

# Connecting the AV industry

Igor Muttik, McAfee Avert Labs

**IEEE, Industry Connections Security Group**

21<sup>th</sup> September 2009



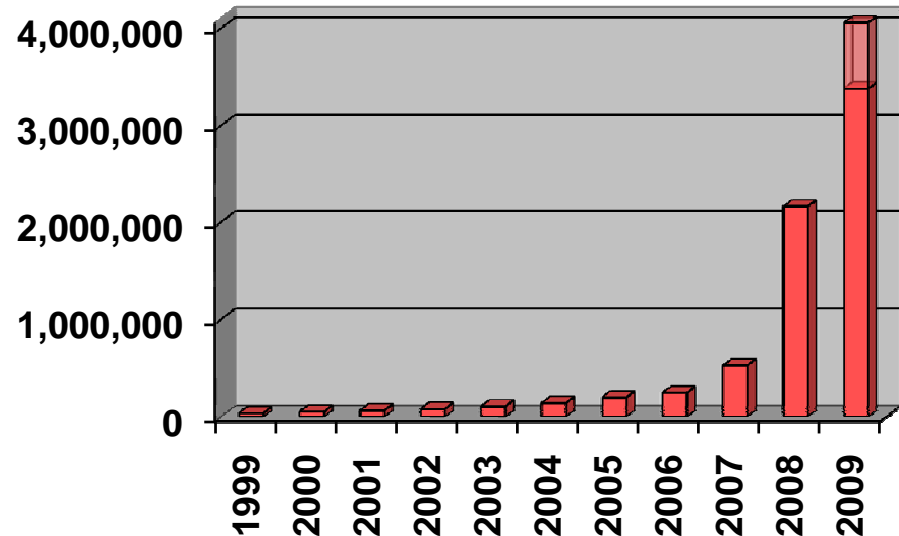
# Agenda

- The problem
- Re-inventing the wheel?
- Introducing ICSG
- Malware Working Group
- XML meta-data sharing format
- Main concepts
- Details of XML standard proposal
- Questions



# The Problem

- Attackers have shifted away from mass distribution of a small number of threats to micro distribution of millions of distinct threats
- The security industry still by and large responds to threats in their individual silo's with 'limited' operational & cross industry co-operation
- Many in the security industry want to pool their experience and resources in response to this systematic and rapid rise in new malware



# Re-Inventing the Wheel ?

- Lots of great examples of working groups focused on specific aspects of security intelligence , incident response, testing, best practices & policies



- ❖ APWG
- ❖ ASC
- ❖ AMTSO
- ❖ CARO
- ❖ Others

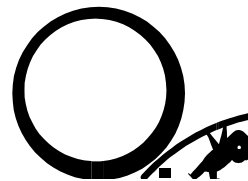
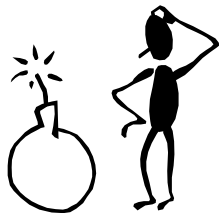
*Anti-Phishing Intelligence & Best Practices*  
*Anti-Spyware Intelligence and Best Practices*  
*Anti-Malware Testing Standards and Best Practices*  
*Computer Anti-Virus Research Organization*  
*AVAR, EICAR, AVPD, MWAAG , FIRST, CWG, etc*



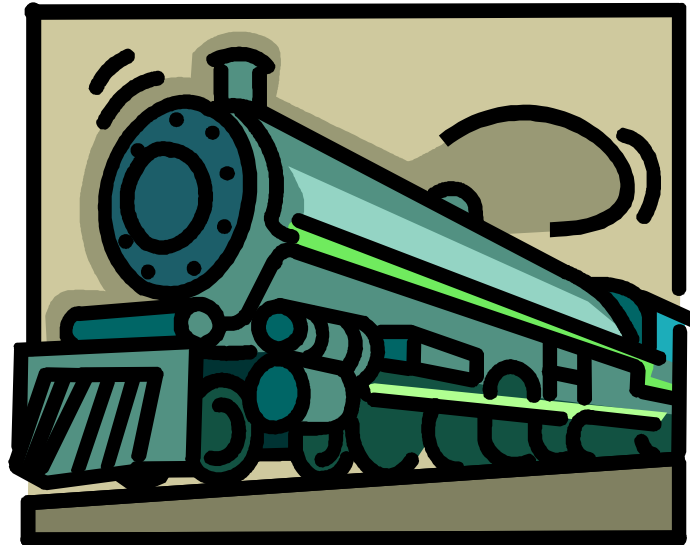
# Re-Inventing the Wheel ?

