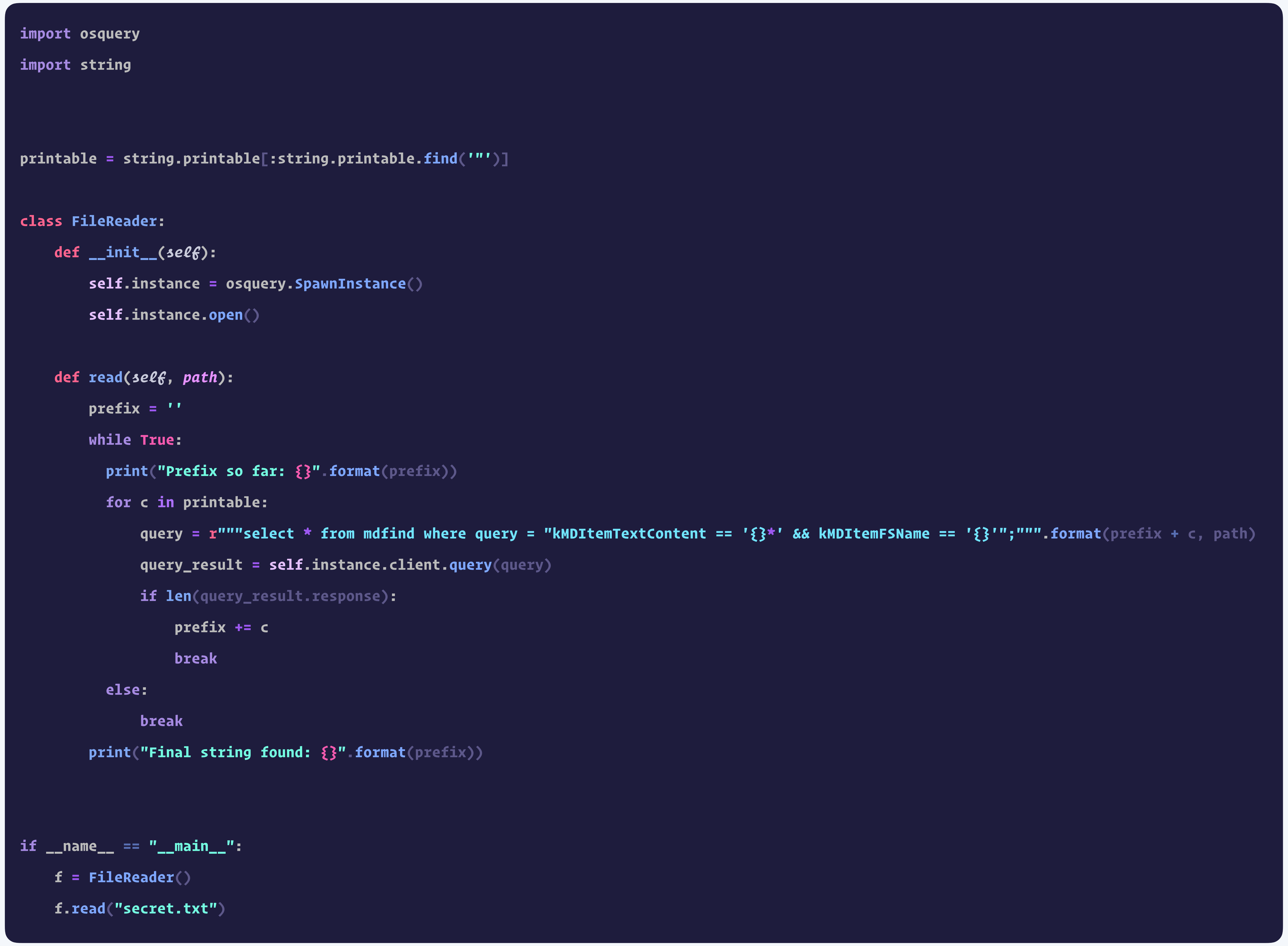
kMDItemDescription =
"Hey you guys going to QueryCon this week? Absolutely, I'm giving a talk this year! Drinks 🍻 afterwards? You know it!"



```
➜  ~ mdls '/Users/fritz-imac/Library/Messages/Archive/2019-06-18/Chat with <202a><202d>+1 (203)
313-2253<202c><202c> et al on 2019-06-18 at 10.22.32.ichat'
_kMDItemDisplayNameWithExtensions  = "Chat with +1 (203) 313-2253 et al on 2019-06-18 at
10.22.32.ichat"
kMDItemAuthorAddresses             = (
    "e:"
)
kMDItemAuthors                     = (
    "Fritz Ifert-Miller",
    "\U202a\U202d+1 (203) 313-2253\U202c\U202c",
    "\U202a\U202d+1 (860) 733-2618\U202c\U202c",
    "\U202a\U202d+1 (203) 490-4876\U202c\U202c"
)
kMDItemContentCreationDate         = 2019-06-18 13:58:28 +0000
kMDItemContentCreationDate_Ranking = 2019-06-18 00:00:00 +0000
kMDItemContentModificationDate     = 2019-06-18 18:29:47 +0000
kMDItemContentType                 = "com.apple.ichat.transcript"
kMDItemCoverage                    = "chat809819206630457684"
kMDItemDateAdded                   = 2019-06-18 18:29:58 +0000
kMDItemDateAdded_Ranking           = 2019-06-18 00:00:00 +0000
kMDItemDeliveryType                = "SMS"
kMDItemDescription                 = "Hey you guys going to QueryCon this week? Absolutely, I'm
giving a talk this year! Drinks 🍻 afterwards? You know it!"
kMDItemDisplayName                 = "Chat with +1 (203) 313-2253 et al on 2019-06-18 at
10.22.32.ichat”
kMDItemDurationSeconds             = 16278
kMDItemFSContentChangeDate         = 2019-06-18 18:29:47 +0000
kMDItemFSCreationDate              = 2019-06-18 14:22:32 +0000
kMDItemFSName                      = "Chat with +1 (203) 313-2253 et al on 2019-06-18 at
10.22.32.ichat"
kMDItemInstantMessageAddresses     = (
    "e:",
    12033132253,
    18607332618,
    12034904876
)
kMDItemInterestingDate_Ranking     = 2019-06-18 00:00:00 +0000
kMDItemIsApplicationManaged        = 1
```

