Reconciling Objective Data with Analytical Uncertainty
Recap – Key Ideas

- **Blind Spots**
  - Ryan Stillions – Lack of visibility from info overload
  - Alex Pinto – “Is your sample data confirming your bias?”
  - Michael Wilburn – Pentest like a real blackhat to find the actual vulnerabilities

- **Context**
  - Ryan Stillions – Context determines the value of intel
  - Michael Wilburn – Not a vulnerability until there is context around it

- **Assumptions/Conclusions**
  - Brian Krebs – Question Assumptions
  - Rick Holland – Jumping to conclusions – don’t do it
Overview

- Traditional Intel
  - Blind Spots
- Cyber Intel
- Uncertainty
  - Intent
  - Managing Uncertainty
  - Estimative Language
  - Examples
- In Real Life
- Discussion
Dealing with Uncertainty

Colin Powell’s Rules for Intelligence:

- Tell me what you know
- Tell me what you don’t know
- Tell me what you think
- Always distinguish which is which
Data vs Assessment

Data
- Information
- Black & White
- Precise
- Intransitive

Assessment
- Estimate
- Grey
- Fuzzy
- Transitive
Traditional Intel – “INT” Review

- SIGINT
  - Signals Intelligence
- IMINT
  - Imagery Intelligence
- HUMINT
  - Human Intelligence
- OSINT
  - Open Source Intelligence
- RUMINT
  - Rumor Intelligence
Traditional Intel – Blind Spots

- Radio Intercepts
  - Enigma/Code Talkers
  - Information not transmitted/intercepted
- Aerial Imagery
  - Decoys and concealment
  - Access and area denial (Air Defense)
- Informants/Spies
  - Double Agent
  - Poor access
- Media/Propaganda
  - Misinformation
Serious Blind Spots
Be aware of “Blind Spot”

Multi-INT whenever possible
  - Multi-Source where not possible

Call out sources and level of uncertainty present

Bad analysis =
  - Bad defense
  - Bad policy
  - Bad press

Avoid Blind Spots:
  - Multi-Source
    - Host + Network
  - Collaborate
Cyber Intelligence

- Newest –INT
  - Tradecraft still being developed
  - Combines many traditional “INTs”
Cyber Intelligence

- Technical analysis can answer:
  - When
  - What
  - How
  - Where (sometimes)
- Uncertainty surrounds:
  - Who
  - Why
  - Where (sometimes)
  - When (sometimes – Recon)
Intent

- Ultimate Source of Uncertainty
- Impossible to know for certain
  - Flip Flops
  - Double Cross
  - Indecision
  - External force (blackmail)
- MUST never be stated as a certainty
  - Impossible to know what goes on inside a person’s head
Managing Uncertainty

- Compound sources of uncertainty
  - Measuring how much you don’t know
  - How much uncertainty have you eliminated
- Managed with clear communication
How solid is your data?
- Confirmed
- Probable
- Possible
- Unconfirmed

How sure are you?
- High Confidence
- Medium Confidence
- Low Confidence
Estimative Words

- **High Confidence**
  - Supported by preponderance of evidence
  - Significant evidence missing

- **Medium Confidence**
  - New evidence would invalidate
  - Little evidence available to support
  - Other equally likely hypotheses exist

- **Low Confidence**
  - No evidence against
  - All but certain

- More on this in the next presentation
Assess with high confidence that unconfirmed reports of UFOs will happen next month.

Assess with low confidence that Dori from “Finding Nemo” was possibly only pretending to have a bad memory to prolong her time with Nemo’s father Marlin.

There is a confirmed connection between the events in “The Hobbit” and the events in “The Lord of the Rings”. (High Confidence)

Possible Iran-based actors may have carried out attacks against Sands Casino in February 2014 (Medium Confidence)
When to Use Estimative Language

**Rarely Use – Cite your data**
- Log Data
- Network Traffic
- Forensic Data

**Always Use – Cite your uncertainty**
- Why network traffic means X
- What malware capabilities imply about adversary sophistication
- How log data translates into targeting list
In Real Life

- **Disguise Poor Analysis:**
  - Weasel Words
  - Non sequitur statements
- **Echo chamber effect:**
  - One poorly sourced report
  - Spawns multiple child/related reports
  - Total number of reports creates confident feeling
  - Hides flaws in the original source
  - Confidence is based on total reports + last comment on topic
- **Uncertainty should shrink when consuming any intelligence, not grow**

**Intelligence Rules**
- Tell me what you know
- Tell me what you don’t know
- Tell me what you think
- *Always* distinguish which is which

-Eric Hutchins
Estimative language provides reader with guidance on how seriously to take the assessment.

- There is uncertainty built into the underlined words, telling the reader that this assessment may be wrong.

“Other countries have spied on American companies, and they have stolen from them, but this is likely the first time—occurring months before the late November attack on Sony Pictures Entertainment—that a foreign player simply sought to destroy American corporate infrastructure on such a scale.”

-Bloomberg news
Example – Op Cleaver

- Page 18 – IRGC logo, with no tie in to the text presented
- Page 66 – Listed in conclusion as a source of adversary activity coming out of Iran
  - But not directly tied to the activity reported
- Page 72 – states that the IRGC logo inspired the Cleaver logo because of the close connection between the activity and the IRGC
  - **NO data provided to support the connection**
    - **NO estimate language to guide reader confidence**
Conclusion

**Do**
- Clearly mark unknowns in the data
- Clearly communicate level of certainty

**Don’t**
- Make analytical leaps
- Make unstated assumptions

**Payoff**
- Create higher quality cyber intel
- Improve network defenses