- Lots of great examples of working groups focused on specific aspects of security intelligence , incident response, testing, best practices & policies

*However, this co-operation typically has not been **standardized** or **documented** in a format that lends itself to **systematic improvement** in operational efficiency, or **visibility and review** by people **outside the vertical industries** and in many cases that was **not their mandate***



# Introducing



## ICSG

*Industry Connections Security Group*

*Google for: "IEEE ICSG"*

*[joinicsg@ieee.org](mailto:joinicsg@ieee.org)*

# ICSG Goal & Structure

- Established under the umbrella of the the IEEE-SA Standards Association
- Facilitate the pooling of industry experience and resources
- A forum for development of proposed standards and best practices related to computer security
- ICSG was started by AVG, McAfee, Microsoft, Sophos, Symantec and Trend Micro, but open to others
- *Goes beyond Malware Issues !*



# Why the IEEE?

- Need to reach outside the traditional groups to pool as many different contributors as possible
- IEEE is a recognized brand known to deliver standards
- The existing infrastructure of the IEEE allowed us to start working on the crux of the issues, instead of wasting time on the org side
- We leverage the brand to attract the non-traditional players into the pool





# What to focus on ?



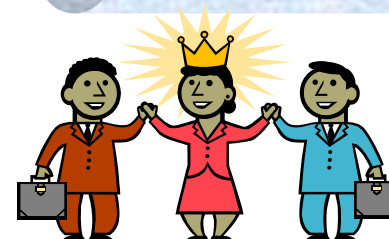
*How do we improve the efficiency of the collection & processing of the millions of malware file samples we all handle each and every month ?*

## Malware Working Group

- Focused on development of a XML based metadata sharing standard to augment existing malware sample sharing
- Website & Wiki located at <http://ieee.sanasecurity.com>
- Home for the schema for validation purposes  
<http://ieee.sanasecurity.com/schema/1.0/metadataSharing.xsd>

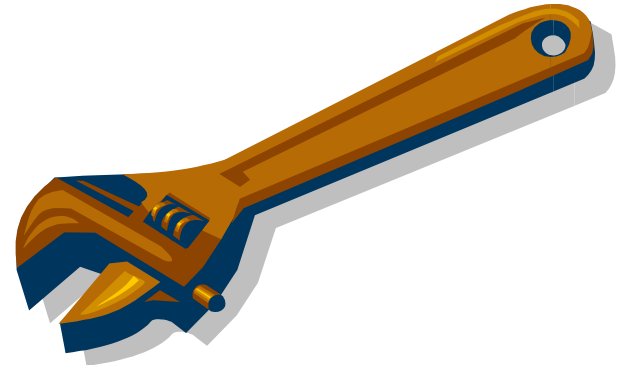
### Additional Contributors

*Support Intelligence  
Immunit  
Team Cymru  
ShadowServer  
Arbor Networks  
Cisco  
WebSense  
AV-Test  
SonicWall  
Avira  
and many others..*



