Academic License for UNITN DISI Security Databases for Scientific, Non-Profit, Non-Commercial Purposes

The Dipartimento do Ingegneria e Science dell'Informazione, University of Trento
Via Sommarive 9, I-38123 Trento, Italy

Hereinafter UNITN by means of

Fabio Massaccci (Professor)		Niculae Sebe (Head of Department)
UNITN PROVIDING SCIENTIST		UNITN REPRESENTATIVE

-
- Hereby grants a free non-exclusive non-transferable license for the Security Dataset(s)
- (delete as appropriate)
    - NVD-EDB-EKITS-SYM
    - FFV-GCV-IEV-ASV
    - ESEJ
- to:

DEPARTMENT, ORGANIZATION
ADDRESS, COUNTRY

Hereinafter RECIPIENT by means of

Name (Position)		Name (Position)
RECIPIENT SCIENTIST(S)		RECIPIENT REPRESENTATIVE

Under the terms and conditions stated herein:

1. The RECIPIENT asks rights to access the DATASETS mentioned above among the following ones

- **NVD**: is the reference database for the population of vulnerabilities. It collects the data from the *National Vulnerability Database* from NIST.[1]

Via Sommarive, 9. – 38123 Trento, Italy - Tel. +39 0461-28-3936

- **EDB** is the reference database for public (proof-of-concept) exploits. It collects the data from the *Exploit-DB* web site.
  - **EDB-files** contains all actual exploits referenced in EDB, categorized by platform.
- **EKITS** is a database of vulnerabilities and exploits traded in the black markets. We have built an update infrastructure that allows us to keep our database well ahead of any public source on such vulnerabilities publicly available (such as Contagio's Exploit Pack Table).
- **SYM** is a database of vulnerabilities exploited in the wild as reported by Symantec's sensors worldwide. This dataset is a collection of publicly available vulnerability data through Symantec's *Threat Explorer* and Attack Signatures websites.
- **FFV** collects the vulnerabilities of the *Firefox* browser. It is the most comprehensive database. It integrates the Mozilla Foundation Security Advisory (MFSA) bulletin, the Mozilla Bugzilla bugtracker and the NVD.
- **GCV** reports the vulnerabilities of the *Google Chrome* Browser extracted from Chrome Issue Tracker, integrated with the NVD to reconstruct affected versions and checked for consistency with the code distribution. It does not include all vulnerabilities of the browser as some of the third party software such as WebKit are only partly included.
- **IEV** lists the vulnerabilities for *Internet Explorer* extracted from the Microsoft Security Bulletin and integrated with the NVD to reconstruct affected versions.
- **ASV** Vulnerabilities of the Apple Safari Web Browser extracted from the Apple Knowledge Base and integrated with the NVD to reconstruct affected versions.
- **ESEJ** is the list of vulnerabilities in Google Chrome and Mozilla Firefox along with ranges of major versions affected by each vulnerability. For each vulnerability, the dataset contains two affected version ranges: (1) vulnerable versions according to the NVD; (2) vulnerable versions based on the vulnerable code evidence (identified by our algorithm).

A description of the tables and entries in of the DATASETS is provided as ANNEX A.

2. The RECIPENT intends to use the dataset for the following scientific, non-profit, non-commercial purposes

   Write here the broad goal of the research

   The RECIPENT agrees that such purposes, the name of the RECIPIENT's scientist(s) and affiliation, and any publications by the RECIPIENT that uses the DATASETS will be listed by UNITN on the web site http://security-data.disi.unitn.it.

3. "MODIFICATIONS" of DATASETS is software or database tables or database columns created by RECIPIENT which contains/incorporates DATASETS or a part thereof or SQL code and data table values of DATASETS or a part thereof.

4. The RECIPIENT is free to make MODIFICATIONS of DATASETS by making in-house copies of tables and to modify copied tables, for example by the addition of columns and changing of data. The RECIPIENT is free to use the database in-house as he/she wishes and to create logical objects containing original and/or derived data so long as such use does not violate terms specified in this agreement.

Via Sommarive, 9. – 38123 Trento, Italy - Tel. +39 0461-28-3936

5. The license is free of charge so long as DATASETS, or any component of DATASETS, or any derivative work that includes or depends on DATASETS in whole or in part, is used for scientific, non-profit, non-commercial use only. Any other use of DATASETS and use of MODIFICATIONS of DATASETS for other purposes, alone or integrated into other databases or software, requires prior written consent by UNITN.

