Technische Hochschule
Brandenburg
University of
Applied Sciences
Institute for
Security and Safety

Attribution of Cyber Incidents – A Framework for an International Cyber Attribution Council

Institute for Security and Safety (ISS)
at the Brandenburg University
of Applied Sciences

Potsdam/Germany
June 2019

Attribution of Cyber Incidents – A Framework for an International Cyber Attribution Council

Institute for Security and Safety (ISS), Potsdam/Germany

http://www.uniss.org/

info@uniss.org

About This Report

The Institute for Security and Safety (ISS) is an associated institute of the Brandenburg University of Applied Sciences and a NGO in consultative status with the Economic and Social Council of the United Nations. The ISS focuses on independent, nonpartisan research, as well as international education and training activities in the areas of cybersecurity and nuclear security. This report was sponsored by the German Federal Foreign Office and written by Marco Macori, ISS research fellow, with input provided by Guido Gluschke, ISS co-director. ISS' publications do not necessarily reflect the opinions of its sponsors.

The report evaluates existing information and academic literature regarding the attribution of cyber incidents. More precisely, it focuses on an appropriate framework for the creation of an international body tasked to perform independent, objective and transparent attribution of significant/major cyber-attacks on a global scale. Importantly, the body's task does not include incident response, (law) enforcement activities or network defense actions in any regard. As a matter of course, more work will be needed to further elaborate on the proposed framework.

Key Words

Attribution, Cyber-attacks, Cyber Incidents, Cybersecurity, Cyberstabilty, Intelligence, International Cyber Attribution Council (ICAC), Transparency

CONTENTS

Introduction

Not least since January 2017, when the U.S. Intelligence Community (IC) publicly attributed cyber-enabled interference activities (or "active measures") in the 2016 U.S. presidential election to Russia[1], the complex issue of attribution of major cyber incidents has become a "hot topic" in international security, diplomacy and politics. In its Intelligence Community Assessment the U.S. IC went out of its way to declassify as much highly classified support material as possible in order to lend public credibility to its conclusions. However, for obvious reasons it is impossible for national intelligence agencies to fully reveal their sensitive sources and methods. As a matter of fact, the U.S. IC states as much in their 2017 assessment.[2] Consequently, those critics who argue that the IC did not really provide full proof for their conclusions are – strictly factually speaking – not entirely wrong. All of this just serves to illustrate the need for the creation of an international body tasked to perform independent and transparent attribution of major cyber incidents, one that does not have to rely on intelligence from nation states.

In this report, we'll describe – based on already existing information and academic literature – an appropriate framework for such a body, for which we recommend the name "International Cyber Attribution Council (ICAC)". The report does not focus though on technical, historical and contextual aspects regarding the attribution of cyber-attacks. We recommend for an overview, if needed, on these areas the following excellent work:

- Davis II, John S., Benjamin Boudreaux, Jonathan William Welburn, Jair Aguirre, Cordaye Ogletree, Geoffrey McGovern, Michael S. Chase, "Stateless Attribution: Toward International Accountability in Cyberspace", RAND Corporation, 2017, https://www.rand.org/pubs/research_reports/RR2081.html

- Herpig, Sven, and Thomas Reinhold, "Spotting the bear: credible attribution and Russian operations in cyberspace", in: Chaillot Paper of the European Union Institute for Security Studies, October 2018, p. 33-42, https://www.iss.europa.eu/sites/default/files/EUISSFiles/CP_148.pdf

[1] U.S. Office of the Director of National Intelligence (DNI 2017), "Assessing Russian Activities and Intentions in Recent US Elections", Intelligence Community Assessment, 6 January 2017, https://www.dni.gov/files/documents/ICA_2017_01.pdf

[2] Ibid., p. 1: "The Intelligence Community rarely can publicly reveal the full extent of its knowledge or the precise bases for its assessments, as the release of such information would reveal sensitive sources or methods and imperil the ability to collect critical foreign intelligence in the future. Thus, while the conclusions in the report are all reflected in the classified assessment, the declassified report does not and cannot include the full supporting information, including specific intelligence and sources and methods."