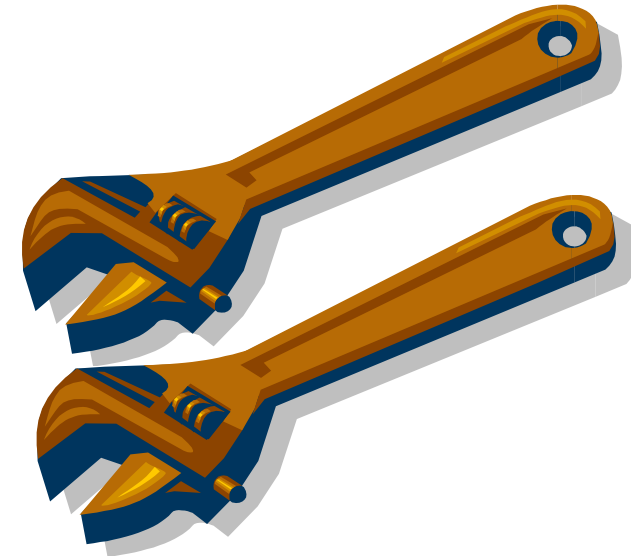
# The Use Cases (1)

- Prioritize samples in a queue
  - E.g. by commonality
  - By geography
- Sharing clean files is problematic but meta-data is OK
- Weed out improper samples
  - Corrupt
  - Clean from malware test set
  - Malicious or unwanted from a clean set
- Build knowledge about a sample (or a collection) before you get it



## The Use Cases (2)

- Grouping non-static malware
  - Polymorphic
  - Server-side polymorphics
  
- Create better test set
  - E.g. with old malware excluded
  - With more common malware
  
- Link malware and URLs
  - “fast-flux” URLs
  - domains operated by entities (malware-writing groups)
  - URLs <-> IPs
  
- More...



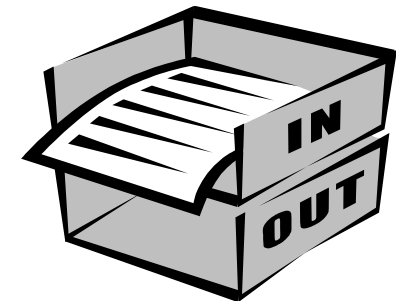
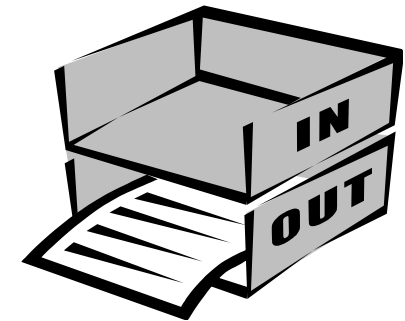
# XML schema

- Why XML?
  - XML is friendly for RDBMS
  - Friendly for humans too
  - Extendable
  - Common and supported everywhere



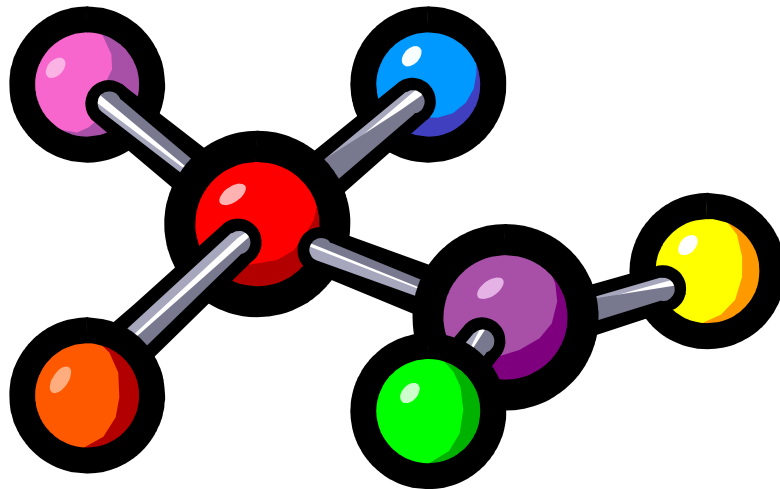
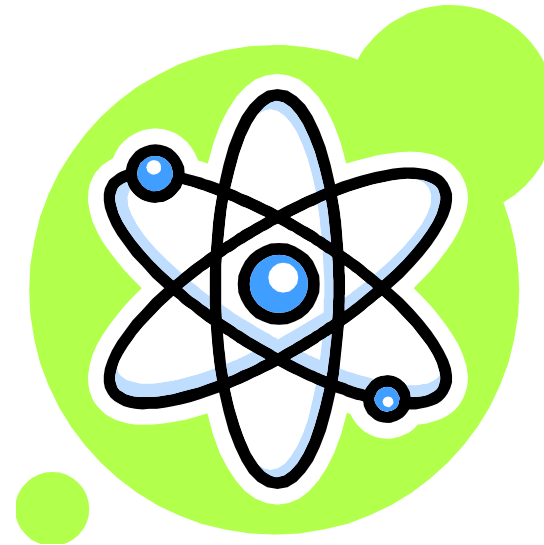
# How?

