14. Too Late - Civil Defense in Cyberwar 2.0

The only defense against a Hacker-AI is to be technically prepared so that computer devices would not fall into the hand of the adversary. With technical preparation, only technically unprepared devices could be used in Cyberwar 2.0. In that situation, there will be a threshold in which the main objective, i.e., the covert coordinated recruitment of people from targeted countries' populations for being part of a government overthrow, is not feasible anymore.

Where the threshold is and if technical preparation could deter an assailant from taking Cyberwar 2.0 action is unknown. With misdirection and offensive Cyberwar 2.0 steps still being prepared, it would probably depend on the outcome of simulation or exercises to know if government overthrow/regime change is a likely outcome. It will also depend on assailant's determination if they want to start this conflict anyway.

The war's outcome will depend if cyberwar weapons are used by fully and comprehensively prepared nations or by a group of criminals that enable wealthy and disgruntled groups within the attacked country. It is conceivable that Cyberwar 2/0 weapon usage is only part of a cyber-uprising or an internal civil war. However, it is better to assume that a potential outcome or even its goal could be an AI-based surveillance state.

In approaching non-technical preparedness done in defense against a foreign nation or a rogue group waging cyberwar, readers are briefly reminded about the capabilities that defenders could expect from a party using Hacker-AI in waging Cyberwar 2.0-related actions:

(1) Surveillance (audio, video, and usage) of smartphones or other IT devices - at scale (every system)

(2) Selective access denial for some (key-) people/positions caused by their device's malware

(3) Direct threatening of targeted people via AI bots on devices

(4) Realtime Deep-Fakes and their use in re-defining facts/truth

(5) Reduction of costly war consequences via pre-war intelligence and preparation

(6) Misdirections on who is the culprit in cyber activities

An effective defense against the above capabilities is only achievable by technically restricting the capabilities of malware/Hacker-AI and having countermeasures deployed comprehensively on all network-capable devices. However, it will take time to deploy sufficient defense capabilities, during which an assailant could still start waging Cyberwar 2.0.

Still, I believe governments and their people could and should prepare with non-technical means for Hacker-AI and Cyberwar 2.0 before sufficient technical capabilities are deployed to protect a country. Still, it isn't easy to forecast whether the defenders could prevent a cyberwar victory. Only extensive simulation with varying decisions or actions from recruitment and responses by the attacked government on additional secret preparation could give the assailant probabilities, i.e., how often his war script results in a successful outcome. Additionally, these estimates will depend on how reliable Cyber Reconnaissance delivers truthful and comprehensive data.

The defender could run similar simulations to determine or improve his survival odds.

Advanced digital forensic capabilities and human spies who provide reliable, potentially undeniable information on attackers' war plans, in combination with defenders' technical preparation, could potentially deter nations from starting a Cyberwar 2.0. Because cyberwar is deniable, the assailant could start and abandon it if it fails.

I introduced six **Threat-Levels (TL)** depending on defender's knowledge based on the assumed or proven existence, capabilities, or actions of Hacker-AI regarding the development/deployment of technical countermeasures. These levels could also be used in the preparation of non-technical countermeasures to Cyberwar 2.0:

- TL-0 means that there is no current threat from Hacker-AI.
- In TL-1, Hacker-AI is potentially developed but not detected yet.
- TL-2 means that all involved in developing/deploying of countermeasures were warned that Hacker-AI has tried to attack the security of development or preparation efforts covertly.
- Sub-level TL-2X means people in development or preparation were directly attacked.
- TL-3 would show that an adversary has used Hacker-AI in a presumingly successful Cyberwar 2.0 campaign.
- TL-4 means that adversary's Hacker-AI is making the development of countermeasures impossible.

I mentioned that it is unknown if Hacker-AI already exists. Therefore, the current threat level is below TL-2. This could change if experts determine that the amount of used zero-day vulnerabilities would increase steeply, leading to the conclusion that hacking was automated. If this assessment would lead to the conclusion that Hacker-AI features are also turned into deployable Cyberwar 2.0 weapons is then a political decision.

Situation / Scenario

I am not predicting future events or political developments, or intentions. I want to describe a scenario that requires urgent attention due to its attractiveness to adversaries and the catastrophic outcome for its targets.

In a previous chapter, I discussed three main scenarios; the US facilitates Cyberwar 2.0 capabilities for Russian dissidents to help them transition Russia into a post-Putin world more peacefully, China (PRC) attacking Taiwan (ROC), and rogue / private actors facilitate Cyberwar 2.0 as a Service for regime change.