Privacy Implications cont.

Open issue for data leakage via brute forcing mdfind wildcard capabilities



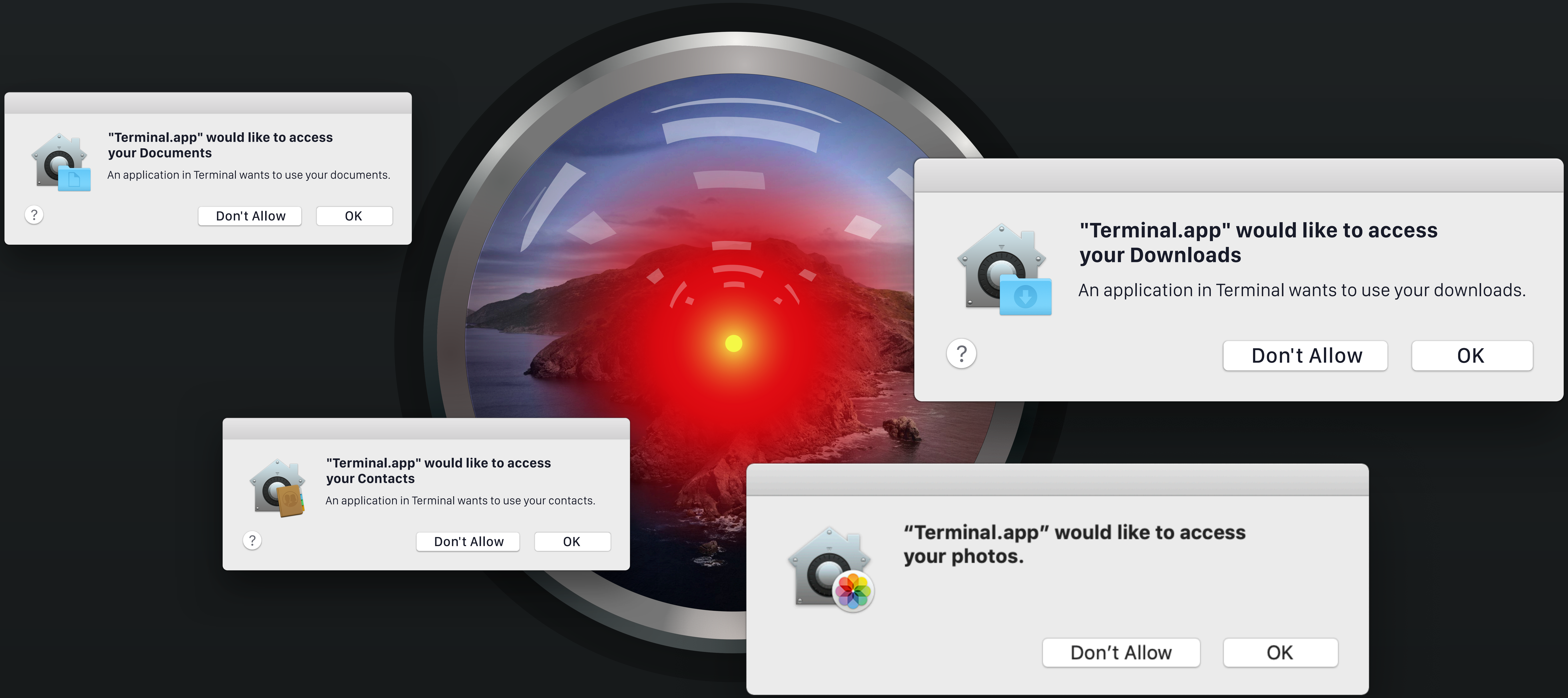
```python
import osquery
import string


printable = string.printable[:string.printable.find('"')]

class FileReader:
    def __init__(self):
        self.instance = osquery.SpawnInstance()
        self.instance.open()

    def read(self, path):
        prefix = ''
        while True:
          print("Prefix so far: {}".format(prefix))
          for c in printable:
              query = r"""select * from mdfind where query = "kMDItemTextContent == '{}*' && kMDItemFSName == '{}'";""".format(prefix + c, path)
              query_result = self.instance.client.query(query)
              if len(query_result.response):
                  prefix += c
                  break
          else:
              break
        print("Final string found: {}".format(prefix))


if __name__ == "__main__":
    f = FileReader()
    f.read("secret.txt")
```