6. RECIPIENT shall have the right to publish its findings and results related to DATASETS, provided that UNITN researchers are cited as the source of DATASETS and the references below are cited in the publication. The published references for DATASETS are listed below:
    a. **NVD, EDB, EKITS, SYM:**
        i. **Luca Allodi and Fabio Massacci. 2014. Comparing Vulnerability Severity and Exploits Using Case-Control Studies. *ACM Transactions on Information and System Security*. 17(1), 20pp, 2014. DOI=10.1145/2630069**
        ii. Its preliminary version whenever historical attribution is important:
            1. Luca Allodi and Fabio Massacci. 2012. A preliminary analysis of vulnerability scores for attacks in wild: the ekits and sym datasets. In *Proc. of the 2012 ACM Workshop* BADGERS '12. DOI=10.1145/2382416.2382427
    b. **FFV, IEV, ASV, GCV:**
        i. **Fabio Massacci and Viet Hung Nguyen. An Empirical Methodology to Evaluate Vulnerability Discovery Models. *IEEE Transactions on Software Engineering* 40(12):1147-1162, 2014. DOI=10.1109/TSE.2014.2354037**
        ii. Its preliminary versions whenever historical attribution is important:
            1. Fabio Massacci, Stephan Neuhaus, Viet Hung Nguyen. After-Life Vulnerabilities: A Study on Firefox Evolution, its Vulnerabilities and Fixes. In *Proc. of the 3rd Int. Symp. on Engineering Secure Software and Systems (ESSoS'11)*, 2011. Springer Verlag. DOI=10.1007/9783642191251.
            2. Viet Hung Nguyen, Fabio Massacci. An Independent Validation of Vulnerability Discovery Models. In *Proc. of the 7th ACM Symp. ASIACCS'12,* 2012. DOI=10.1145/2414456.2414458.
    c. **ESEJ:**
        i. **Viet Hung Nguyen, Stanislav Dashevskyi, Fabio Massacci. 2015. An Automatic Method for Assessing the Versions Affected by a Vulnerability. In *Empirical Software Engineering* (to appear). DOI=10.1007/s10664-015-9408-2**
        ii. Its preliminary version whenever historical attribution is important:
            1. Viet Hung Nguyen, Fabio Massacci. The (Un)Reliability of Vulnerable Version Data of NVD: an Empirical Experiment on Chrome Vulnerabilities. In *Proc. of the 8th ACM Symp. ASIACCS'13*, 2013. DOI=10.1145/2484313.2484315.

The RECIPIENT undertakes to notify UNITN of the existence of the publication by

Via Somm
461-28-3936

email at security-data@disi.unitn.it.

7. It is the responsibility of the RECIPIENT to read the articles mentioned in Article 6 in order to understand the scientific limitations of the DATASETS.

8. No property rights with respect to DATASETS shall transfer to RECIPIENT through this agreement. UNITN may demand compensation for uses other than those granted in this license according to article 15.

9. The RECIPIENT, the user and any research assistants, co-workers or other workers who may use DATASETS agree to not give the data to third parties or grant licenses which include DATASETS, alone or integrated into other databases, to third parties without prior consent of UNITN.
   RECIPIENT may not place DATASETS on public servers unless prior agreement is given by UNITN.
   UNITN may demand compensation from RECIPIENT for transfer of information and licenses granted to third parties by RECIPIENT.

10. The RECIPIENT undertakes to refer any requests by third parties for the provision of DATASETS to UNITN at security-data@disi.unitn.it

11. Where the research involving DATASETS results in an invention or patentable MODIFICATION of DATASETS, RECIPIENT and its Researcher/s shall promptly disclose this development to UNITN. RECIPIENT and UNITN shall decide in common about the inventorship, taking due consideration UNITN's contribution to the invention through DATASETS. Decisions about further proceedings, such as filing of a patent application or exploitation, shall be made after inventorship is determined.

12. DATASETS is not guaranteed as suitable for use with any application. UNITN gives no warranty express or implied of any kind with regard to the distribution, content or operation of DATASETS, in particular but not limited to any warranty of suitability or fitness for any purpose. UNITN will not assume liability for damages occurred through the use of DATASETS.