Authoritarian regimes like Russia or PRC could potentially use Hacker-AI and Cyberwar 2.0 capabilities against their population at scale to prevent their citizens from participating in cyberwar-triggered uprisings. These countermeasures are not acceptable to liberal democracies.

The US and its allies could have already developed some Hacker-AI technology. Because of legal considerations, they are likely using it sparely in targeted ways and not deploying it offensively. However, it is assumed that intelligence services use malware-based Cyber Reconnaissance against a small group of people to gain intelligence to determine the intention of certain criminals / terrorists and locate / arrest them before their planning could turn irreversible into actions. However, skilled / smart terrorists or criminals are not using smartphones.

For our scenario in this chapter, I assume that nations like PRC prepare and then wage a Cyberwar 2.0 on Taiwan. PRC has or will invest in espionage, propaganda-based, damage-creating cyberwar weapons, and offensive malware-generating Hacker-AI and Cyberwar 2.0 capabilities.

Because Taiwan (ROC) has a high density of smartphones / IT devices, it is almost an ideal candidate for this Cyberwar 2.0. Therefore, the annexation of Taiwan by China is the most likely example of an offensive Cyberwar 2.0 and not damage-creating cyberweapons.

This baseline scenario assumes that PRC has the technical skills to create cyberattack tools and the political will to replace (decapitate) ROC's leadership, incl. governmental bureaucracy. In the aftermath, China could create / operate an AI-based surveillance apparatus that cover all 23 million inhabitants of Taiwan to fortify its gains.

Based on this baseline scenario, the targeted country (ROC) and seemingly unaffected countries (USA, NATO countries) are not technically prepared. No country has sufficient software-based defense tools or hardware-based cybersecurity. In the worst case, they may not even have started to develop the required defense tools.

I expect that we have two distinct situations. (A) The target country is directly surveilled, attacked, and defeated, and (B) directly unaffected countries (bystanders), basically the rest of the world, must respond/adapt to the fact that Cyberwar 2.0 is now an existential threat to every country and their government.

Only the USA, NATO countries, and a few other countries have the engineering that could contribute to the development/production of countermeasures. The cyberwar target (Taiwan) is disabled to participate in that development effort.

(A) Preparation: Goals/Measures for Cyberwar 2.0 Target

Overview

The scenario assumes that the assailants' main goal is decapitating the government and society. From that assumption, defenders can establish measures to counter Cyberwar 2.0 activities by keeping the government in place and operable as long as possible. Additionally, it is assumed that revealing that Hacker-AI was used in a government overthrow/regime change could have incalculable political and potential economical costs. Therefore, the assailant's costs of the conflict are significantly increased when the defender can prove to the world that Hacker-AI was used in this Cyberwar 2.0. Unfortunately, both goals (i.e., keep in power as long as possible and tell the truth about Hacker-AI) are very difficult to achieve under Cyberwar 2.0 conditions; it is likely impossible without dedicated and extensive preparation.

Preparing for Cyberwar 2.0 using digital or Internet-based means to report suspicious deep-fakes or intimidation by AI bots is a serious mistake. Digital channels will be suppressed or manipulated by flooding false or manipulated reports, even if the targeted country is only partly technically prepared. Assuming that defenders could somehow work around attacker's capabilities is most likely wrong. Instead, here are attackers' likely capabilities: (1) covert surveillance via mobile phones, (2) denial of service for critical people/organizations,

(3) intimidations via AI bots or (4) deep-fakes, (5) comprehensive attacker preparation/planning via simulation, and (6) misdirection on who is the attacker.

Governments must proactively prepare rules to remain in charge during active cyberwar actions; I called that cyberwar phase: CWP-II. These (emergency) rules would make its illegal or abrupt replacement by a puppet government as difficult and time-consuming as possible. Unfortunately, the legitimate government's probability of surviving Cyberwar 2.0 is slim. Therefore, it should be considered a victory if the old government could publicly announce CUCA ("Country is under Cyber-Attack") and warn the world community about what happened with irrefutable evidence.