https://github.com/osql/osql/issues/49

```
$ cat secret.txt
1d0puzt9880y25qquwamcbzng6jn5x2d2g
$ time python read-file-with-osquery.py
Prefix so far:
Prefix so far: 1
Prefix so far: 1d
Prefix so far: 1d0
Prefix so far: 1d0p
Prefix so far: 1d0pu
Prefix so far: 1d0puz
Prefix so far: 1d0puzt
Prefix so far: 1d0puzt9
Prefix so far: 1d0puzt98
Prefix so far: 1d0puzt988
Prefix so far: 1d0puzt9880
Prefix so far: 1d0puzt9880y
Prefix so far: 1d0puzt9880y2
Prefix so far: 1d0puzt9880y25
Prefix so far: 1d0puzt9880y25q
Prefix so far: 1d0puzt9880y25qq
Prefix so far: 1d0puzt9880y25qqu
Prefix so far: 1d0puzt9880y25qquw
Prefix so far: 1d0puzt9880y25qquwa
Prefix so far: 1d0puzt9880y25qquwam
Prefix so far: 1d0puzt9880y25qquwamc
Prefix so far: 1d0puzt9880y25qquwamcb
Prefix so far: 1d0puzt9880y25qquwamcbz
Prefix so far: 1d0puzt9880y25qquwamcbzn
Prefix so far: 1d0puzt9880y25qquwamcbzng
Prefix so far: 1d0puzt9880y25qquwamcbzng6
Prefix so far: 1d0puzt9880y25qquwamcbzng6j
Prefix so far: 1d0puzt9880y25qquwamcbzng6jn
Prefix so far: 1d0puzt9880y25qquwamcbzng6jn5
Prefix so far: 1d0puzt9880y25qquwamcbzng6jn5x
Prefix so far: 1d0puzt9880y25qquwamcbzng6jn5x2
Prefix so far: 1d0puzt9880y25qquwamcbzng6jn5x2d
Prefix so far: 1d0puzt9880y25qquwamcbzng6jn5x2d2
Prefix so far: 1d0puzt9880y25qquwamcbzng6jn5x2d2g
```

macOS 10.15 *Catalina*

```
"I'm sorry Fritz, I can't let you search that"
```



"Terminal.app" would like to access your Documents

An application in Terminal wants to use your documents.

"Terminal.app" would like to access your Downloads

An application in Terminal wants to use your downloads.