- Rid, Thomas, and Ben Buchanan, "Attributing Cyber Attacks," in: Journal of Strategic Studies, Vol. 38, Nos. 1–2, 2015, https://ridt.co/d/rid-buchanan-attributing-cyber-attacks.pdf

- Romanosky, Sasha, and Benjamin Boudreaux, "Private Sector Attribution of Cyber Incidents: Benefits and Risks to the U.S. Government", RAND Corporation, 2019, https://www.rand.org/pubs/working_papers/WR1267.html

In a nutshell, attribution is important for two main reasons: "First, attribution imposes responsibility on the party or parties involved in the cyberattack. Second, attribution deters future cyberattacks by raising the cost of state-sponsored offensive activity."[3] Furthermore, international cybersecurity norms remain ineffective without reliable and credible mechanisms for attribution, which – taken together – hinders the creation of stability in cyberspace. However, "analysis of recent cases indicates that the practice of attribution has been diffuse and discordant, with no standard methodology used in the investigations to assess evidence, nor a universal confidence metric for reaching a finding … Further, public statements of attribution have been met with suspicion, confusion, and a request for greater transparency about the investigation and the evidential basis."[4]

[3] Beyer, Jessica, et al., "Cyberattack Attribution: A blueprint for private sector leadership", The Henry M. Jackson School of International Studies, University of Washington, 2017, p. i, https://jsis.washington.edu/wordpress/wp-content/uploads/2017/07/ARP-2017-Report-FINAL.pdf

[4] Davis II, John S., Benjamin Boudreaux, Jonathan William Welburn, Jair Aguirre, Cordaye Ogletree, Geoffrey McGovern, Michael S. Chase, "Stateless Attribution: Toward International Accountability in Cyberspace", RAND Corporation, 2017, https://www.rand.org/pubs/research_reports/RR2081.html

Framework for an International Cyber Attribution Council (ICAC)

Council Blueprint

The International Cyber Attribution Council (henceforth "the Council" or infrequently "the Consortium") should not be dependent on international agreements between nation states. Rather, its legitimacy and authority needs to be solely based on its independent, effective, objective and transparent work. This stands in contrast to, for instance, the International Atomic Energy Agency (IAEA), the Organisation for the Prohibition of Chemical Weapons (OPCW) or the International Telecommunication Union (ITU). The Councils mission can be roughly described in the following way: "The mission of the organization is for a broad team of international experts to conduct independent investigations of major cyber incidents for the purpose of attribution … The international community could use the Consortium's findings to bolster network defenses, thwart future attacks, and pursue follow-on enforcement actions to hold the perpetrator(s) accountable. In addition to providing a credible and transparent judgment of attribution, the Consortium's investigations would help standardize diffuse methodological approaches, naming conventions, and confidence metrics that would advance shared understanding in cyberspace and promote global cybersecurity."[5]

The Council's membership should be set-up as follows[6]:

- A total membership of 31 persons; members need to come from a wide array of countries and regions in order to ensure a fair diverse geographic representation.
- Technical experts from leading private cybersecurity consultancies[7], technology companies, IT/cybersecurity laboratories and academia.
- Cyber policy, diplomacy and legal experts from research/policy institutes, think tanks, academia and civil society with proven records of excellence in their respective fields.

[5] Ibid., p. 27.
[6] See ibid., Beyer et al. (2017), and Charney, Scott, Erin English, Aaron Kleiner, Nemanja Malisevic, Angela McKay, Jan Neutze, and Paul Nicholas, "From Articulation to Implementation: Enabling Progress on Cybersecurity Norms", Microsoft Corporation, June 2016, p. 11f, https://query.prod.cms.rt.microsoft.com/cms/api/am/binary/REVmc
[7] Regarding some potential downsides of the involvement of private cybersecurity consultancies see: Rich, William G., "The US Leans On Private Firms To Expose Foreign Hackers", in: WIRED Opinion, 2018-11-29, https://www.wired.com/story/private-firms-do-government-dirty-work/

- Technical/forensic investigations and assessments should be performed by dedicated small teams of experts; the rest of the Council's members provide non-technical, relevant input, review and general oversight.