Still, I propose that governments and citizens must prepare measures to achieve many goals within the following categories:

(a) Facilitating Information/Intelligence Gathering
  - Governments need reliable info on threats or demanded tasks from intimidated people asap

(b) Preservation of Structures and organizational missions
  - Preventing government's decapitation by maintaining (reduced) command and control
  - Preservation of existing bureaucratic/security structures and hierarchies
  - Increased organizational resilience against external influence or personal intimidation

(c) Protection against painful economic disruptions/damages
  - Reduction of economic disruption for defenders
  - Prepared methods to slow down detrimental decisions, accelerating beneficial decisions

(d) Protected (unaltered) access to or communication with citizens
  - Dependable announcement that a comprehensive cyberwar has started (CUCA)
  - Establishing (reliable) methods of authorized information flow to all citizens

(e) Maintaining capability for reliable actions during cyberwar (CWP-II) and its aftermath (CWP-III)
  - Preparing a command/control backup (i.e., underground) for retaking governmental control

(f) Protection of people
  - Preventing arrests of innocent people in bureaucracy, security, or leadership - except it is done by officers who have first-hand knowledge or irrefutable evidence of treason
  - Protection of people who have given information despite threats

Keeping secrets around most measures or processes is a waste because the adversary will gain this information anyway. However, electronic traces must be avoided proactively (as much as possible) if people are involved or at risk. The strength of preparation should come from making it public (open source) so that people in different positions can contribute with their detailed know-how or ideas to make it better.

(a) Information / Intelligence Gathering

The government needs to have reliable information on what the assailant demanded within intimidating calls to people as soon as possible (within 24 hours or faster). Informants must be given a safe method to report these demands despite vicious threats to their people and families.

I have listed below a few ideas in bullet-point form

- Routinely, e.g., every 4-6 hours, dedicated people in each organization would collect an envelope with a paper form in which informants can anonymously report demands by assailants. This form must be simple (e.g., checklist-type) for quick processing. Victims of these threats should not be encouraged to disregard these demands - their protection is more important if we expect reliable or accurate intelligence. As an additional rule, everyone must fill out these forms by hand in video-cam-free zones; only that way, more anonymity could be secured.
- Paper-based, standardized forms or reports are locally aggregated by removing empty forms to reduce reporting to a smaller size. The hand-filled, paper-based (summary) reports are then provided to the next aggregation level / office in person or by courier. Using only electronic means, incl. drones, could be a mistake even if some governmental systems are already hardened against Hacker-AI; the original forms / summaries should be preserved. Finally, the reports are collected/ aggregated in different independent main information/ intelligence offices and reported to the nation's political / military leadership so that they can decide on declaring a cyberwar emergency, i.e., if the country and its government are under cyber-attack (CUCA).
- Stationed analysts do threat-level assessments at these intelligence offices, preferably without IT support, i.e., mainly using wallpaper or erasable blackboards only. With hardened IT equipment, these systems should still be isolated, i.e., have no network access or unnecessary (USB) interfaces. Aggregation happens by regularly analyzing pre-sorted forms with summary forms, so tasks at higher-level aggregation/ intelligence offices are less labor-intensive. Threat-level assessments should be produced quickly. They should also include handwritten details on unexpected demands to improve the reports' quality.

- These analysts need certain character traits that help them work as a team over an extended time together. These professionals would create threat-level assessments multiple times per day. They would work without phones, and advanced entertainment (like electronic games), separated and isolated from their family for their own safety. After some time, they are rotated out, similar to missile launch crews responsible for nuclear weapons. This schedule would give assailants less chance to have them compromised and misused within operational goals.
- The locations of these offices can't be kept secret, but still, they should be protected against electronic surveillance. Some should be close to the government's leadership. Thereby, in-person meetings of the head of each office with the government's political and military advisors could be held regularly. From the outside, anyone could infer the threat assessment status from the frequency of meetings or who has been invited.
- Ideally, everything done by the information/intelligence offices is paper-based and handwritten. However, technology can make some processes more efficient and faster if hardened computers are available. Still, printing information is restricted - supervised by humans who know what they expect to get. Only old copy machines or scanners (with no network access) should be allowed and used.

(b) Preservation/Continuity of Governance

The assailant's main war mission is to decapitate the existing government - violently or via cyber-means by isolating the leadership or certain bureaucratic layers from each other. During this time, the risk is that the assailant will use fake orders or deep-fake audio/video calls to create new structures with new or compromised/intimidated leaders. Defender's goal is that the existing governmental/bureaucratic/security structures and hierarchies and their organizational missions must be preserved and operated in pre-determined modes.

Preventing government decapitation means that the command and control over countries' institutions must be maintained in a (potentially significantly) reduced intensity.

Because it is unlikely that technically not prepared countries can detect prewar CWP-I surveillance activities, I assume that only (some) demands (issued by AI bots on user devices) within the active CWP-II phase could potentially be reported (safely). The announcement that the "country is under a cyberattack" (CUCA) should only be issued if there is sufficient evidence that the decapitation is adversary's goal. Due to the worldwide consequences, the evidence must withstand being intensely scrutinized later.

