Devs Against the Dark Arts

attacus (Lilly Ryan) — 🧿@attacus_au

(**N.B.** Please be warned that this paper will contain spoilers for all of the Harry Potter books. If you're not familiar with these stories, some explanations are included, but you're welcome to put this aside, go and read the books, and come back to it later, if you like.)

Greetings, student wizard.

Today's *Devs Against the Dark Arts* class is an introduction to getting into the mindset of Lord Voldemort so that we can write better web applications.

> “ It was once my job to think as dark wizards do.
>
> *"Alastor Moody",* ***Harry Potter and the Goblet of Fire***
>
> ”

It is a pentester's job to think as malicious Internet users do.

Courses on how security vulnerabilities work and how to mitigate them can often pack in a lot of technical detail, but understanding security concepts doesn't have to start there. It doesn't have to even be about software. We can learn a lot about app security by imagining that we're fighting evil wizards, because the basic concepts are almost exactly the same.

We're going to look at this by analysing how Professor Albus Dumbledore, the Headmaster of Hogwarts School, and widely recognised as "pretty good at magic", thought about and constructed defences around various important objects and places in the Harry Potter books.

Then we'll look at how Lord Voldemort - a.k.a. "The Bad Guy" - approaches his own security, examine the difference between the two wizards' security practices, and see what we can learn from them both when it comes to web app security.

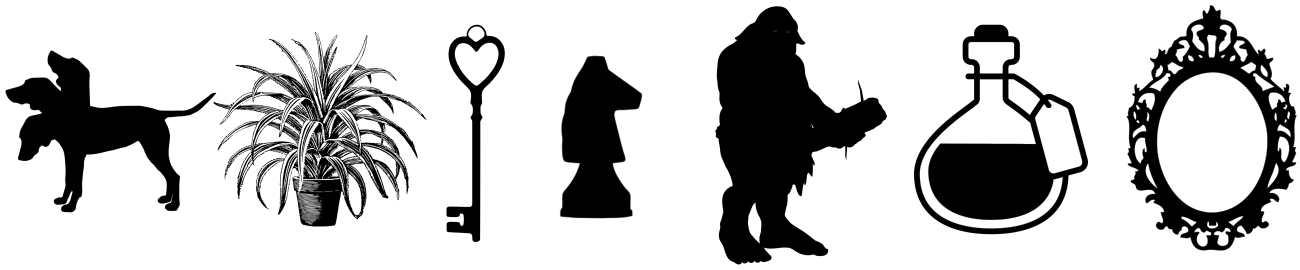
Case Studies

I've chosen three security case studies for us to look at in order to pick apart Dumbledore's defence methods.

These are:

- the guarding of the Philosopher's Stone, in the first book, *Harry Potter and the Philosopher's Stone*
- the security mechanisms around the Goblet of Fire, in the fourth book, *Harry Potter and the Goblet of Fire*
- the defences around the house located at Number 12, Grimmauld Place, London, in the fifth to seventh books of the series

The Philosopher's Stone



The Philosopher's Stone is a shiny stone with magic powers that the good guys have. Lord Voldemort wants the stone because it will help him kick his evil goals.

Dumbledore is tasked with protecting this stone from Lord Voldemort, which he does by getting several of his staff members to set up a series of seven defences in front of the Stone, in sequential chambers hidden behind a classroom door in a third-floor corridor at Hogwarts.

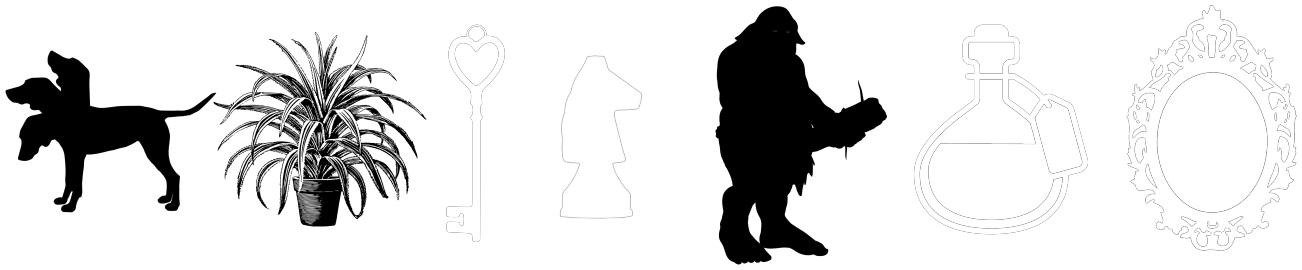
The seven defences are (in the order that they are encountered by anyone trying to reach the Stone):