The Council can request information and other support from nation states regarding specific investigations if needed, but states shall not be in a position to volunteer information/ intelligence to the Council[8]. This approach stands in contrast to earlier proposals for a global attribution organization, such as those from the Microsoft Corporation and the Atlantic Council, a U.S. think tank in the field of international affairs.[9] Attribution decisions will be communicated publicly through established, standardized processes and channels, and will include transparent confidence levels.

Initially, the Council's work will be funded by two groups: first, IT and telecommunications companies, which possess vested interest in transparent, neutral and credible attribution of major cyber incidents; and second, philanthropic and civil society organizations. After its establishment, the Council shall evaluate if and how contributions from, most likely, the United Nations should be sought.[10] The Council may require a budget of $40 million for the first year, and a budget of $30 million per year for the following years.[11]

Description of the Attribution Process[12]

I) Which cyber incidents will be investigated by the Council?

The Council will have the prerogative to decide which cyber incidents it will investigate. However, that does not mean that the Council will select its "cases" for review. Rather, victims of cyber-attacks shall have the opportunity to bring their case to the Council, which then decides, based on regularized criteria and thresholds, whether that case merits "attribution action" by the Council. This process guarantees a victim's privacy on

[8] See Davis II et al. (2017) for reasons for the exclusion of nation states, e.g. on p. 26: "… although states offer unique intelligence capabilities for attribution, their involvement can create complications that have a negative impact on the objectivity, transparency, and independence of the finding." Further reasons can be found on p. 29f.
[9] See Charney et al. (2016) and Healey, Jason, John C. Mallery, Klara Tothova Jordan, and Nathaniel V. Youd, "Confidence-Building Measures in Cyberspace: A Multistakeholder Approach for Stability and Security", Atlantic Council, November 2014, http://www.atlanticcouncil.org/images/publications/Confidence-Building_Measures_in_Cyberspace.pdf
[10] See Davis II et al. (2017): p. 44f.
[11] Estimate by Beyer et al. (2017): p. iv.
[12] The description of the attribution process is significantly based on the work of Davis II et al. (2017): p. 35ff.

the one hand, and on the other hand, gives the Council independent authority in its actions and decisions.

II) How will evidence be collected by the Council?

One can assume that an attack victim (or a group of victims, of course) that comes with an attribution request to the Council will be willing to provide the necessary information, network data, artifacts and other relevant assets. Nonetheless, if critical intellectual property or other significantly sensitive information of the victim were involved in the process, the Council shall be able to conclude appropriate and narrowly focused confidentiality agreements with the victim.

III) How will evidence be assessed and evaluated by the Council?

The Council will, as clearly and transparently as possible, define the assessment methods and the analytic framework that will be employed concerning evidence evaluation. The analytic framework, for its part, should include mechanisms to compare cyber incidents regarding scope, scale, sophistication and severity. After the assessment and evaluation of evidence in a case is finished, the members of the Council shall aim to reach a consensus attribution decision. In case a consensus cannot be reached among the members, the simple majority will rule (with the possibility of abstention from voting). At the same time, the opinion of the minority will be recorded and communicated publicly together with the overall "ruling" by the Council. That same procedure has been employed, for instance, by the U.S. Supreme Court. It is an expression of democratic transparency, which should contribute to public trust in and long-term credibility of an organization.

IV) How will attribution decisions be communicated by the Council?