The governance must significantly change its operational mode after CUCA is confirmed or declared. For example, meetings and orders should (or better, must) be given in person only. Note-taking or orders are handwritten (or done with an old typewriter). With hardened computer systems in the government's hands, some of these measures could be reduced to accelerate progress in preparing or configuring the country's defenses.

The government's overall goals are to maintain command, control, and resilient/stable organizations. In simple terms, a sieged government focuses on keeping the lights on. Additionally, it should not make spontaneous decisions based on rumors or emotions, i.e., no (unnecessary) changes from the plan that was designed before.

Unlike other wars, in Cyberwar 2.0, every decision, order, or change of laws/rules could be fake, and its authenticity must be questioned because of Hacker-AI, i.e., malware from Hacker-AI could manipulate data representations of events to its advantage. A freeze in making modifications to rules/laws is a significant limitation on the government's sovereignty; however, this is necessary due to the nature of Cyberwar 2.0 that was waged against the targeted country. A hardened government IT and communication system is potentially changing this situation, but not for insufficiently informed outsiders; in their assessment, every order, info, or change request could be fake.

Here are a few suggestions for concrete measures in bullet-point form

- Proactively defining/establishing rule simplifications and prioritization, including a reduced due process. This implies the freeze on major decisions (triggered by CUCA) must be maintained as a matter of principle. Except for emergency rules, the bureaucracy should not accept additional changes or decisions because assailants could generate a distorted picture of what is actually going on. Instead, clerks, officials, and officers could operate in a mode in which they are allowed, i.e., at their discretion, to cut through previous red tape in serving their citizens.
- After CUCA is declared, some critical (security-related) administrational processes must be turned into paper-/form-based processes (i.e., done by hand) so that malware has less impact on them. If these changes were not planned, thought through in advance, and decided and potentially trained ahead of CUCA - attackers would likely take advantage of the confusion around the trial and error during the initial implementation.
- Organizations must be resilient against undue external influence or personal intimidation; i.e., external issues should not impact the service people expect from that organization. E.g., multiple persons should al-

ways be prepared to take over any task if required. Additionally, organizational changes, particularly in key positions, should happen transparently (overtly) and be crosschecked with intelligence extracted from gathered data about assailants' threatening demands. Leadership changes are void if there is evidence of a plot or undue external threats.

- The lack of sufficient control over most governmental organizations could be stress-tested or spied out by attackers ahead of the actual war via fake attacks. However, CUCA should enable organizations to self-regulate/control their operations. They should also handle most of their problems from internal/external influence or fake orders for many months without direct oversight by the central government. Reviews on what happened during the cyberwar crisis could be adjudicated later, and all participants should be aware of that.

Unfortunately, the performance of operations at governmental organizations after the CUCA declaration will likely have no impact on the government's decapitation. The government overthrow is more likely accelerated by events that have nothing to do with operational activities. The theater leading to a new government controlled by the assailant is later assessed based on established (political) conventions and not how poorly the legitimate executive continued their governmental business.

Realistically, the old government is being replaced quickly by a new puppet regime; cyberwar activities will stop then. Preparation for the continuation of governance is probably a waste of effort. In Cyberwar 2.0, the attacker will win predictably and gain the necessary institutional power.

The most important goal by defenders is likely the announcement of CUCA to warn the world community of the cyberwar. If the government can provide evidence for its claim, it has probably scored an important victory before being finally defeated.

(c) Protection against Economic Damages/Disruption

Economic consequences for the assailant will come from sanctions, a less productive workforce, or via sabotage. Increasing assailants' costs without increasing the personal risks of people involved in acts of protest is extremely difficult and potentially unachievable.

Limiting painful economic consequences from disruptions or damages from assailants is largely outside the control of the government during war and aftermath phases: CWP-II or CWP-III. The government should still try to reduce economic disruption for defenders' population. Government's analysts should identify methods to slow down detrimental decisions for its citizens while accelerating beneficial decisions.

Suggesting concrete proposals is outside author's competence.

(d) Protection of Communication with Citizens

The government must stay in touch with its citizens, i.e., making reliable announcements. Also, as consumers of received information, citizens and businesses must be sure that they receive unaltered messages that the government has authorized. In short: People must trust the message, the medium (i.e., printed or verbal from an event), and the messenger, who the receiver should (at least) know.

