OWASP DEF (Data Exchange Format)

Background:

At the moment exchanging data between pentest tools it is far too difficult.

So ... the purpose of this project is to define a simple, open format for exchanging data between pentest tools!

Involvement is encouraged, so if you would like to contribute to this project then please join the mailing list (https://lists.owasp.org/mailman/listinfo/owasp-data-exchange-format) and / or contact one of the project leaders.

Theres also a Google Code project http://code.google.com/p/owasp-def/ which we're using to store things like example formats used by pentest products. Contact Simon or Dinis to get commit access to this project.

Requirements:

The format must be open, and licensed so that it can be adopted by all products, whether open, closed, free or commercial.

It must be as simple to adopt as possible, and ideally based on existing open formats.

The OWASP DEF (from now on refrenced to as DEF) is a work in progress.

This document will describe the DEF format.

1st. Let me say that i suck at writing documentation so bare with me on this.

q. What is the purpose of OWASP DEF?

a. Well the purpose of the OWASP DEF is to have a unified format that should cover any / all information related to Dynamic, Static, Information gathering and Attack / Brute force scanning software.

q. Will this format be free for anyone to use?

a. Yes it will be free for any one to adapt and use in their projects.

OK lets descripe the format as it looks at the time of writing this.

This is the Root, and it should hold the information about the version of DEF.

<DEF SpecVersion=*"0.1"*>

Then there is the Session-refrence, this should hold an overall refrence the the session that is contained in this XML file.

<Session-Reference>Scan specific reference</Session-Reference>

Then there is the Date-Time refrence, this is when the scan was done, atm. in any date time format.

<Date-Time>Date and time the session was started</Date-Time>

Then there is the Scan refrence, this holds the type of scan that is done, and atm. it can be either dynamic, static or info. And I will start with describing the Dynamic format.

<Scan type=*"dynamic"*>

In the Dynamic section of this format we will have some basic information about the target that we scanned, for this we have a Host section, wich holds the name of the host, and the IP Address.

<Host name=*"Hostname"* ip-address=*"Either IPv4 or IPv6 Address"*>

Then there is the Port section, this is used for the basic information about the protocol and port the scan was done against.

<Port protocol=*"The Protocol used"* portid=*"The Port Number"*>

Then there is the Service section, and this holds some basic information about the service name, service product and the service product version.

<Service name=*"Name of the Service"* product=*"Product Name"* version=*"Product Version"* />

Then there is the Software-name section, and this holds the name of the Software that did the scan.

<Software-Name>Name of the tool that found the issue</Software-Name>

Then there is the Software-version section, and this should hold the version of the software that did the scan.

<Software-Version>Version of the tool that found the issue</Software-Version>

Then there is the Software-arguments, and this should hold the arguments that where used to preform the scan.

<Software-Arguments>Arguments used to perform the scan</Software-Arguments>

Then there is the Vulnerability section, and this holds the severity of the current vulnerability, wich can be one of: Critical, High, Medium, Low, Informational, (and Best Practies (At the time of writing this it is not decided if the latter should be included or not, this is dependent on how many of the softwares that actually uses this)(Note: Should this be converted into a number representation, Critical = 1, High = 2, Medium = 3, Low = 4, Informational = 5, and so on).

<Vulnerability Severity=*"The Severity"*>

Then there is the Finding section, and this hold the NativeID and IdentifiedTimestamp, wich are the unique plugin / test ID from the software, and a time stamp of when this vulnerability was found.

<Finding NativeID=*"The internal Test ID"* IdentifiedTimestamp=*"DateTime stamp for when we found this vulnerability"* uniqueID=*"The Software unique ID for this Finding"*>

Then there is the Summary section, and this holds a short one line summary of the vulnerability.

<Summary>A sort (one line) description</Summary>

Then there is the Description section, and this hold a longer description of the vulnerability.

<Description>More detailed description</Description>

Then there is the Confidence section, this hold the confidence level that the software have about this vulnerability, and can be a number between 0 and 10, where 0 equal not provided, and 10 equals the highest confidence we have.