- a three-headed dog guarding a trap door
- a murderous plant called a Devil's Snare
- a room full of flying keys
- a giant Enchanted Chess Board
- a literal troll
- a table with seven potions and a logic puzzle on it
- a magical mirror called the Mirror of Erised

By layering different defences on top of each other to protect his sensitive content (i.e. the Stone) Dumbledore is using the strategy of **"defence in depth" - multi-layered defence strategies working together to protect sensitive information**.

"Defence in depth" is an established security pattern. But how well was this pattern implemented when it came to protecting the Philosopher's Stone?

Physical Defences



A few of these defences are things that are broadly classed as physical defences. These are:

- the three headed dog
- the Devil's Snare plant
- the troll

These things act just like locked doors and giant walls with spikes on the top: they're meant to put off opportunistic attacks and buy the defenders some time. However, they're also supposed to be difficult to penetrate. This is not the case for any of the defences used here, which all have widely known vulnerabilities and bypasses.

The three-headed dog is fierce and scary, but it also goes to sleep whenever it hears music. This is common knowledge because the dog's owner, Hagrid, has terrible OPSEC[†], and told pretty much everybody.

The Devil's Snare is a plant that will bind you tightly to stop you from going anywhere, but it also recoils in the presence of light, which is a common fact that you can read in wizard textbooks for eleven year olds. If you bring a Zippo or a wand with you, you'll be fine.

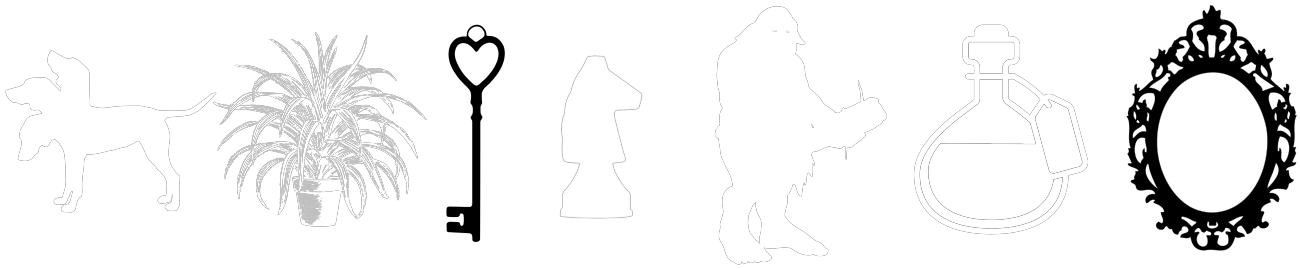
The troll is… a troll. It doesn't bathe, isn't very intelligent, and probably says hurtful things on the Internet. It's vulnerable to becoming unconscious if it is hit on the head really hard. Also, if it is knocked out, the troll doesn't reset itself or get replaced with a fresh troll, so after Voldemort and his henchman knock it out, when Harry, Ron, and Hermione find it, it is still knocked out and they can just walk around it. **Your defences should not be one-use-only, especially if there is nothing else reinforcing them at the same time.**

Finally, before you even get to this set of chambers full of allegedly Dark-wizard grade magical defences, the whole thing is protected by a wooden door in a frequently used corridor, and it opens with a very basic unlocking spell, which they also teach in class to eleven year old wizards. It's so simple that our protagonists actually end up in this room *by accident* once. **You don't want to implement defences that can be breached because an eleven year old**

panicked.

(Also, the books are filled with portraits that enforce magical passwords. *Why* one of these wasn't used here is completely beyond me.)

Obfuscating Defences



Next, we run into another common pattern that is designed to stop attackers by being just annoying enough that most people will give up and go away. This is done by treating sensitive information like a needle in a haystack, also known as "security through obscurity". The info is right there, but it will take patience to go through everything to find it.

You see this kind of thing a lot on the web, where apps will make sensitive documents or endpoints public, but give them weird names and hope that nobody will spend the time to try lots of different combinations to find them. This is a commonly used method for hiding sensitive information because it is quick and easy to implement, but it is *also* not a very strong defence, and has some easy bypasses, as you're about to hear.