Unfortunately, digital communication channels must be distrusted, even if the computer systems of the government and media organizations are hardened. Propaganda and misinformation could still create damage. The following preparation goals are for the war and aftermath phases (CWP-II and CWP-III):

(i) There must be a dependable announcement to its citizens that the country is in a comprehensive cyberwar (CUCA). After this declaration, citizens are informed and prepared that services or messages are compromised (i.e., under assailant's influence), and official follow-up messages could be deep-faked. It should be publicly announced that all electronic publications, incl. videos/TV and all not-in-person audio/video communication, could be faked and used for propaganda or disinformation purposes.

(ii) Tasks within civil defense measures are assigned (preferably) to teams of trained citizens. Potentially spontaneously formed (autonomous) teams could help maintain and improve society's living conditions when the government's more centralized command/control is failing.

(iii) The legitimate government has prepared reliable methods of authorized information flow down to all citizens based on word of mouth and redundant trusted messengers.

I acknowledge that I do not have a comprehensive plan on how the above objectives can be accomplished reliably and sustainably (under surveillance). However, a few suggestions should be made here in the following bullet points:

- New methods or processes using prepared publications and trained volunteers should be developed on an ongoing basis and tested/established in advance. All communication measures should be independent of IT technologies and created with minimal data traces to protect the people involved.
- Local (info) events should be cell phone-free. Only cell phones with clearly removed battery/ power supply are trusted. Otherwise, they are surveillance devices, even if manually switched off. Putting cell phones in tin foil reduces connectivity, but microphones can still record sound.

- Phone jammers are insufficient. Instead, EM-detectors could help locate devices based on their emitted network activity, incl. Wi-Fi and Bluetooth. But malware could deactivate these network activities while keeping the microphone or video cam active for recording. All mobile-device users must be told to be mindful about having these devices with them all the time.

(e) Maintaining the Capability to do Reliable Actions

During cyberwar (CWP-II), the government's leadership is under assault. I discussed the command/control and continuity of governance under item (b) preservation/continuity of government. However, the cyberwar will likely be lost, and a puppet regime will take control.

In the aftermath (CWP-III), surveillance continues and is likely being enhanced or made more overt by public surveillance measures if some people try to avoid being surveilled by their smartphones or personal devices. Avoiding surveillance could already be considered suspicious behavior. Taking action against the new order could become dangerous.

Still, the old government should not give up on helping their fellow citizens to regain their freedom and stop the intrusive AI-based surveillance. Even if it is unlikely that there is a back from total AI-controlled surveillance, there should be no stone unturned in which we study the reversal. The following bullet points contain some ideas:

- Preparing policies proactively for the aftermath (CWP-III) via establishing an (informal) command/control backup (i.e., an underground) could help people to retake their governance or mobilize people to start an uprising in a coordinated manner.
- Dedicated experts should prepare conceptional plans on what key positions need to be retaken or occupied by sympathizers. Also, it should be determined which technical key components should be deactivated, potentially sabotaged, that could favorably change the result of civil uprisings or help the resistance network.
- Small, independent teams or cells with people who know each other remain silent until they prepare actions when the signal for the uprising comes. These groups are trained and enabled to receive and send covert messages under surveillance conditions.
- The uprising could only be successful if Hacker-AI-based malware is being stopped on enough IT devices. Once persistent malware has occupied the hardware, it is too late for software-based security fixes. Only security hardware components included in devices as retrofit could make a difference. These security-hardware retrofits could be miniaturized to very small components, i.e., as small as network plug-

connectors (about 1 cm^3 or less). They would need to be smuggled into the country and distributed to people who want to regain control over their devices.

- Existing smartphones are likely, not retrofittable. They would need to be destroyed as potential spying devices and replaced by simple burner phones or new smartphones with security components.
- The idea of liberating a cyber-occupied country requires:
  (a) fixing/retrofitting IT equipment with security hardware around the same time in all devices so that early movers doing that would not become a target for the puppet regime's security,
  (b) removing power from all (non-retrofittable) mobile or IoT devices and
  (c) having a plan of switching off or removing adversarial control over public surveillance/ infrastructure IT systems via (dormant underground/resistance) teams focusing on assigned tasks.

However, the occupier or new regime could have changed many systems within the aftermath (CWP-III), expecting moves like those mentioned.

Returning to the old order/pre-occupation might be impossible - it is too early to tell.