- Outgoing XMLs
  - Along with collection distribution (daily or ad-hoc)
  - RAR-archived (for integrity checking)
  - PGP-encrypted (for authenticated access)
  - Distributed via FTP/SFTP/HTTP/HTTPS (same as already used for collection distribution)
  - At least four companies already started
  
- Incoming XMLs
  - Level of details will depend on the source
  - Several entities already consume XML feeds



# Atomic

- ▣ Simple basic elements



- ▣ To cleanly link “atomic” elements with “relationships”

# Main concepts (1)

- Header
  - Source of meta-data
  - Author
  - Timestamp
  
- Object1..ObjectNN
  - File
  - URI, domain, service (protocol:port)
  - Environment
  - Registry
  - Entity
  
- Classification1.. ClassificationMM
  - Clean/dirty/unwanted
  - Malware category
  - Detection name, product, company





# Example (minimal)

```

<?xml version="1.0" encoding="UTF-8"?>
<malwareMetaData xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="http://xml/metadatasSharing.xsd metadatasSharing.xsd"
  xmlns="http://xml/metadatasSharing.xsd" version="1.0" id="1234">
  <company>McAfee</company>
  <author>Raiden</author>
  <comment>This is minimal - just some files</comment>
  <timeStamp>2008-11-25T21:34:56</timeStamp>
  <objects>
    <!-- files -->
    <file id="2f437c1c8f73c2d6ffbb6214d3f1ccfe994151b3bd80fe2b3934a1bc89384599">
      <!--<attribute type="filename">116.exe</attribute-->
      <md5>8b31da6402d850ce94e7c19bc97effe1</md5>
      <sha1>850e5b037c799f86f04ee63da786f9ee139ebf57</sha1>
      <sha256>2f437c1c8f73c2d6ffbb6214d3f1ccfe994151b3bd80fe2b3934a1bc89384599</sha256>
      <size>32769</size>
      <crc32>34efdbca</crc32>
    </file>
    <file id="3a437c1c8f73c2d6ffbb6214d3f1ccfe994151b3bd80fe2b3934a1bc89384599">
      <!--<attribute type="filename">116.exe</attribute-->
      <md5>aa31da6402d850ce94e7c19bc97effe1</md5>
      <sha1>990e5b037c799f86f04ee63da786f9ee139ebf57</sha1>
      <sha256>22437c1c8f73c2d6ffbb6214d3f1ccfe994151b3bd80fe2b3934a1bc89384599</sha256>
    </file>
  </objects>
</malwareMetaData>

```

## Main concepts (2)

- Relationships1.. RelationshipsXX
  - Child
  - Parent
  - isClassifiedAs, droppedBy, hosts, installed, runs, exploits, downloads, resolvesTo, verifiedBy, usesCNC, contactedBy, operatedByEntity, isnameServerOf, causesToInstall, ...
  
- fieldData1..fieldDataYY
  - firstSeen
  - Origin (e.g. country/collection/honeypot/...)
  - Commonality, priority
  
- Reference - file[@id="12345"].



