GIS and its Information Security Applicability

Author: Alexandru Mircea Rotaru, Larisa Gabudeanu

What is GIS?

Geographic Information Systems (GIS) represent data as points on a map. Most commonly, it helps monitor how incidents cluster and correlate, thus informing its users about where they need to invest effort and resources next. This way, your organization ensures every dime and every second used is accounted for at all times, keeping all wastes at a minimum.

A Brief History of GIS

Though the first modern GIS only appeared in 1963[1], its principles date back to Victorian England. A doctor by the name of John Snow (sadly, no iron throne appears in this story) sought to find the source of a cholera outbreak in London's suburb of Soho. He mapped out the cases he found, and most converged at a nearby water pump. Snow then used these findings to show how cholera transmits through water. As a result, when the authorities removed the handle from said water pump, the nearby cholera cases sharply decreased[2].

GIS Today

With time, more and more fields adopted GIS, for a wider range of uses. Public health still uses GIS to monitor various vector-borne diseases, such as West Nile Virus and Malaria. Part of the Public Health Accreditation Board requirements for any Health Department in the United States to receive accreditation involves using GIS in daily operations[3]; this includes over 200 state, local, and tribal health departments in the US spanning 46 states and the District of Columbia[4]. Some nations also use GIS to map COVID-19 cases for contact tracing.

Beyond public health, GIS can be used in land surveying and mapping, in natural resource exploitation, in logistics and transportation, in housing and urban development, and in everything else that needs to identify geographic areas where data clusters or correlates.

[1] https://www.esri.com/en-us/what-is-gis/history-of-gis
[2] https://www.ph.ucla.edu/epi/snow/snowcricketarticle.html
[3] https://phaboard.org/wp-content/uploads/PHABSM_WEB_LR1-1.pdf, p. 44
[4] https://phaboard.org/who-is-accredited/

That's all Fine and Dandy, but what does GIS have to do with Information Security?

There are two ways in which GIS and information security management can rely on each-other to work better. On the one hand, you can use GIS to monitor and prevent threats by identifying which geographic areas pose the greatest vulnerabilities for your organization. Once you have collected the data, you need to represent it somehow, and if a map makes the most sense, you need to use GIS. You can then use this information to intervene before anything goes wrong.

GIS will also help you correlate data, and incidents in your area as a whole. That way, your organization will know that, should they impact factor A, it could also have a collateral impact on factors B and C. For instance, if you see that in communities with high rates of smoking there are low rates of high-school graduation, implementing programs to keep children in school would likely reduce the rate of smoking in the neighborhood. Likewise, if you see that in communities with high alcohol consumption rates there are low anti-depressant consumption rates, implementing campaigns to reduce the amount of alcohol consumed would likely increase anti-depressant usage.

In short, the various factors GIS monitors do not happen in a vacuum, and any collateral effects of any interventions (or lack thereof) are the organization's responsibility. Using GIS to identify reasonable correlations would at least help you identify a solid amount of these collateral effects, so that you can intervene.

Therefore, GIS can be used to provide a greater overview of past incidents affecting the confidentiality, integrity, or availability of data, taking into account incident data available through incident reports, logs, and circumstantial information. This is useful to identify root causes and remedies given the entire set of circumstances surrounding an incident. Furthermore, previous incidents and their geographic locations can be correlated with other information, such as identified vulnerabilities and threat modelling, to prevent similar incidents from happening in the future. This is particularly useful for certain types of organizations, such as agriculture companies using IoT and petroleum extraction companies using sensors for quality of services and security, as well as for business continuity and disaster recovery planning.

On the other hand, GIS still collects and processes data – often personal data – no matter what you use it for, so you need to keep that data safe. Most countries have devised some sort of data privacy laws, such as GDPR in the European Union, and HIPAA and FERPA in the United States. Every information system risks getting breached on the daily, so, when working with personal data, you need to make sure it is safe wherever it may be used, including in GIS.