(f) Protection of People

Citizens of the target country are in danger for three reasons:

(i) The new regime could arrest someone who is a member of the political leadership, i.e., politically opposed to the PRC. They could also be a key member of ROC's/government's bureaucracy or a member of the government's security (military, intelligence services, or police).

(ii) Anyone who spoke about being contacted by the assailant's AI bots ignored to comply with their demands and their threats. These threats might be real, and machines would not forget. Retaliation could be implemented in the war script and executed via drones almost automatically.

(iii) Once assailant controls security and the justice system, many more could be arrested because they fit a profile (i.e., being a potential saboteur) and therefore jailed in a re-education camp.

Events related to (i) and (ii) happen during the actual cyberwar (i.e., in phase CWP-II). Events related to (iii) are the aftermath (CWP-III); they are all outside the control of the legitimate (and replaced) government. There are a few ideas mentioned in the following:

- Everyone on the list mentioned in (i) should have a personal escape plan (with their family) and access to escape resources provided by the legitimate government before and after CUCA is announced. The announcement of CUCA could be interpreted by many citizens exposed because of reason (i) that they should leave with their families their country asap. Some of them may have a second passport that would give them diplomatic support from other governments. They may also have financial resources outside the country. The targeted country, in this scenario, Taiwan, could proactively negotiate treaties with its neighboring countries to handle predictable refugee issues via financial agreements or (secret) transactions.

- Preventing innocent leaders, bureaucrats, or security personnel (i) from getting arrested during cyberwar CWP-II falls under the legitimate government's purview. Strict rules should prevent normal security personnel from arresting a member of this "special category of people" without specially authorized officers involved and present. The arrests must be accompanied by at least two higher-ups from a special investigative department who must be involved in investigating crimes committed by a member of this special group. Alternatively, two or more investigators with first-hand knowledge of treasonous behavior must be present. At least one must give a written testimonial about the case, which is then provided to the dedicated aggregating/intelligence office - which has the authority to overwrite (at least temporarily) arrests and potentially also court decisions (like warrants) based on the quality of their reports and intelligence.
- There should be a national security rule about people who are qualified or authorized to give testimonials, i.e., being accusers against members of this special category of people. If the arrestees can prove via valid ID credentials that they are members of a special category with security or government-related responsibilities, then police officers must release them immediately or wait until two legitimate accusers appear. Without these accusers present in person, the arrest is invalid. If a police officer agrees that the above rule applies and the credentials check out, this should suffice to set them free. It is almost like being a member of a diplomatic choir. For Identification, all officers and executives within the special investigative department are stored in printed binders with images/personal data that police officers could consult when accepting or executing the arrest. The goal is that the release of these people could give people in this special group and their families a chance to leave their country immediately, even before CUCA is announced. The philosophy is based on the "No man left behind" rule of the US military so that everyone gives all they can until it makes no sense to continue. Unfortunately, with Cyberwar 2.0, the mentioned special group operates in their own country, which could turn hostile against them for no apparent reason except they belong to a group that is in the way of an external adversary.

- Everyone who thinks that the assailant’s AI bot contacted them on their smartphone/device should safely report this event via handwritten forms to be filled out (in video-free zones). A report must be provided and inserted in an envelope even if nothing reportable happened. Informants are not instructed to ignore or defy the demanded task, even if that act would violate their loyalty to the organization. It’s more important to keep informants safe during and after reporting.
- Everyone is at risk of being contacted by AI bots or deep-fakes; they should be trained and informed on how to react and report incidents (with details) as soon as possible without creating suspicion. The training for dealing with threats happens in video-free and cell-phone-free events to give people more freedom to discuss personal concerns without having Cyber Reconnaissance or surveillance know about it.
- People at risk should be trained to use their common sense when and how they inform non-verbally on their coercive situations, i.e., when they believe they should be taken out of a case related to a demand from the AI bot. Conversations should be trained to give or allow others to do the task instead. Without this training, it is unlikely that people know what to do. Training could be essential in throwing a wrench into assailant’s plans and thereby preventing many actions from happening smoothly via causing unexpected problems. However, disrupting a step within an assailant’s larger plan could be helpful but dangerous. If people do this, there must be a way to help them avoid repercussions and re-education camps.
- Training should give informants advice on how they identify themselves later as informants; this could be done with unique code names they are giving themselves, left in their messages.

It would be a significant success if many people targeted via arrest could be rescued before being thrown into re-education camps. Helping many people to escape the country affected by Cyberwar 2.0 or after CUCA is announced is considered an important long-term contribution and potentially a good investment into a better future for that country and the people left behind.