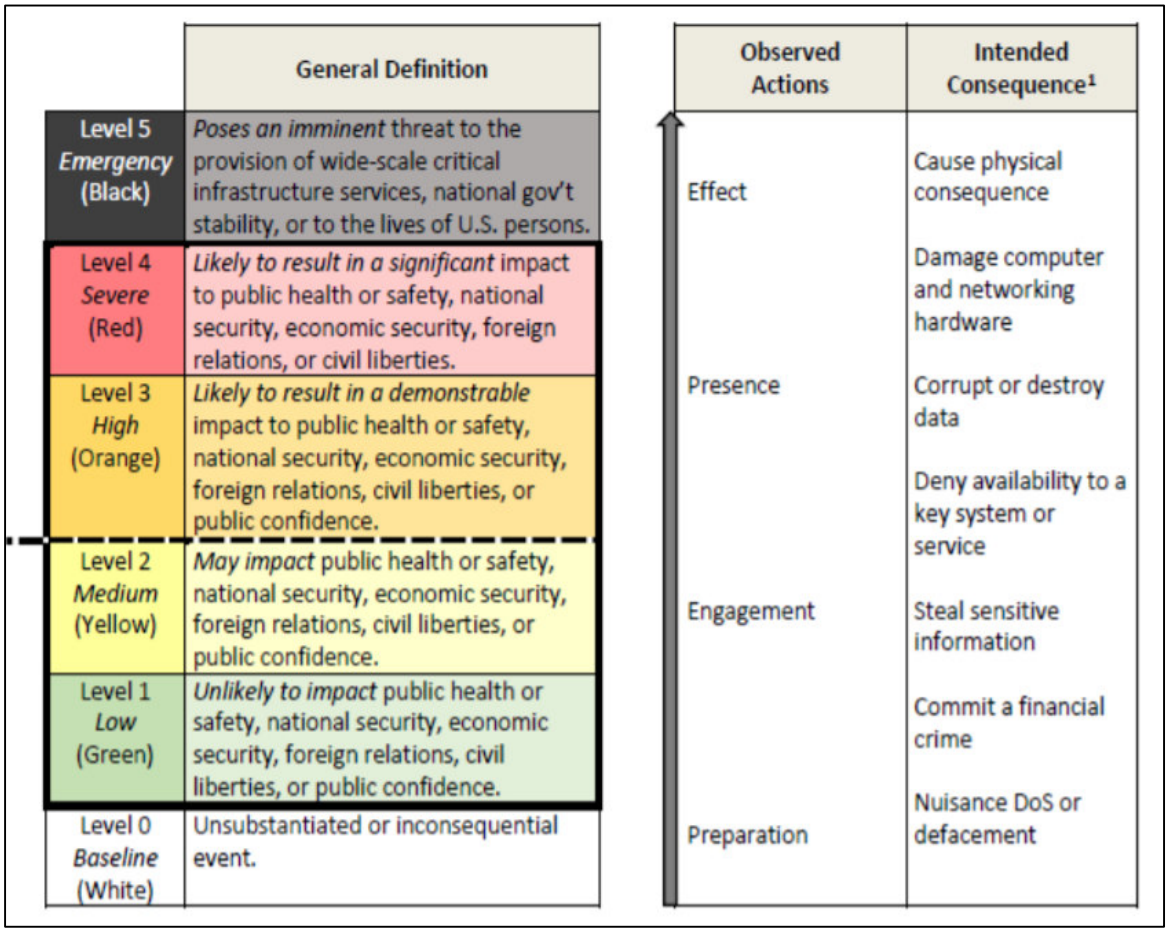
The Council shall communicate all its attribution decisions to the public in a timely fashion. This will include verbal statements by a professional spokesperson, as well as written reports, not least because only a minority of the general public will read such a report. Moreover, in today's modern media landscape verbal statements have a far wider reach since they can be used in TV/online news pieces, etc. Having established that, the written reports will include as many technical details as possible for consumption by an

expert audience. More generally, written reports will describe the investigative process and the reasons for the decision.

In order to ensure the reputational authority, credibility and integrity of the Council, the attribution decisions have to be communicated in a transparent way. Consequently, it will be essential to express how confident the Council is in its evaluations, judgements and decisions. In this context, the Council will basically use the analytic standards developed some years ago by the U.S. Office of the Director of National Intelligence (DNI)[13], which leads to the following confidence levels to be employed: low, medium-low, medium-high and high. Brief explanations/illustrations of these confidence levels at the beginning of attribution reports by the Council will help the understanding of non-expert readers.