# Example (file+ref+classification)

```

<objects>
  <!-- one file -->
  <file id="2f437c1c8f73c2d6ffbb6214d3f1ccfe994151b3bd80fe2b3934a1bc89384599">
    <!--<attribute type="filename">116.exe</attribute-->
    <md5>8b31da6402d850ce94e7c19bc97effe1</md5>
    <sha1>850e5b037c799f86f04ee63da786f9ee139ebf57</sha1>
    <sha256>2f437c1c8f73c2d6ffbb6214d3f1ccfe994151b3bd80fe2b3934a1bc89384599</sha256>
    <size>32768</size>
  </file>

  <!-- one classification -->
  <classification id="AVG:Virut.BK" type="dirty">
    <classificationName>Virut.BK</classificationName>
    <companyName>AVG</companyName>
  </classification>
</objects>

<!-- this file is Virut -->
<relationships>
  <relationship type="isClassifiedAs">
    <parents>
      <ref>file[@id = '2f437c1c8f73c2d6ffbb6214d3f1ccfe994151b3bd80fe2b3934a1bc89384599']</ref>
    </parents>
    <children>
      <ref>classification[@id='AVG:Virut.BK']</ref>
    </children>
  </relationship>
</relationships>

```

# Example (field data)

```

<!-- this is the prevalence data -->
<fieldData>
  <!-- by file -->
  <fieldDataEntry>
    <references>
      <ref>file[@id = '2f437c1c8f73c2d6ffbb6214d3f1ccfe994151b3bd80fe2b3934a1bc89384599'] </ref>
    </references>
    <startDate>-1999-11-25T00:00:00</startDate>
    <endDate>2008-11-26T00:00:00 </endDate>
    <origin>user</origin>
    <commonality>8</commonality>
    <location type="countryCodeISO3166-2">US</location>
  </fieldDataEntry>
  <fieldDataEntry>
    <references>
      <ref>file[@id = '2f437c1c8f73c2d6ffbb6214d3f1ccfe994151b3bd80fe2b3934a1bc89384599'] </ref>
    </references>
    <startDate>2008-11-26T00:00:00</startDate>
    <endDate>2008-11-27T00:00:00</endDate>
    <origin>user</origin>
    <commonality>5</commonality>
    <location type="countryCodeISO3166-2">US</location>
  </fieldDataEntry>
  <fieldDataEntry>
    <references>
      <ref>file[@id = '2f437c1c8f73c2d6ffbb6214d3f1ccfe994151b3bd80fe2b3934a1bc89384599'] </ref>
    </references>
    <startDate>2008-11-27T00:00:00</startDate>
    <endDate>2008-11-28T00:00:00</endDate>
    <origin>user</origin>
    <commonality>1</commonality>
  </fieldDataEntry>

```

# Next Steps

- We're looking for active members of the Malware Working Group
- We need more participants in the pilot
- We need ideas on critical areas we should focus on
  - Blacklisting of malicious packers
  - Improving sample exchange mechanisms