We can see this pattern used in the case of the room full of flying keys, and also in the case of the final defence, the Mirror of Erised.

Harry, Ron, and Hermione are able to access the flying key that opens the door to the next chamber by spotting it amongst the other keys and then using a broomstick to fly up and grab it. Because the room contains multiple broomsticks, it allows for multiple threads of attack to run at the same time, because more than one person could be in the air looking for the key at any point. There is no rate limiting in place here at all. In fact, multiple simultaneous attempts are explicitly *supported* because the room has a bunch of extra broomsticks just lying around. As a defender, this doesn't make much sense.

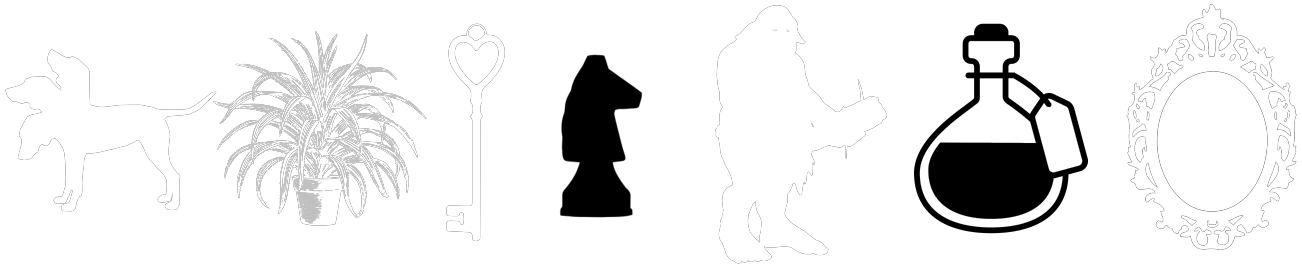
It's the same thing with the Mirror of Erised. The mirror contains the Philosopher's Stone, but it will only release the Stone to somebody who wants to find it, but not use its powers. Voldemort and his henchman can't get the Stone out of the Mirror because *they* want to use it. But Harry *can* get the Stone out of the Mirror, because he just wants to find it in order to stop Voldemort.

Ironically, Harry would have achieved his goal more easily if he hadn't gone after the Stone in the first place, because it was his presence in the room that allowed the Stone to be released from the Mirror. Once it was out, the bad guys tried to grab it from him. It was only Dumbledore turning up at the last minute to fight them off that stopped this from happening, like a server admin SSHing in to do a live patch in prod.

Likewise, there didn't appear to be any rate limiting on this Mirror. Theoretically, Voldemort could have grabbed the mirror, taken it out of the chambers, and then set it up and have a long line of people look into it until one of them fulfilled its conditions - that is, they wanted to find the Stone, but not use it.

Then he could have hit them over the head with a brick, and stolen the Stone.

CAPTCHA Defences



Our last type of defence in front of the Philosopher's Stone is what amounts to a couple of CAPTCHA tests.

CAPTCHA stands for "Completely Automated Public Turing test to tell Computers and Humans Apart", and it classically looks like a collection of warped text. More recently, it tries to get you to train self-driving cars. It's supposed to be a task that only an attentive human would be able to complete, in order to stop a bot or a script from using the feature it is protecting.

These CAPTCHA-style defences are the enchanted chess game, and the logic puzzle with potions.

The enchanted chess game is just a game of chess, except all the chess pieces are animate and will knock you out if you are standing in for the piece that gets taken. Knowledge of how to play chess isn't secret - it's just a task that takes time and concentration to complete. This slows you down a little bit, but it doesn't guarantee you will be stopped from getting to the other side.

In the same way, the logic puzzle with the seven potions is something that can't be brute forced easily. First, you have to solve the logic puzzle to determine the contents of each of the bottles of potions. You get *one* chance to drink the potion out of the lineup that will allow you to either go forward to the next chamber or go back to the previous one, so you need to make sure that you've

made the right choice. Furthermore, if you're thinking that you could just drink them all and YOLO your way out of there, *one* of the bottles contains a poison that will kill you, so it's not a good strategy. This task, like the chess game, is designed to make sure that the person thinks about it, but it isn't designed to be all that difficult.

Both of these things are also designed to take time. You can't just run through the room to the next thing; you are forced to stop and engage on an intellectual level with what is going on. If you implement a defence that takes time to bypass, you give yourself more time to respond to an attacker.