13. UNITN will disclose any known rights of third parties to DATASETS to the best of its knowledge. However, UNITN does not warrant for any possible infringement of rights. RECIPIENT has to acquire on his own all necessary licenses for supporting DATASETS if not otherwise agreed in writing.

14. This license covers usage of DATASETS only, and does not of itself entitle the RECIPIENT to any support from UNITN in the installation, maintenance or use of DATASETS. Distributions of DATASETS and updates to it will be accompanied by a document clarifying issues related to support. RECIPIENT will inform UNITN of any defects found in DATASETS by email at security-data@disi.unitn.it.

15. In case the DATASETS is or will be under the control of RECIPIENT before this

Via Sommarive, 9. – 38123 Trento, Italy - Tel. +39 0461-28-3936

agreement is signed UNITN gives consent to use of DATASETS under the condition of RECIPIENT'S prior consent to this agreement.

16. This agreement may be terminated by either party with two months notice.

17. Disputes and requests for compensation related to this license will be adjudicated by the International Court of Arbitration of the International Chamber of Commerce (ICC). Both parties agree to consider ICC arbitration final and agree to abide by its ruling and pay its awards without further appeals.

Fabio Massaccci (Professor)		Niculae Sebe (Head of Department)
UNITN PROVIDING SCIENTIST		UNITN REPRESENTATIVE
Date and signature	Stamp of the Organization	Date and signature

Name (Positions)		Name (Position)
RECIPIENT SCIENTIST(S)		RECIPIENT REPRESENTATIVE
Date and signature	Stamp of the Organization	Date and signature

Via Sommarive, 9. – 38123 Trento, Italy - Tel. +39 0461-28-3936

ANNEX A – DATASETS TABLES AND ATTRIBUTES

1. **NVD** Note that each entry in the NVD dataset **does not correspond to a vulnerability**. A vulnerability ID can be associated with more than one software or vendor. The same ID can be reported in different tuples.

a. **CVE_ID:** Id of the vulnerability.
b. **Pub_date:** First publication date of the vulnerability
c. **Mod_date:** Date of last update to the entry
d. **CVSS_score:** CVSS v2.0 Risk Score of the vulnerability
e. **CVSS_Imp:** CVSS v2.0 Impact score of the vulnerability
f. **CVSS_Expl:** CVSS v2.0 Exploitability score of the vulnerability
g. **CVSS_AV:** CVSS Exploitability assessment: Access Vector
h. **CVSS_AC:** CVSS Exploitability assessment: Access Complexity
i. **CVSS_Au:** CVSS Exploitability assessment: Authentication
j. **CVSS_Conf:** CVSS Impact assessment: Confidentiality
k. **CVSS_Integ:** CVSS Impact assessment: Integrity
l. **CVSS_Avail:** CVSS Impact assessment: Availability
m. **Aff_Sw:** Software affected by the vulnerability
n. **Vendor:** Vendor of the software
o. **Description:** English description of vulnerability

2. **EDB**
(*) in EDB-files only

a. **E-id:** Exploit-DB record ID
b. **Cve-id:** CVE_ID of vulnerability to which the exploit refers
c. **Date:** date of emission of exploit
d. **Osvb-id:** ID to third-party vulnerability database: OSVDB
e. **(*) File:** Path to the exploit
f. **Description:** Description of exploit
g. **Author:** name of the researcher who published the exploit
h. **Platform:** operating system of the vulnerability/exploit
i. **Type:** type of exploit (e.g. remote, webapp, denial-of-service)
j. **Port:** remote access port to the vulnerability as reached by the exploit (iff *type==remote*)

3. **EKITS**

a. **Ek_id:** Id of exploit kit
b. **E_name:** exploit kit name
c. **Version:** version of exploit kit
d. **Date:** date of release of exploit kit on the black markets (month)
e. **Price:** advertised price
f. **Per:** license duration (year,month,week)
g. **Service1:** Services sold alongside the product (not available for all ekits)
h. **Service2:** Services sold alongside the product (not available for all ekits)
i. **Service3:** Services sold alongside the product (not available for all ekits)

Via Sommarive, 9. – 38123 Trento, Italy - Tel. +39 0461-28-3936

j. **Cve_id:** CVE_ID of vulnerability exploited by the kit
k. **P_source:** primary source of information
l. **S_source:** secondary source
m. **Notes:** english notes on the ekit/advertisement/services