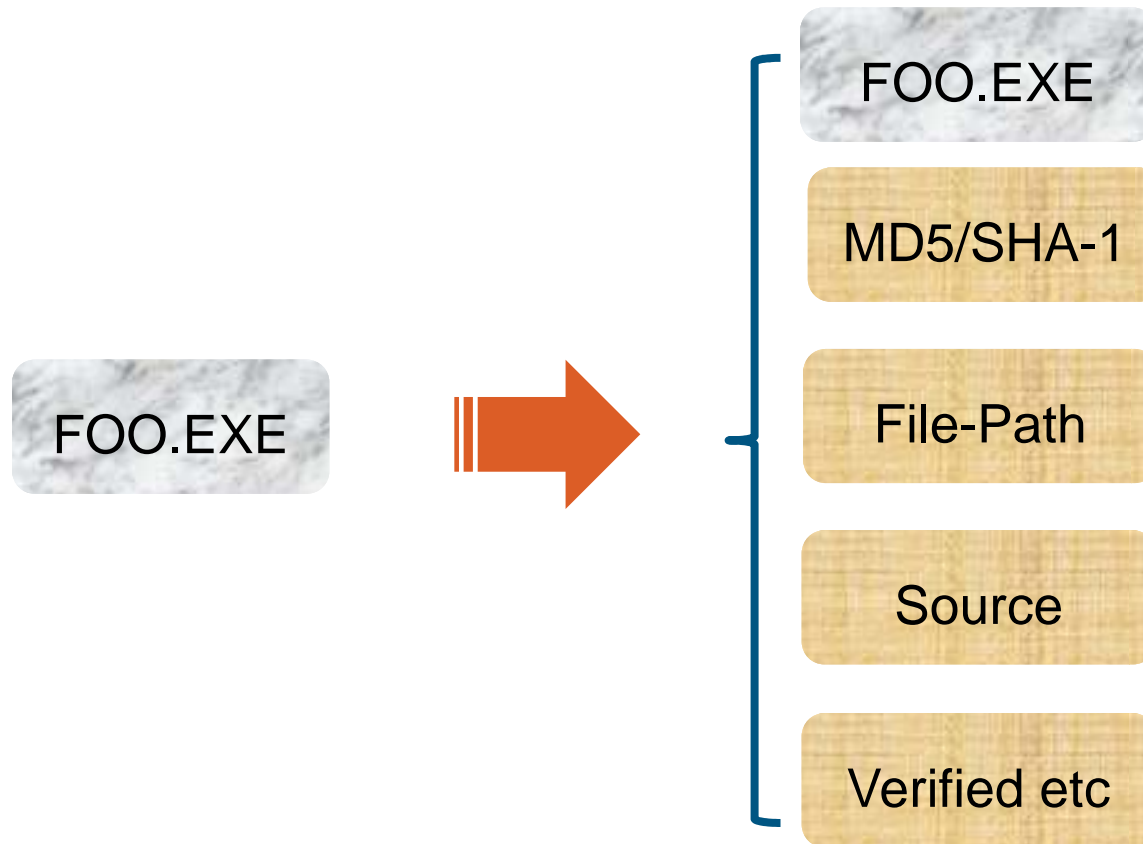


# Questions



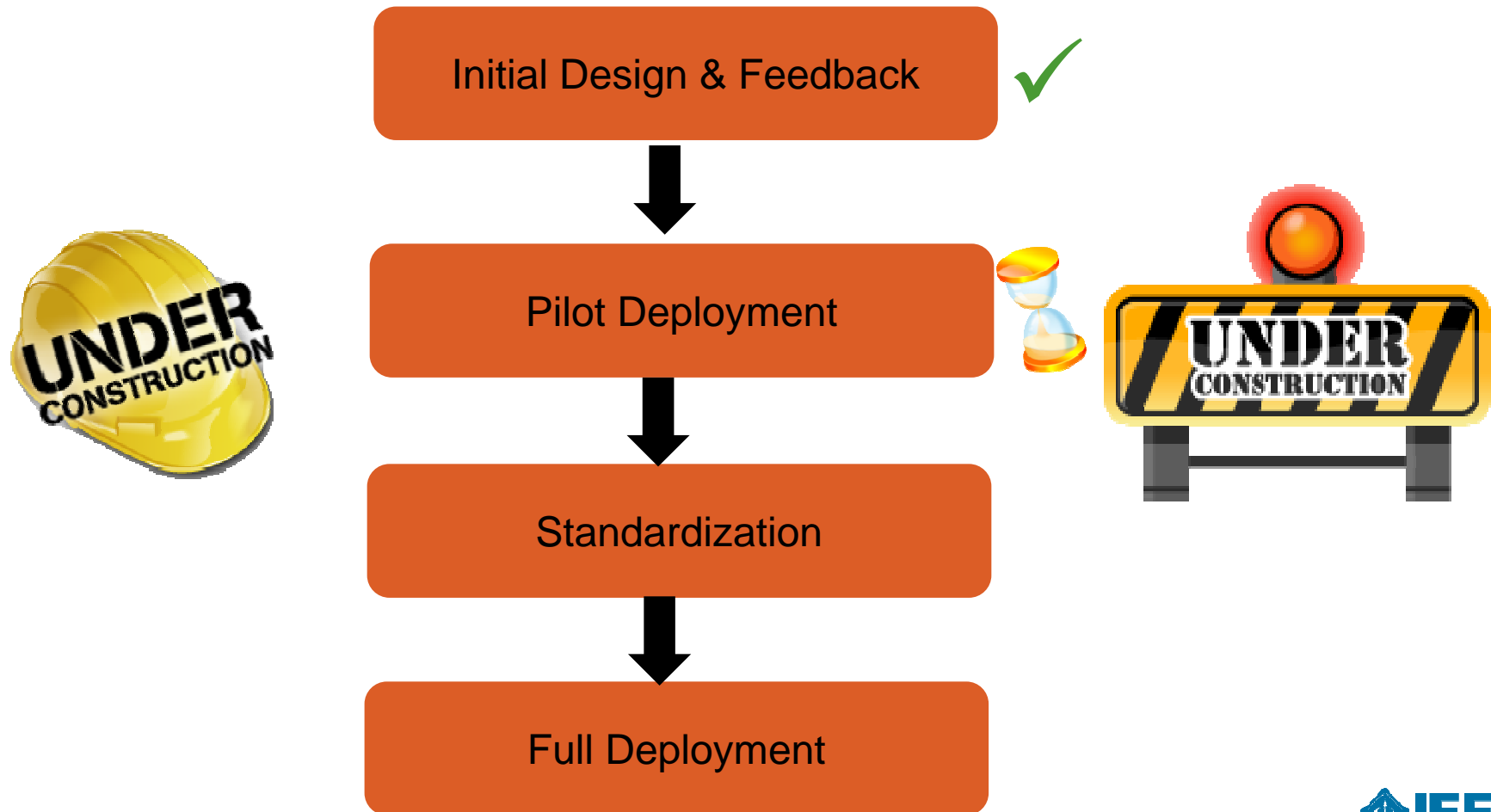
Emails: [joinicsg@ieee.org](mailto:joinicsg@ieee.org) , [mig@mcafee.com](mailto:mig@mcafee.com)

# How it works



We add available metadata to file sample during transfer (XML format)

## Development Phases





## Key Milestones

Deliverables	Date
Malware Meta-Data Exchange Format (MMDEF) V1 XML Schema document ready for Beta testing by initial WG Participants	9 <sup>th</sup> April 2009
Final XML Review Meeting by initial WG Participants	17 <sup>th</sup> April 2009
MMDEF V1 XML Schema document (draft) complete and ready for review with Invitees	22 <sup>nd</sup> April 2009
MMDEF V1 Review 1	1 <sup>st</sup> May 2009
MMDEF V1 Review 2	8 <sup>th</sup> May 2009
MMDEF V1 Review 3	15 <sup>th</sup> May 2009
MMDEF V1 Review 4	22 <sup>nd</sup> May 2009
MMDEF V1 XML Schema document (final) complete and sent for informal WG ballot of readiness for piloting	29 <sup>th</sup> May 2009
Approval of MMDEF V1 XML Schema document for piloting	17 <sup>th</sup> June 2009
Piloting of Schema begins	18 <sup>th</sup> June 2009
TargetPiloting concludes	31 <sup>st</sup> July 2009
MMDEF V1.1 Schema final edits and review complete (if needed)	August 2009
MMDEF V1.1 Schema balloted	Late August 2009

- The next section of the presentation gives a brief outline of the XML schema