On the other hand, these defences aren't all that difficult to bypass. All the information needed to get past them is either common knowledge, in the case of how to play chess, or right there in front of you, in the case of the logic puzzle. And much like real CAPTCHAs, they're incredibly annoying, especially when you have to deal with them on your way in and you're a legitimate user.

The main point of employing defence in depth is to buy you time to respond to an attack, if you are a defender.

The main downfall of all of the above mechanisms is that you can't buy time to respond to an attack if you don't even know the attack is happening, which is exactly the case with the Philosopher's Stone.

There is no useful logging, monitoring, or alerting going on here at all: no ghosts running around the halls screaming that someone broke into the room on the third floor corridor, no owls showing up to deliver messages to the wizards on call.

The ridiculous thing about this is that we *know* the wizarding world has access to spooky levels of monitoring and alerting, because any time an underage wizard does magic outside of Hogwarts, they get an owl with a letter about it almost immediately, right down to the timestamp and the name of the spell performed.

This happens to Harry Potter at the start of *the Chamber of Secrets*, at the start of *the Prisoner of Azkaban*, and also at the start of *the Order of the Phoenix*. Later in the books, Voldemort also gets alerts every time someone says his name.

For Dumbledore not to have implemented anything like this is a serious oversight in the protection of the Philosopher's Stone. Just because he turned up right at the end to save Harry from being killed by Lord Voldemort, it doesn't make him a hero - it makes him a kind of negligent ops engineer who got lucky. Dumbledore should have been waiting for Voldemort before he even got past the first room.

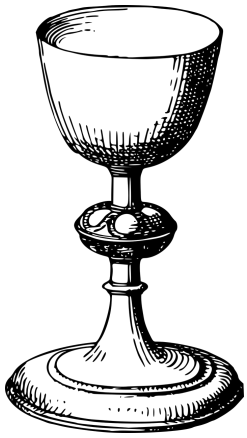
Not to mention that because of his design, he had to fight past all of his own defences *the same way* in order to get to the Stone, just like any old attacker, which doesn't seem like the most efficient way to do things.

Assessment

Philosopher's Stone: 4/10

Okay attempt at using defence in depth, but uses multiple known vulnerable defences, and is severely let down by a lack of monitoring and alerting.

The Goblet of Fire



The Goblet of Fire is a magical cup with the job of deciding which students will compete in the Triwizard Tournament, which is an inter-school magical competition. It appears in the fourth book, *Harry Potter and the Goblet of Fire*.

Students can write their name and school on a slip of paper that goes into the cup. Submissions are open for 24 hours, and then there is a ceremony where the cup will return the names of one candidate per school, who are then bound to compete in the Triwizard Tournament.

It was decided that the only students who could compete in the Triwizard Tournament during the course of Harry Potter's fourth year at Hogwarts should be aged 17 and over, because the competition was so dangerous. The only other rule for selection of candidates for the Goblet was that there should be only one candidate chosen per school. Traditionally, three schools compete in the Triwizard Tournament, hence the name.

The plot of *the Goblet of Fire* revolves around the fact that fourteen year old Harry Potter is mysteriously chosen to be the fourth Triwizard Champion. How this happened, when Harry was under the age of seventeen and a champion from his school had already been selected, is revealed at the end of the story, and we'll take a look at the vulnerabilities that were exploited there in just a second.

Firstly, let's have a look at the Goblet's configuration and security settings.

We aren't given much information about how the Goblet makes its decisions about the Triwizard candidates, but we do know that Dumbledore put one defence mechanism in place around the Goblet - an Age Line. This was supposed to stop anyone under the age of 17 from entering the space and putting the paper with their name into the Goblet.

The Age Line is a literal line on the ground that can only be crossed by people aged 17 and over. Otherwise, there's nothing to stop anyone approaching the Goblet, which is left unguarded for 24 hours in the middle of the Great Hall at Hogwarts.

> "To ensure that no underage student yields to temptation," said Dumbledore, "I will be drawing an Age Line around the Goblet of Fire once it has been placed in the entrance hall.
> Nobody under the age of seventeen will be able to cross this line."
>
> *Albus Dumbledore,* ***Harry Potter and the Goblet of Fire***

Despite setting this up to best temptation, people try to bypass this control mechanism almost immediately. In the 24 hours it is open, we learn that four students try to trick the Age Line by taking an Aging Potion to make themselves several months older. This doesn't work, and the spell throws them out of the ring immediately.