For the European Union, these are the most essential principles you need to take into account:

- **Data minimization:** The organization should limit the data collected or processed through GIS only to that personal data needed for the purpose of the GIS use. For example, if it is relevant to know the household contact details or behavior data for the purpose of the organization's activity, these should be collected. Otherwise, if such data is not needed (i.e. the activity of the organization can be fulfilled without this data), it should not be collected.
- **Need to know principle:** The access to the GIS data should be limited only to the individuals that need access to perform their job description, at the lowest level of access that would allow them to fulfill their functions within the organization. This applies to both employees of the organization and the employees of any third parties working on the GIS data.
- **Purpose limitation and basis for processing:** Data collected for one purpose (such as traffic monitoring) generally cannot be used for other purposes (such as marketing, creating heat maps of households). In addition, the organization has to identify if there is a basis for data processing before it starts any actual processing of the data; these include fulfillment of contract with individuals, legitimate interest, and public interest.
- **Limitations in data disclosure:** The organization should have limited third-party involvement, only when needed and within the scope and purpose of the data processing. When transferring to other data controllers, the organization should analyze if there is a basis for such data processing.
- **Proper transparency of data processing:** The organization must inform any individuals whose data gets processed, per the transparency requirements, about how their personal data is being used. Furthermore, when consent is needed, it must be properly obtained prior to any data collection.

How can you Avoid a Collision with Something you don't Notice?

Per Murphy's law, risks never go away. This is why many organizations today prefer to be reactive when addressing threats and breaches. Though, wouldn't it make more sense to know where threats can come from so they don't catch you off-guard? That's exactly where GIS comes in.

Still, organizations may not see why they would need to invest time, energy, and funds into getting an operational GIS system in order. Going with the collision example, your eyes can help you see obstacles and incoming traffic if you're driving a car, and some may say that's enough. However, what if you are flying a plane at half the speed of sound? Cyber security incidents are more like an airplane flying through a flock of Canada geese than a road accident, as they can occur very quickly, and you are likely going down and have to prepare for disaster by the time you notice them.

So, how can anyone expect pilots flying at half the speed of sound to notice the flock of geese in the way using nothing but their eyes? They don't. Instead, the airline industry adapted, with radars, forecasts, and air traffic control making sure planes don't crash into each-other, fly into dangerous updrafts, or collide with any local birds. Also, in case something does go wrong, airplane manufacturers design engines that withstand the impact of flying through birds as a last resort.

So, why shouldn't an organization invest in something similar to prevent cyber security incidents and other risks? This is where some sort of monitoring and prevention system comes into play, particularly when operating with personal or confidential data. If the threats are related to roads, geographic areas, or underground resources, a GIS system will serve that role.

Data Security is Key

Naturally, every monitoring and prevention system comes with drawbacks of its own. Otherwise, we would all be using these perfect systems, and information security as a field would not even exist. Sadly, we do not live in that perfect world, and your organization still needs to look out for threats whenever using information systems, including GIS.

In many cases the data that gets analyzed through GIS is personal data and/or confidential, mainly through its identifiers - such as names, addresses, phone numbers, social security numbers (or equivalents), and bank account numbers. By identifier, we mean any data that can reasonably lead to inferring an individual's identity; this definition includes personal data. Particularly when doing analytics using GIS to see which geographic areas or points of interest on a map need the most attention from the company, you need to secure the data in case of a cyber security incident. This is in line with the data minimization principle, which entails collecting, storing and processing only the data needed for the data processing purpose.

The easiest way to do so is to remove all identifiers before analysis. In the United States, studies in the sciences and social sciences require that the final results be stripped of any and all identifiers, so it may very well be part of the local regulations to do that. However, if that is not possible, you need to secure the data and the identifiers.

Increase Security the More Confidential the Data

With GIS, you can layer data, and select who gets to see what. That way, if your organization's information systems get breached, you can contain the attackers to the level they penetrated. With GIS, you can have the lowest level of security be the heatmap with no identifiers, and then provide higher level access to more confidential data as needed. That way, if the breach only reaches the "dots on a map with exact address labels" without knowing anything about what exactly each dot represents or the details of any sub-category said dot may be in, you can contain the breach at that level. However, every second matters, and if the breach gets to the highest level of security, it's time to call the company's management, lawyers, and incident handling team.