4. SYM (malware + network attacks)

a. **attack_ID:** ID of attack referenced by Symantec (network attacks table)
b. **threat_ID:** ID of malware referenced by Symantec (malware table)
c. **Type:** where in the text the vulnerability is mentioned (i.e. description of attack or references)
d. **CVE:** CVE_ID of vulnerability
e. **String:** name of to the attack on Symantec's website

5. FFV

a. **bugID:** the identifier of a bug responsible for this vulnerability.
b. **cve:** the identifier of an CVE entry referring to this vulnerability.
c. **mfsa:** the identifier of an MFSA entry referring to this vulnerability.
d. **bugDate:** the date when the corresponding bug is filed to Bugzilla.
e. **cveDate:** the date that the corresponding CVE is filed to NVD.
f. **minVersion:** the earliest major version that this vulnerability affects to.
g. **maxVersion:** the latest major version that this vulnerability affects to.

6. GCV

a. **bugID:** the identifier of a bug responsible for this vulnerability.
b. **cve:** the identifier of an CVE entry referring to this vulnerability.
c. **bugDate:** the date when the corresponding bug is filed to ChromeIssueTracker.
d. **cveDate:** the date that the corresponding CVE is filed to NVD.
e. **minVersion:** the earliest major version that this vulnerability affects to.
f. **maxVersion:** the latest major version that this vulnerability affects to.
g. **codeMinVersion:** the earliest major version where the vulnerable code footprint is found.
h. **codeMaxVersion:** the latest major version where the vulnerable code footprint is found.

7. IEV

a. **cve:** the identifier of an CVE entry referring to this vulnerability.
b. **mssb:** the identifier of an MS Security Bulletin entry referring to this vulnerability.
c. **cveDate:** the date that the corresponding CVE is filed to NVD.
d. **minVersion:** the earliest major version that this vulnerability affects to.
e. **maxVersion:** the latest major version that this vulnerability affects to.

8. ASV

a. **cve:** the identifier of an CVE entry referring to this vulnerability.

Via Sommarive, 9. – 38123 Trento, Italy - Tel. +39 0461-28-3936

b. **akb:** the identifier of an Apple Knowledge Base entry referring to this vulnerability.
c. **cveDate:** the date that the corresponding CVE is filed to NVD.
d. **minVersion:** the earliest major version that this vulnerability affects to.
e. **maxVersion:** the latest major version that this vulnerability affects to.

8. ESEJ

a. cve: the identifier of a CVE entry refferring to this vulnerability.
b. bugID: the identifier of a bug responsible for this vulnerability (Bugzilla or Chrome issue tracker).
c. cveDate: the date that the corresponding CVE is filed to the NVD.
d. minVer: the earliest major version that this vulnerability affects.
e. maxVer: the latest major version that this vulnerability affects.
f. bugFix: the bug fix commit that was successfully located.
g. esminVer: the earliest major version that this vulnerability affects, according to the code evidence identified by our algorithm.
h. esmaxVer: the latest major version that this vulnerability affects, according to the code evidence identified by our algorithm.

Via Sommarive, 9. – 38123 Trento, Italy - Tel. +39 0461-28-3936

This is the human readable summary of your rights and obligations. It is provided for your convenience only. The formal text of the agreement is the only binding document.

- You can
  1. share these datasets in whatever format with any member of your institution—faculty, administration, students, research associates in the case of universities, and employees in the case of government ministries and research organizations;
  2. use these datasets in creative ways for scientific, not-profit, non-commercial use including publications under the terms of the agreement.

- You cannot
  1. post any of these datasets on your website such that it becomes available to non-members of your institution, or make copies that circulate outside of your institution;
  2. use it for profit or commercial purposes unless agreed in writing.

- You agree to
  1. cite the appropriate reference work in all your publications that make use of the datasets or its derivatives;
  2. provide us the information on the publication where you used the data by email to **security-data @disi.unitn.it** for the purposes of posting it on our web site **http://security-data.disi.unitn.it** with your name and affiliation**;**
  3. refer to us any person or organization outside your institution who would like to use the data.

- You are aware that
  1. other parties may have rights or set licensing obligations in some of the data contained in the datasets and it is up to you to obtain permissions from these parties if needed;
  2. the datasets may contains errors or may be unfit for your purposes and we bear no liability for any problem you might encounter.

Via Sommarive, 9. – 38123 Trento, Italy - Tel. +39 0461-28-3936