A successful Age Line bypass is made on behalf of Lord Voldemort by Barty Crouch Jr, in disguise as Mad-Eye Moody, one of the Hogwarts teachers. He enters Harry Potter's name into the Goblet, and is successful in having Harry chosen as a fourth Triwizard Champion, which is part of Voldemort's evil plan.

Crouch circumvents the protections on the Goblet in two ways:

Firstly, he is over the age of 17. There's nothing in the books to say that a person over the age of 17 couldn't enter someone else's name into the Goblet. It's not clear if any other students tried this, but if they did, they weren't chosen so we don't know about it.

Going on what we *do* know about Crouch's hack, the Age Line only checked if the person crossing it was over the age of 17. It didn't check if that person was adding their own name to the Goblet, or someone else's.

The authorisation of app requests is important, especially when those requests are making changes to a database. The Goblet didn't have any kind of authentication or authorisation protections whatsoever, which is a serious oversight.

> It would have needed an exceptionally strong Confundus Charm to bamboozle that Goblet into forgetting that only three schools compete in the Tournament … I'm guessing they submitted Potter's name under a fourth school, to make sure he was the only one in his category …
>
> *~~Alastor Moody~~ Barty Crouch, Jr.,* ***Harry Potter and the Goblet of Fire***

Secondly, Crouch uses a Confundus Charm on the Goblet to trick it into thinking that there is a fourth competing school in the competition, and that Harry Potter belongs to the school. Because

Harry Potter is the only candidate for the fourth school, he is chosen to compete in the Triwizard Tournament. There don't appear to be any data injection defences on the Goblet of Fire at all, or if they are, they're not very strong ones.

If the Goblet had validated the names of the schools that were submitted against its own known list of schools, it may have been more likely to reject the presence of a fourth school, and then Harry wouldn't have been guaranteed to be chosen.

Also, the data that was entered into the Goblet should have been sanitised before it was added, and then again for bad data before the final vote was run. This could include stripping out data with the names of bogus schools as well as matching and removing names that corresponded to students under the age of 17.

The long and short of this is that **you should never trust user input into your app**, or your Goblet of Fire! Your user might be legit, or they could be submitting POST requests on behalf of Lord Voldemort!

Assessment

The Goblet of Fire: 2/10

The Age Line worked well against some basic attacks, but without authorisation and authentication, and without validating user input, it was vulnerable to data injection and the plots of evil wizards.

12 Grimmauld Place

For our last case study, we'll examine the protections that Dumbledore placed around Number 12, Grimmauld Place, London. This address is the house of Harry Potter's godfather, Sirius Black. For a few years, it acted as the headquarters of the Anti-Voldemort activist group the Order of the Phoenix.

Because of its status as Order HQ, the house became a target for Dark Wizards everywhere. Lots of prominent members of the Order, including Dumbledore, Severus Snape, and Harry Potter, were known to spend time there. This made it important to defend properly, so the Order looked to its leader, Dumbledore, to harden the configuration of the house as much as possible.

The main defence that Number 12 had going for it was the Fidelius Charm. In the words of one of the Hogwarts teachers:

> “ [The Fidelius Charm is] an immensely complex spell involving the magical concealment of a secret inside a single, living soul.
> The information is hidden inside the chosen person, or Secret-Keeper, and is henceforth impossible to find — unless, of course, the Secret-Keeper chooses to divulge it.
>
> *Filius Flitwick,* ***Harry Potter and the Prizoner of Azkaban***
>
> ”

The secret, in this case, was the location of the Headquarters of the Order of the Phoenix.

Dumbledore was the Secret Keeper for this particular spell, which effectively meant that he was the only one who could tell someone where the Headquarters was. Once someone else knew about it, they were unable to pass that knowledge on, even if they were tortured or coerced.

The main problem with this protection is that the Fidelius Charm has a single point of failure.

That's the Secret Keeper. That person has to be the one to tell *everybody* who needs to know the secret.

However, Harry Potter learns about Number 12, Grimmauld Place because Dumbledore writes the secret down on a piece of paper, hands it to someone else, who then passes it to Harry. This is an extremely insecure method of transport. This note could have been accidentally dropped and read by *anyone*. It was totally in plain text. The information on it is only kept secure after Harry reads it because the note is set on fire by the person who gave it to him. If they hadn't done that, presumably the note could have been handed around any number of people and the secret would have been shared multiple times. Dumbledore clearly had not heard about transport encryption.