"Terminal.app" would like to access your Contacts

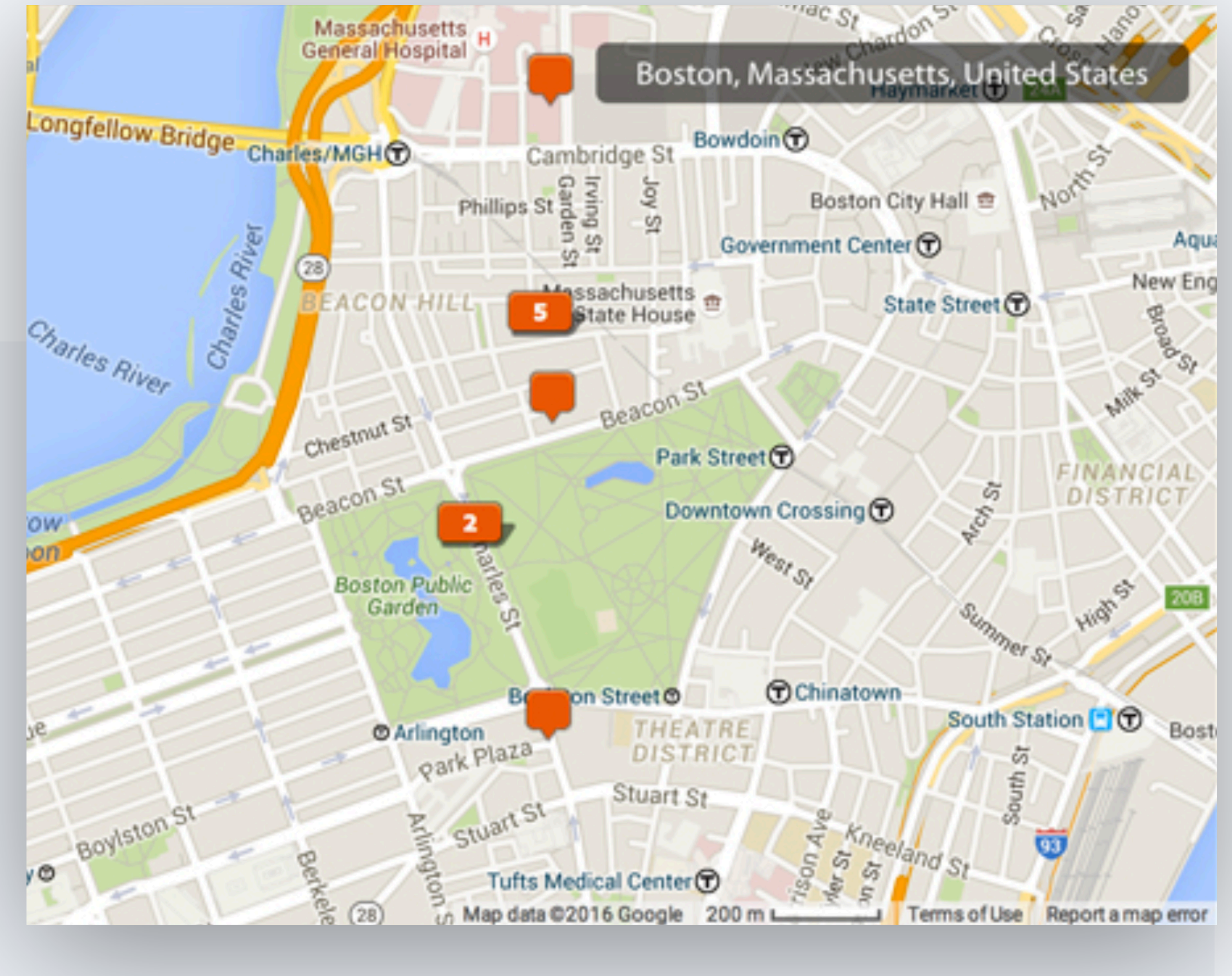
An application in Terminal wants to use your contacts.

“Terminal.app” would like to access your photos.

Find while the mdfind’ing is good.

Find photos taken within a given coordinate range

Adaptable to if you worked at an R&D facility with strict no photo policy



```
SELECT f.path,
       ROUND((f.size * 10e-7),2) size_megabytes,
       f.btime file_created_epoch,
       datetime(f.btime, "unixepoch") file_created
FROM file f, mdfind md USING (path)
WHERE md.query =
"kMDItemLatitude > 42.146512 && kMDItemLatitude < 42.48565 &&
kMDItemLongitude < '-70.74414' && kMDItemLongitude > '-71.3241'";
```

Find items downloaded from a specific source

Great for finding forgotten 2FA backup codes and data exports

```
SELECT f.path,
       ROUND((f.size * 10e-7),2) size_megabytes,
       f.btime file_created_epoch,
       datetime(f.btime, "unixepoch") file_created
FROM file f, mdfind md USING (path)
WHERE md.query =
"kMDItemWhereFroms =
'https://myaccount.google.com/_/two-step-verification/
backupcodes/download?*'";
```

