Negative Samples from Actual Code

A scan of this file should have zero sensitive findings.

False positive alerting can swamp your SecOps team and automated remediation. Regex and heuristic-based DLP solutions perform adequately in recognizing contrived negative test cases. However, when scanning actual code and data in a Git repository, they fail on a wide range of negative samples.

To solve this challenging problem, Nightfall AI has developed a natural language processing (NLP) based solution trained on 100s of thousands of lines of public code to recognize complex negative samples and eliminate false positives.

The 10 code snippets below demonstrate Nightfall's effectiveness in dramatically reducing false positives. The snippets are not contrived. The code or data in samples support commonly used software technologies. We grabbed them from public repos on GitHub. Each snippet contains data that, when scanned by regex and heuristics, looks to be a secret from a popular API service; however, in each case, Nightfall recognizes it as a negative sample and does not alert.

AWS MWS

A package lock json file contains a token matching an AWS marketplace secret.

```
{ "@babel/helper-optimise-call-expression": {
      "version": "7.12.10",
      "resolved:
"https://registry.yarnpkg.com/@babel/helper-optimise-call-expression/-/helper-optimise-call-expression-7.12.10.tgz#94ca4e306ee11a7dd6e9f42823e2ac6b49881e2d"
       "integrity: "sha512-4tpbU0SrSTjjt65UMWSrUOPZTsgvPgGG4S8QSTNHacKzpS51IVWGDj0yCwyeZND/i+LSN2g/O63jEXE
      "dev": true,
      "requires": {
        "@babel/types": "^7.12.10"
      }
    }
}
```

Confluent

A checksum in a configuration file that will likely generate Confluent key.

```
# Template file for 'mg'
pkgname=mg
version=20220614
revision=1
build_style=gnu-makefile
make_install_args="mandir=/usr/share/man"
hostmakedepends="pkg-config"
makedepends="libbsd-devel libmd-devel ncurses-devel"
short_desc="Micro GNU/emacs"
maintainer="Orphaned <orphan@voidlinux.org>"
license="Public Domain"
homepage="https://github.com/hboetes/mg"
distfiles="https://github.com/hboetes/mg/archive/${version}.tar.gz"
checksum=d09efde300c1027e0289de1a1d2da093ecce6b182250e03f05215fb044f24f8b
```

Datadog

An HTML snippet containing a file name and an HTML tag look like a Datadog token and key word.

```
<dl><dt>${keyword}</dt><dd>${description}</dd></dl><img
src="http://upload.wikimedia.org/math/0/c/b/0cb40a68f97d8790e8ffed3527adb639.png" />
```

Jira

A JSON snippet containing the substring “jira” and an alpha numeric string that looks like a Jira api token.

```
{ "_id" : { "$oid" : "5328e6f265f3f4d17089a11d" }, "domain" : "jira.mongodb.org", "banned_by" :
null, "media_embed" : {}, "subreddit" : "mongodb"}
```

Facebook

A GUID in a Unity YML file matches Facebook pattern and contains “Facebook” in context.

```
%YAML 1.1
%TAG !u! tag:unity3d.com,2011:
```

```
--- !u!114 &11400000
MonoBehaviour:
  m_ObjectHideFlags: 0
  m_PrefabParentObject: {fileID: 0}
  m_PrefabInternal: {fileID: 0}
  m_GameObject: {fileID: 0}
  m_Enabled: 1
  m_EditorHideFlags: 0
  m_Script: {fileID: 11500000, guid: 461bfb74be57ce74eac44498472f0ad0, type: 3}
  m_Name: FacebookSettings
  m_EditorClassIdentifier:
  selectedAppIndex: 0
  - 479945615483473
  appLabels:
  - Foundation
  cookie: 1
  logging: 1
  status: 1
  xfbml: 0
  frictionlessRequests: 1
  iosURLSuffix:
  appLinkSchemes:
  - list: []
```

An XCodeProject file snippet containing a several strings matching Facebook's OAuth token pattern and a substring "FB" in the near context.

```
4376945FEA4E29DB5A31133BBE2EA3A8 /* Frameworks */ = {
      isa = PBXFrameworksBuildPhase;
      files = (
            7C08906E192A9A36007307FB /* SAppDelegate.m in Sources */,
            7CD8E0AC193B3D5D006F0778 /* SCoreDataStore.m in Sources */,
            7CD8E0A2193B2234006F0778 /* SCoreService.m in Sources */);
      runOnlyForDeploymentPostprocessing = 0;
};
```