<Confidence>0 thru to 10</Confidence>

Then there is the Background section, and this hold some background information about the vulnerability.

<Background>More info on the type of issue</Background>

Then there is the Remediation section, and this holds the remediation information for the vulnerability.

<Remediation>Advise on how to fix the issue</Remediation>

Then there is the Further-information section, and this holds further information about the vulnerability.

<Further-Information>

Then there is the Further-info section, and this holds the further information, wich could a link to the OWASP top 10 page for the vulnerability, or any other page with information regarding the vulnerability.

<Further-Info>More information about this specific issue</Further-Info>

Then there is the Classifications section, and this hold the classification information.

<Classifications>

Then there is the Classification section, and this hold the classification type, id and url to the specified id for this vulnerability, this could as the example shown be the CWE classification.

<Classification type=*"The Classification System"* id=*"Classification ID"* href=*"The URL to the Classification description"*>The Title for the Clasasification</Classification>

Then there is the Page section, and this hold the page information that vulnerability was found on.

<Page>

Then there is the Page-reference, and this hold the page specific information.

<Page-reference>Product specific reference e.g. Page Title</Page-reference>

Then there is the URL section, and this hold the URL the vulnerability was found at.

<URL>The actual URL</URL>

Then there is the Method section, and this hold the HTTP method used.

<Method>HTTP method (GET, POST, etc)</Method>

Then there is the HTTPVersion section, and this holds the version of the HTTP communication.

<HTTPVersion>The HTTP Version</HTTPVersion>

Then there is the StatusCode section, and this holds the Status code that the request returned.

<StatusCode>The HTTP Status code</StatusCode>

Then there is the Language section, and this hold the information about the detected language of the page / web application.

<Language>The detected Language of the Web Application</Language>

Then there is the Parameters section, and this hold the Parameters used to exploit the and find the vulnerability.

<Parameters>

Then there is the Parameter section, and this hold the parameter(s) used to find the vulnerability.

<Parameter>The parameter</Parameter>

Then there is the Request-response section, and this hold the Request and Response sent to and from the Web application.

<Request-Response>

Then there is the Request section, and this hold the information about the Request that was sent to the web application. (NOTE. Should this be base64 encoded for this format?)

<Request>

Then there is the Raw section, and this holds the raw request as it was sent.

<Raw>The RAW HTTP Request</Raw>

Then there is the Headers section, and this holds the Headers information.

<Headers>

Then there is the Headers Data section, and this holds the header data.

<Data name=*"The name of the Header Data"* value=*"The value for the Header Date"*/>

Then there is the Cookie section, and this holds the Cookie information.

<Cookie>

Then there is the Cookie Data section, and this holds the Cookie data.

<Data name=*"The name of the Cookie Data"* value=*"The value for the Cookie Date"*/>

Then there is the Additional-RequestData section, and this holds the Additional Request Data.

<Additional-RequestData>

Then there is the Additional-RequestData Data section, and this holds the Additional-RequestData Data.

<Data name=*"The name of the Additional Data"* value=*"The value for the Additional Date"*/>

Then there is the Response section, and this holds the Response data.

<Response>

Then there is the Raw section, and this holds the raw response data.

<Raw>The RAW HTTP Response</Raw>

Then there is the Headers section, and this holds the Headers information.

<Headers>

Then there is the Headers Data section, and this holds the header data.

<Data name=*"The name of the Header Data"* value=*"The value for the Header Date"*/>

Then there is the Cookie section, and this holds the Cookie information.

<Cookie>

Then there is the Cookie Data section, and this holds the Cookie data.

<Data name=*"The name of the Cookie Data"* value=*"The value for the Cookie Date"*/>

Then there is the Additional-ResponseData section, and this holds the Additional Response Data.

<Additional-RequestData>

Then there is the Additional-ResponseData Data section, and this holds the Additional-ResponseData Data.

<Data name=*"The name of the Additional Data"* value=*"The value for the Additional Date"*/>