Generally Suggested Methods/Rules or Behavior

Being under constant (covert) surveillance via smartphones or other IT devices (audio, video, or usage) is a new quality in surveillance. Switching them off is often not enough, as electronic devices can appear to be switched off. People will deal with this situation differently. Some will surrender and accept the new reality, while others will give up on smartphones or IoT devices, keep electronic devices disconnected from the power supply, or remove their batteries.

The novel 1984 by George Orwell illustrates the risk of (actively) avoiding surveillance; it could make people doing that more suspicious. Still, personal methods and organizational policies could be implemented to deal with mobile devices already pre-cyberwar. I will suggest a few ideas.

- People could have less sophisticated burner phones instead of smartphones. However, switching between these phones should be supported by product features that could make SIM cards more easily exchanged between different phones.
- Organizations should offer more smartphone boxes for storing them temporarily and safely within their building. They need to be dampened so that recording surrounding sound is useless. Meetings are done mandatorily without smartphones. There should be (simple) detectors preventing the undetected presence of phones in meeting rooms.
- All offices should have parts of their space unmistakably marked as cell phone-free and video-free zones
- Pre-Cyberwar, some people could analyze their homes for potential IoT spy devices. They can start protecting their privacy immediately with the removal of devices. Taking this step after CUCA is declared is potentially a more dangerous move.

Not all people will be subject to constant surveillance, but they cannot know that (for sure). After the pre-war phase (CWP-I), attackers will likely narrow the surveillance focus to a few million citizens or less. Still, low-ranking clerks or operators could be on the list of potentially useful assets.

In Cyberwar 2.0, attackers use and misuse regular people as tools and collaborators. It is likely dangerous to refuse collaboration. Because it is war, the assailant gives orders to follow; they won't negotiate, but collaborators could be bribed, e.g., with encouraging benefits. These people should be prepared to become retribution victims, in particular, if they ignore orders or reveal what happened. Additionally, this bot could claim to be the perpetrator of some freak accident, which happens regularly by chance - or it could create fake news showing what happens in case of illegality. Therefore, a normal reaction to an attacker's demands or orders is full compliance.

A few people will, despite the risk, quietly signal to the outside that they are being coerced into collaboration. Here are a few ideas on how to support quiet dissent:

- Based on interviews with members of an organization, trained analysts should be able to determine the risk potential of having people (clerks, officials, security, etc.) within that organization being attacked in a cyberwar.

- Organizations should give coerced people a quiet method of dealing with being a collaborator. If people report and even show that they are coerced, they could become potentially double agents, in particular, if this is being reported up the chain.
- Organizations, i.e., people working there, must know via prior education about the significant personal risk they would accept by exposing themselves as a coerced collaborator to others. People must be trained via publications, education, and exercises to send or spot these potentially hidden signals.
- Additionally, organizations should have a culture of (sufficient) transparency to quickly detect suspicious (potentially treasonous) decisions/actions. This could be used to release pressure on some of the coerced collaborators.
- Providing advice on how to deal with intimidation within a cyberwar is essential. This training, together with role-play, should be done within the organizations by security professionals educated in this topic.
- As a rule and policy, all decisions or orders done or given under distress are null and void. Their outcome must be reverted immediately or soon after coercion has been detected and confirmed.
- Subordinates have the responsibility, even the duty, to decline orders from supervisors when they have doubts (via non-verbal signals or from the context) about coercion. However, at least one peer or collaborator must be consulted and accept or tolerate the declining of specific orders in their discussion (within a safe zone) to make the decision unassailable later. The compromised person's name should not be mentioned in these discussions.

Is There More?

For defenders, war preparation aims to deny the adversary his main goals while increasing the costs of this conflict. In Cyberwar 2.0, assailant's main goal is a rapid governmental overthrow and low follow-up costs from the war and its aftermath. Denying these goals in a conventional war is done by destruction and sanctions from the world community. However, cyberwar is a remote data operation with surveillance, intimidation, and misdirection designed to decapitate a government and its society and replace it with coerced puppets recruited from within the targeted population. It seems defenders will lose on all fronts.

It is very difficult to fight a war where the covert frontlines go trace-less through society's population. In cyberwar, the assailant will give orders that unwilling civilians must follow while it points out severe consequences when ignored. The assailant is determined to do whatever it takes to accomplish his mission. In war, this means accepting the death of innocent people, even actively killing innocent people. Announcing any form of opposition against the likely winner of the cyberwar is a dangerous decision that could later bring people into re-education camps.