[13] See DNI (2017) and U.S. Office of the Director of National Intelligence (DNI 2015), "Analytic Standards," Washington, D.C., Intelligence Community Directive 203, 2015-01-02, https://www.dni.gov/files/documents/ICD/ICD%20203%20Analytic%20Standards.pdf

Furthermore, attribution communications by the Council will include an assessment of the severity of the analyzed cyber incident. Here, the color-coded severity schema used by the U.S. Department of Homeland Security (DHS)[14] will be employed:



	General Definition
Level 5 *Emergency* (Black)	*Poses an imminent* threat to the provision of wide-scale critical infrastructure services, national gov't stability, or to the lives of U.S. persons.
Level 4 *Severe* (Red)	*Likely to result in a significant* impact to public health or safety, national security, economic security, foreign relations, or civil liberties.
Level 3 *High* (Orange)	*Likely to result in a demonstrable* impact to public health or safety, national security, economic security, foreign relations, civil liberties, or public confidence.
Level 2 *Medium* (Yellow)	*May impact* public health or safety, national security, economic security, foreign relations, civil liberties, or public confidence.
Level 1 *Low* (Green)	*Unlikely to impact* public health or safety, national security, economic security, foreign relations, civil liberties, or public confidence.
Level 0 *Baseline* (White)	Unsubstantiated or inconsequential event.

Observed Actions	Intended Consequence[1]
Effect	Cause physical consequence
	Damage computer and networking hardware
Presence	Corrupt or destroy data
	Deny availability to a key system or service
Engagement	Steal sensitive information
	Commit a financial crime
Preparation	Nuisance DoS or defacement

One major rationale for including a severity schema is to encourage post-attribution decision and actions by third parties. E.g., by corporations/institutions/governments etc. that may only employ additional cybersecurity measures in response to severe cyber-attacks.

[14] U.S. Department of Homeland Security, "National Cyber Incident Response Plan", Washington, D.C., December 2016, p. 38, https://www.us-cert.gov/sites/default/files/ncirp/National_Cyber_Incident_Response_Plan.pdf

Lastly, at a later stage the Council may also include an assessment of sophistication of a cyber-attack. Relevant research has already been performed in this area, for instance by B. Buchanan of the Belfer Center at Harvard[15], but more of it will be required to include a standardized framework or schema in attribution communications by the Council. Alternatively, the Council itself could endeavor to develop it.

V) Will the Council recommend enforcement action or punishment?

Most importantly, the Council shall not be in a position to refer "cases" to institutions such as the United Nations Security Council or the International Court of Justice in The Hague. Additionally, it will not recommend concrete, punitive follow-on actions such as criminal sentencing. Rather, it will be a decision to be made by the victim in which way to make potential further use of the attribution decision by the Council. Appropriate processes and norms should develop over time in this important context. Here, it is encouraged that a core group of institutions/states etc. should take the lead instead of waiting for all relevant stakeholders to agree on appropriate processes and norms.

[15] Buchanan, Ben, "The Legend of Sophistication In Cyber Operations", Belfer Center for Science and International Affairs, Harvard Kennedy School, January 2017, https://www.belfercenter.org/sites/default/files/files/publication/Legend%20Sophistication%20-%20web.pdf

Organizational Chart and Best Practices

In 2017, a research team of the Henry M. Jackson School of International Studies at the University of Washington worked on a project sponsored by Microsoft detailing current best practices for cyber-attack attribution and response.[16] After creating a referenceable data set by studying 23 existing "attribution organizations", such as the IAEA or the United Nations Al-Qaida Sanctions Committee, and investigative processes, such as the "Sony Pictures Hack Investigation" and the "Stuxnet Investigation", the team drew upon private sector expertise from a wide array of countries to create an organizational blueprint for a cyber attribution organization. Although not totally congruent with all aspects laid out in this report, the Council should strongly consider for its operation over time the proposed best practices and organizational structures recommended by the Henry M. Jackson School team.

Chief among the proposed best practices are the following:

- Equitable geographic representation
- Organizational transparency
- Stakeholder outreach
- Internal accountability
- Inclusion of technical and geopolitical experts
- Private sector membership[17]

[16] Beyer et al. (2017).
[17] Ibid., p. iii.