An XCodeProject file snippet containing a strings matching Facebook's OAuth token pattern and the string "Facebook" in the near context.

```
attributes = {
      LastUpgradeCheck = 0610;
      ORGANIZATIONNAME = Facebook;
      TargetAttributes = {
            00E356ED1AD99517003FC87E = {
                  CreatedOnToolsVersion = 6.2;
                  TestTargetID = 13B07F861A680F5B00A75B9A;
            };
      };
};
```

Plaid

This CSS contains a gif matching the token pattern for Plaid.

```
#shop #buttonz {
      width:669px;
      height:375px;
      background: transparent url(http://gf1.geo.gfsrv.net/cdn31/6d9df44ffadc89257e9485c3bb59da.gif)
no-repeat;
      color:#848484;
      font-size:11px;
      line-height:14px;
}
```

Salesforce

A file autogenerated by the USCDataScience Sentiment Analysis Parser containing a filename matching a the pattern of a Salesforce secret.

```
Content-Length: 432
Sentiment: sad
model: ../sentiment-models/src/main/resources/edu/usc/irds/sentiment/en-stanford-sentiment.bin
resourceName: C3FA6CFD12B5A0D51E9CFE1AE83AD1065221FDEC5F89BFDAAFF51076140CD2C4.sent
```

Square

Encoded data containing a string matching the format of a Square token.

```
<key>ANSIBrightYellowColor</key>
<data>
YnBsaXN0MDDUAQIDBAUGIiNYJHZlcnNpb25YJG9iamVjdHNZJGFyY2hpdmVyVCR0
AAGGoKYHCBMXGB9VJG51bGzVCQoLDA0ODxARElxOU0NvbXBvbmVudHNVTlNSR0Jc
b2xvclNwYWNlXxASTlNDdXN0b21Db2xvclNwYWNlViRjbGFzc08QKDAuOTUyOTQx
NSAwLjk3NjQ3MDU4ODIgMC42MTU2ODYyNzQ1IDFPECQwLjk0Mjg3OTIgMC45OTAx
NTU5IDAuNTI1OTMyODQ4NQAQAYACgAXSFA0VFlVOU0lDQ4ADgARPEQ8wAAAPMGFw
EAAAbW50clJHQiBYWVogB+AACgAZABcACgA6YWNzcEFQUEwAAAAAQVBQTAAAAAAA
AAAAAAAAAAAAAPbWAAEAAAAA0y1hcHBsAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAARZGVzYwAAAVAAAABiZHNjbQAAAbQAAAQYY3By
clNwYWNlWE5TT2JqZWN00hkaICFXTlNDb2xvcqIgHl8QD05TS2V5ZWRBcmNoaXZl
JVRyb290gAEACAARABoAIwAtADIANwA+AEQATwBcAGIAbwCEAIsAtgDdAN8A4QDj
7gDwAPIQJhArEDYQPxBMEE8QXBBlEGoQchB1EIcQihCPAAAAAAAAAgEAAAAAAAAA
AAAAAAAAAAAAAAAAEJE=
</data>
```

Twitter

This CSS snippet contains an encoded image matching the OAuth2 token pattern for Twitter.

```
.dx-datagrid .dx-datagrid-rowsview .dx-virtual-row > td:not(.dx-datagrid-group-space):before
     { display: block;
     height: 100%;
     content: '';
     background-image:
url(
AHf5/id9+wxLjX8efvl+WQQRBEEEQRBBEEQQBBEEQQAAAAAAAAAAflRuLmaxdSIIggiCIIIgiCAIIgiCIAgAyyUfMBs3BjHUBUEQQRBEEA
QRBEEEQRAE6cKyNQAAAAAAAADQQuySw9aXDl6gsAYkCIIIgiCCIIggCCIIgiAIcILkA2Y7xR5HM0MEQRBBEEQQBBEEQQRBEAQBAAAAAAAA
AHq78l5WMlsngiCIIAgiCIIIgiCCIAiCALDcrQfMDHUEEQRBBEEQQRAEQQRBEEEAAAAAAAAAALZoecCsqo4NYutEEAQRBEEEQRBBEEQQBE
EQAJbresBsGOoIIgiCIIggCCIIggiCIIIAAAAAAAAc5iPAAIcKFwANDofzAAAAAElFTkSuQmCC);
     background-repeat: no-repeat repeat;}
```