Cyberwar activities are deniable and easily blamed on others. Physical evidence for cyberattacks will likely not exist. However, governments could create a paper trail from anonymous handwritten reports to aggregation at more centralized information/intelligence offices about people being coerced, from which trends or patterns could be derived. If governments have some hardened computer systems, they could send the scanned evidence to other intelligence services. But it will be important that the actual paper be brought as diplomatic mail to other countries, where it is stored and further studied. Only via diplomatic couriers, potentially involving other countries, the world community could receive physical evidence that led to the announcement of a cyberwar (CUCA). This outcome could already be considered an important victory under otherwise futile conditions.

Too late means sometimes what it says: too late. Preparing non-technically for situations that require technical solutions (in which we are presumingly too late) does not necessarily provide different outcomes. For targeted countries that are not technically prepared for a Cyberwar 2.0, preparation could, at most, provide the world community a signal that changes (like a government overthrow) within the targeted country were triggered by a confirmed Cyberwar 2.0.

(B) Preparations for Not-Directly Targeted Countries

A confirmed Cyberwar 2.0 would show that waging war is a (seemingly) risk-free decision. This message from that war could send shockwaves around the world. It is hard to imagine that we could be prepared for that in advance without immediately starting technical preparation on a massive scale. No country will be safe until comprehensive technical means make surveillance on smartphones and IT devices much more difficult or impossible.

The most significant difference between targeted and not-directly targeted countries is that non-targeted countries are probably not the target of total, mass-scale smartphone surveillance. The required size of the computational backend is currently outside technical capabilities. It is unlikely that a cyberwar operator could generate comprehensive data models of all people worldwide, but wait: Hacker-AI could use malware to misuse our own computer devices against us. Additionally, Hacker-AI operators might have already done reconnaissance missions to determine who is important to surveil more intensively. They may even monitor these individuals regularly or continuously, only to keep an eye on many people who could become relevant for whatever reason.

How intrusive Hacker-AI operatives are within not-directly targeted countries is difficult to predict. They might try to penetrate defense systems and make militaries' logistics partly or fully inoperable if they know they can't be detected. Also, assailant's malware might try to understand and then deactivate the (nuclear) retaliation system in a few critical key components without making these changes detectable.

I cannot claim to know whether Hacker-AI's malware can penetrate hardened military systems. However, if these systems use the same architectural principles as regular multithreading, multitasking systems, i.e., von-Neuman architecture, unified memory space, virtual address space, direct memory access (DMA), etc., and no physical separation of security and regular tasks or whitelisting of all apps before entering RAM, then it remains to be seen if malware can do it or if these systems are good enough to resist.

The objectives for countries not specifically targeted after the realization of Hacker-AI-based Cyberwar 2.0 as a viable form of warfare are likely:

(1) Increasing all security and defense measures to prevent the country does not become another victim of Hacker-AI and Cyberwar 2.0.
(2) Creating and protecting a safe environment in which Hacker-AI countermeasures can be developed, manufactured, distributed, and deployed.

Regarding (1): Nothing could prevent or protect a country from being the next target when it has not technically prepared its IT system from being overtaken by malware.

I would still suggest that the not-directly targeted countries adopt many of the already proposed methods of guaranteeing the continuity of governments. However, this is probably not sufficient after governments have seen how quickly they can be replaced.

Another/additional approach is to demand that smartphone use is reduced or even prohibited until new smartphones with security hardware are manufactured and widely deployed. This proposal would probably be too much to ask, but it is conceivable that an order like that is being issued.

Although it might be too late, some computer systems should be taken offline immediately and kept offline until security hardware solutions are used to protect them. Unfortunately, it is not sufficient to reinstall old software (from persistent storage media, like CDs) because if the system is compromised, it will remain compromised after reinstallation with any OS software.

Based on cultural, societal, or political consent, countries will respond to the threat of Cyberwar 2.0 and AI-based surveillance via smartphones in different ways. However, they all must hope that technical security tools will lead out of this problem soon. It is conceivable that countries will use their war mobilization act to create tools to help adapt their bureaucracy, economy, and citizens to the new normal. Inevitably, new tools must be developed and finally deployed within step (2).

I already discussed solution development (2) in the previous chapter. However, every piece of preparation in the absence of Hacker-AI would be tremendously helpful.

The most important goal related to non-technical measures is to safeguard engineering skills, manufacturing capacities, and services, including all steps toward deployment against nefarious malware interferences in almost every imaginable way. The problem with being late is that the adversary might prove late to us that we are too late.