Best Practices in Detail[18]:

The figure below outlines the direction of the information flow in an attribution process:

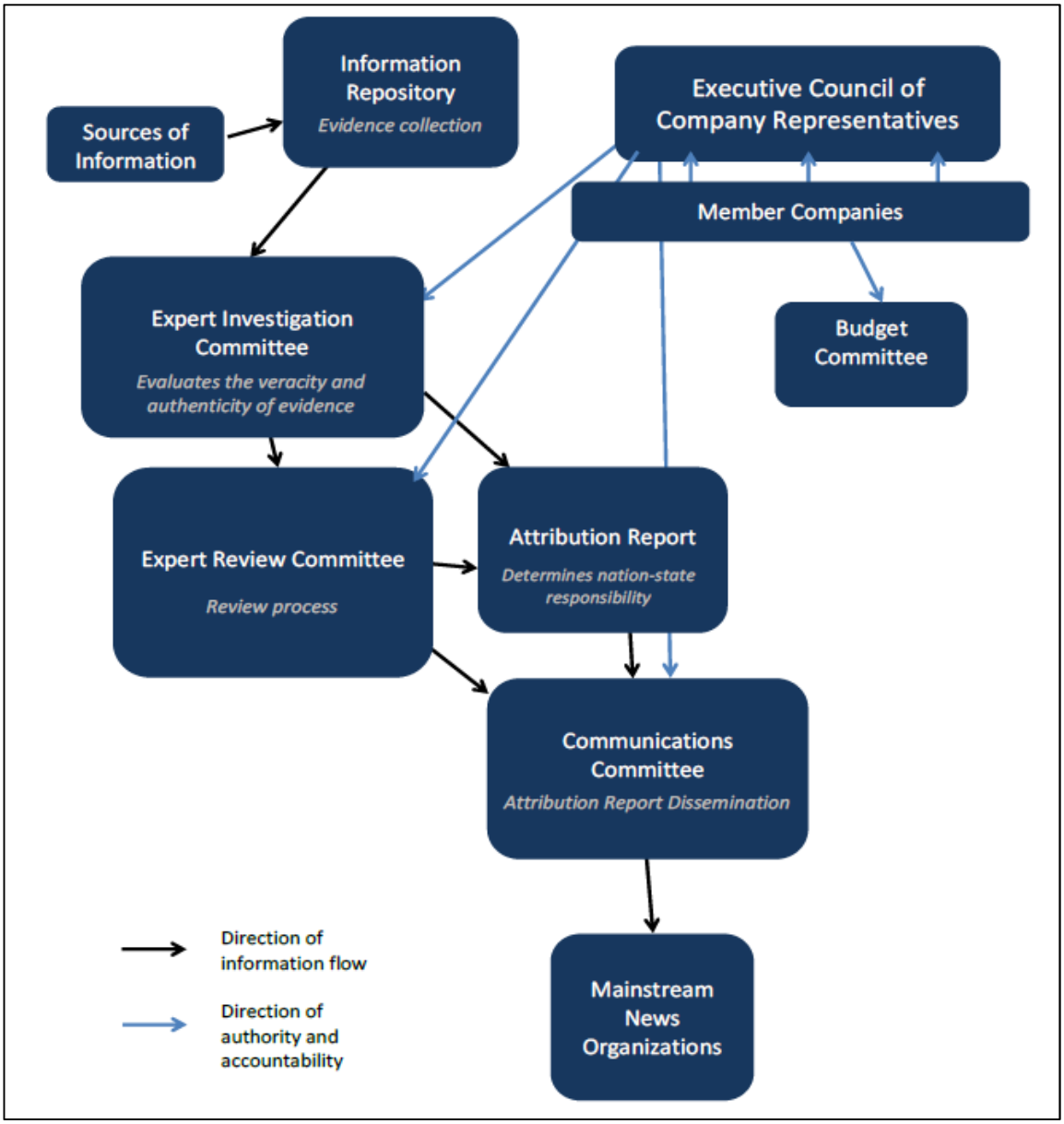
- Information arrives at the Council through an information repository.
- As evidence is collected, an Expert Investigation Committee verifies the veracity and authenticity of the evidence.
- An Expert Review Committee evaluates the evidence and the findings of both groups to create the substance of the attribution report.
- The Expert Review Committee forwards the attribution report to the Communication Committee.
- The Communication Committee works with the media to publicize the results of the review.[19]

[18] Ibid., p. 8.
[19] Ibid., p. 5f.

Moreover, the figure illustrates the Council's authority and accountability hierarchy:

- Member companies make up an Executive Council of Company Representatives and a Budget Committee.
- The Executive Council provides resources and oversight to the two experts groups. It also assists with the dissemination of the Council's findings.
- Term limits are incorporated into the Executive Council's design as a good governance mechanism to ensure diversity within the executive leadership.[20]

Organizational Chart in Detail[21]:



[20] Ibid., p. 6.
[21] Ibid., p. v.