Find items from a specific time range

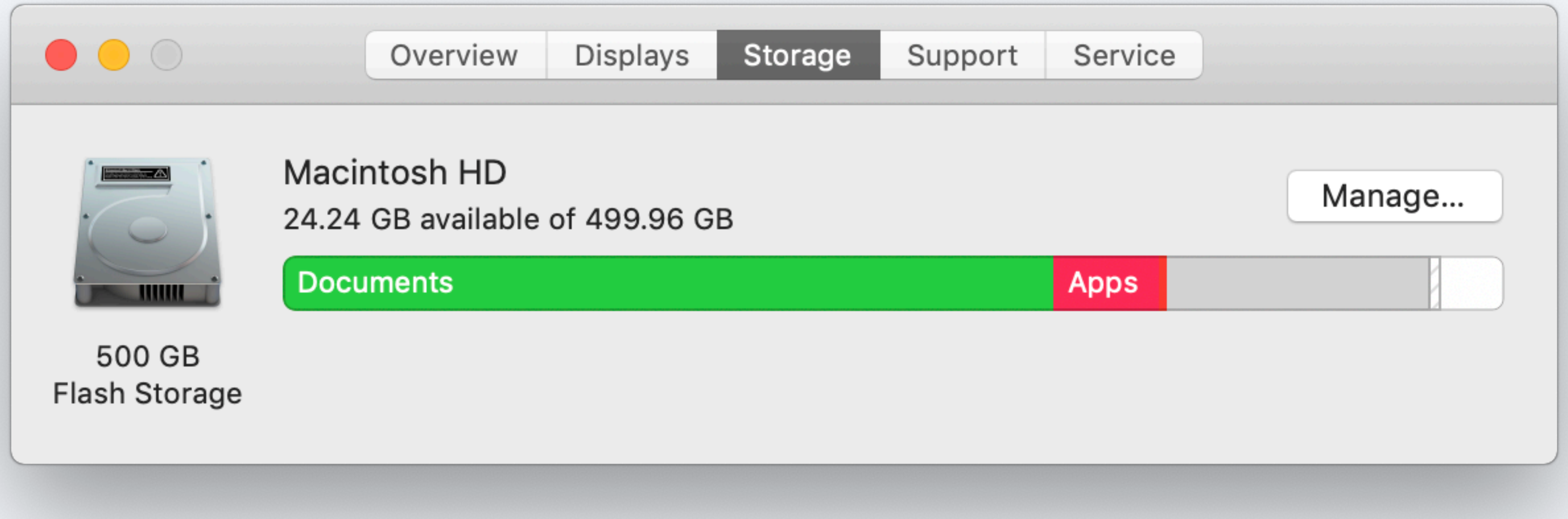
Good for cleaning up old export jobs after they are no longer being used

```
SELECT f.path,
       ROUND((f.size * 10e-7),2) size_megabytes,
       f.btime file_created_epoch,
       datetime(f.btime, "unixepoch") file_created
FROM file f, mdfind md USING (path)
WHERE md.query =
"(kMDItemFSCreationDate >= $time.today(-14)) &&
 (kMDItemFSName = '*.csv')";
```

Determine how much disk space a filetype is consuming

Useful if you are trying to rebuild Apple's storage widget

```
SELECT ROUND(CAST(SUM(f.size * 10e-10) AS decimal(20,2))) AS
sum_GB, COUNT(1) AS number_files
FROM file f, mdfind md USING (path)
WHERE md.query = "kMDItemFSName == '*.sketch'"
```



Macintosh HD

24.24 GB available of 499.96 GB

500 GB
Flash Storage

Joining on extended_attributes to display metadata contents

For example: Where files were downloaded from

```
SELECT mdfind.path,
       ROUND((f.size * 10e-7),2) AS size_megabytes,
       datetime(f.btime, 'unixepoch') AS file_created,
       ea.value AS download_source
       FROM extended_attributes ea
JOIN mdfind ON mdfind.path = ea.path
JOIN file f ON f.path = mdfind.path
  AND mdfind.query =
      "(kMDItemWhereFroms != '') && (kMDItemDateAdded >= $time.today(-14))"
  AND ea.key = 'where_from'
    GROUP BY ea.value
    ORDER BY f.btime DESC
    LIMIT 100;
```

Additional Reading

Kolide mdfind Blog Post:

https://blog.kolide.com/spotlight-search-across-every-mac-in-your-fleet-with-osquery-55789c765986

Apple Spotlight Query Syntax Documentation:

https://developer.apple.com/library/archive/documentation/Carbon/Conceptual/SpotlightQuery/Concepts/QueryFormat.html

Apple Spotlight Metadata Attributes Documentation:

https://developer.apple.com/library/archive/documentation/CoreServices/Reference/MetadataAttributesRef/Reference/CommonAttrs.html

Questions?

Thank you!

✉ **fritz** @ kolide.com

github.com / **Fritzx6**

fritz @ osquery Slack

KOLIDE