Nobody Believes Organizations that Publish Data that's been Tampered with

One reason your organization's GIS system may be breached is to steal personal information. However, particularly in research, public policy, and public health, using data that's been tampered with is the fastest way to kill your organization. Tampering with data can happen for multiple reasons, which mostly boil down to your competition trying to undermine your organization's credibility, a cover-up of how bad the situation truly is (or how good it is, to have access to more funding that would otherwise be inaccessible), or an attempt to increase or decrease the monitored issue's level of urgency for the stakeholders by eliminating existing correlations or fabricating ones that don't actually exist.

This is why you need to set clear protocols as to saving multiple backups of data and working with the data, to ensure that, when someone does tamper with the data, your organization will be able to pick up from where it left off. You also need to make sure your organization is able to identify which sets of data have been tampered with - a quick cross-reference with the backups and/or logs should do.

Also, if you find that your data has been tampered with and there is no unauthorized access to data from outside your organization, you and/or the organization managers might want to have a long conversation with the staff. In the best situation, a quick refresher on how to manipulate data without altering it should do. If you have a malicious employee, you need to make unpopular decisions that limit liability. If nobody knows how the data got tampered with, you might want to upgrade your organization's locks, security cameras, logging and monitoring, as prevention is key in cyber security.

The Data Correlation Process is a (Hacker's) Dream

Even though data correlation is one of GIS' biggest assets, it's also one of its greatest vulnerabilities. The entities that breach your informational systems aim to get as much out of it as possible, and correlations provide multiple sets of data that are also somehow connected. Your organization's data is at the greatest risk when you do the correlation process, so you need to pay extra attention to what is going on during those times. Also, make sure that the data sets you correlate are then separated as soon as you finish the analysis, thus limiting the amount of time they stay correlated.

Don't Forget the Obvious – Confidentiality

Imagine you're presenting your findings about anti-depressant use in your area. You made sure your data is secure, and that nobody breached or tampered with it. Only, you left the data as dots, and one of them is on your address. Next thing you know, you have a PR scandal and everyone calls your sanity into question because you forgot the obvious: when presenting GIS data, make sure to eliminate all identifiers– especially the addresses. Anonymization techniques should be applied to the data. However, this requires a balance between generalizing the data in order to reduce the possibility of individuals being identified and keeping the data useful for the data processing purpose. GIS data provided to the public should, generally, be anonymized completely; many organizations today use heatmaps instead of the actual dots to do just that.

Insight without Action is Worthless

The whole point of identifying threats and weaknesses in your organization's system is for your organization to do something about it. GIS will help inform your decisions and the consequences they will likely have, but it will be all for nothing if your organization doesn't use them. So, before investing in a shiny new GIS, make sure you know what you are using it for and that you have a process in place for integrating data from GIS in its usefulness in the incident prevention process, the business continuity analysis, and the risk management process. As a rule of thumb, if you begin with the end in mind, your work will yield far better results.

So, how do I incorporate GIS in my Organization?

First, you need to ask yourself: what would I use GIS for? If you work with maps, natural resources, transportation tracking, the weather, and the like, GIS will prove ideal for you.

Furthermore, for some of these use cases, GIS can assist with implementing legal requirements such as those under the NIS Directive. Then, you need to implement protocols for using the GIS, for handling the data without altering it, for security levels, and for what to do when a cyber security incident happens. Also, people from different backgrounds using GIS can each develop their own set of abbreviations, which would turn using GIS into a nightmare akin to translating dead languages. To avoid these kinds of situations, a list of all organization-wide standard abbreviations for the GIS would help make sense of what is going on.

Finally, you need to have a GIS technician on your team. You can either hire someone from outside or get one of the present employees certified; it's a very intuitive system that shouldn't take much time to master, and many institutions in Europe, the United States, Canada, Australia, New Zealand, and other parts of the world offer such certifications and/or coursework for credit.