ABBREVIATIONS

DHS	U.S. Department of Homeland Security
DNI	U.S. Director of National Intelligence
IAEA	International Atomic Energy Agency
IC	Intelligence Community
ICAC	International Cyber Attribution Council
ITU	International Telecommunication Union
OPCW	Organisation for the Prohibition of Chemical Weapons

REFERENCES

Beyer, Jessica, et al., "Cyberattack Attribution: A blueprint for private sector leadership", The Henry M. Jackson School of International Studies, University of Washington, 2017, https://jsis.washington.edu/wordpress/wp-content/uploads/2017/07/ARP-2017-Report-FINAL.pdf

Buchanan, Ben, "The Legend of Sophistication In Cyber Operations", Belfer Center for Science and International Affairs, Harvard Kennedy School, January 2017, https://www.belfercenter.org/sites/default/files/files/publication/Legend%20Sophistication%20-%20web.pdf

Charney, Scott, Erin English, Aaron Kleiner, Nemanja Malisevic, Angela McKay, Jan Neutze, and Paul Nicholas, "From Articulation to Implementation: Enabling Progress on Cybersecurity Norms", Microsoft Corporation, June 2016, https://query.prod.cms.rt.microsoft.com/cms/api/am/binary/REVmc8

Davis II, John S., Benjamin Boudreaux, Jonathan William Welburn, Jair Aguirre, Cordaye Ogletree, Geoffrey McGovern, Michael S. Chase, "Stateless Attribution: Toward International Accountability in Cyberspace", RAND Corporation, 2017, https://www.rand.org/pubs/research_reports/RR2081.html

Healey, Jason, John C. Mallery, Klara Tothova Jordan, and Nathaniel V. Youd, "Confidence-Building Measures in Cyberspace: A Multistakeholder Approach for Stability and Security", Atlantic Council, November 2014, http://www.atlanticcouncil.org/images/publications/Confidence-Building_Measures_in_Cyberspace.pdf

Herpig, Sven, and Thomas Reinhold, "Spotting the bear: credible attribution and Russian operations in cyberspace", in: Chaillot Paper of the European Union Institute for Security Studies, October 2018, p. 33-42, https://www.iss.europa.eu/sites/default/files/EUISSFiles/CP_148.pdf

Rich, William G., "The US Leans On Private Firms To Expose Foreign Hackers", in: WIRED Opinion, 2018-11-29, https://www.wired.com/story/private-firms-do-government-dirty-work/

Rid, Thomas, and Ben Buchanan, "Attributing Cyber Attacks," in: Journal of Strategic Studies, Vol. 38, Nos. 1–2, 2015, https://ridt.co/d/rid-buchanan-attributing-cyber-attacks.pdf

Romanosky, Sasha, and Benjamin Boudreaux, "Private Sector Attribution of Cyber Incidents: Benefits and Risks to the U.S. Government", RAND Corporation, 2019, https://www.rand.org/pubs/working_papers/WR1267.html

Smith, Brad, "The Need for a Digital Geneva Convention", in: The Official Microsoft Blog, 2017-02-14, https://blogs.microsoft.com/on-the-issues/2017/02/14/need-digital-geneva-convention/#sm.0000lgk5vy1e5he6gzh99xggdxmu5

U.S. Department of Homeland Security, "National Cyber Incident Response Plan", Washington, D.C., December 2016, https://www.us-cert.gov/sites/default/files/ncirp/National_Cyber_Incident_Response_Plan.pdf

U.S. Office of the Director of National Intelligence (DNI 2015), "Analytic Standards," Washington, D.C., Intelligence Community Directive 203, 2015-01-02, https://www.dni.gov/files/documents/ICD/ICD%20203%20Analytic%20Standards.pdf

U.S. Office of the Director of National Intelligence (DNI 2017), "Assessing Russian Activities and Intentions in Recent US Elections", Intelligence Community Assessment, 6 January 2017, https://www.dni.gov/files/documents/ICA_2017_01.pdf