The other problem with the Secret Keeper method of keeping Number 12 safe is that when the Secret Keeper dies, anybody who was told the secret automatically becomes Secret Keeper themselves. The Order of the Phoenix runs into this problem after Dumbledore dies. In *Harry Potter and the Deathly Hallows*, the protagonists explicitly have to deal with the fallout from this:

> “ Mr Weasley had explained that after the death of Dumbledore, their Secret Keeper, each of the people to whom Dumbledore had confided Grimmauld Place's location had become a Secret Keeper in turn.
> "And as there are around twenty of us, that greatly dilutes the power of the Fidelius Charm."
>
> ***Harry Potter and the Deathly Hallows***
>
> ”

This is not *entirely* unlike someone quitting their job and checking the AWS keys into `git` at 5pm on their last day. **Shared keys are weak keys, because the more people who know the secret, and the more machines that store it, the more likely it is that someone unauthorised is going to obtain the secret and use it to mess up your whole operation.**

In the worst case, if this happened to a software key, we could revoke the original key and rotate it to a new secret value. However, the Fidelius Charm doesn't appear to have the concept of key rotation. There's no way to undo the Secret Keeper spell, unless everyone who knows it dies, or the house it applies to gets knocked down. This is pretty much the equivalent of getting hacked, throwing your hands up, buying a new domain name, and starting over.

Additionally, we learn over the course of the final few books that Dumbledore strongly suspected he was going to die soon. Knowing this, he *still* set himself up as a single point of failure for the safety of the Headquarters of the Anti-Voldemort movement.

This is a jerk move. No points to Gryffindor.

If you know you're going to leave a project, and you're a decent person, you don't

make everything depend on your keys and then just weaken them all with no alternatives or backups. You do handovers, you plan for disaster recovery scenarios, and if you're actively at war with Literally Lord Voldemort, you don't just leave everyone to scramble to find a new secret Headquarters if you die, unless you're a jerk. (Yes, I'm calling Dumbledore a jerk.)

Assessment

12 Grimmauld Place: 0/10

No disaster recovery options, secrets can only be transmitted insecurely, and keys can't be revoked.

Voldemort's Security Model

If Dumbledore, thinking like a "good" wizard, was unable to get into the mindset of a Dark wizard, then how *do* they think? And how does that help us do better?

Let's compare the previous case studies to Voldemort's security model.

Before the start of the Harry Potter books, Voldemort decided that he needed to make backups so that he could recover himself if something bad happened and survive attempts by other wizards to knock him offline.

He did this by creating Horcruxes. Here is how one Hogwarts professor describes Horcruxes:

> “Well, you split your soul, you see, and hide part of it in an object outside the body. Then, even if one's body is attacked or destroyed, one cannot die, for part of the soul remains earthbound and undamaged. —
>
> *Horace Slughorn,* ***Harry Potter and the Half-Blood Prince***
>
> ”

Voldemort decided to create seven Horcruxes, because he knew that if someone tried to murder him and knock one of his Availability Zones offline, he could just failover to any of the other ones while he resolved the incident. He also decided that these Horcruxes needed to be stored in geographically different places so that it would be more difficult to take them all out at once, so he went to a lot of effort to ensure they were placed safely out of the way in a wide range of locations.

The main weakness in the Horcrux method, which was the thing ultimately used to destroy Voldemort in the end, is that he used personally relevant information when he created them. He chose artifacts that were significant to him, and could be identified with some strategic intelligence gathering. This meant that Dumbledore was able to deduce the existence of the

Horcruxes and what they were likely to be, and pass that information on to Harry so he could carry on destroying them.

The other problem with Voldemort's system was that, despite having mastered the art of monitoring and alerting in his day-to-day operations, he had no alerting on the status of the Horcrux infrastructure, which was a real oversight. He set up lots of defences around the Horcruxes, and just assumed they would defend themselves. This is not good security posture. The lack of monitoring and alerting meant that it was easy for Harry and Dumbledore to systematically hunt all the Horcruxes down and destroy them without Voldemort noticing, even though he had alerts going on for all kinds of other important things.

But overall, the Horcrux plan worked really well. It took a team of dedicated opponents a lot of time and effort to find them and destroy them all.

Ultimately, **the time that Voldemort spent drawing up a Disaster Recovery plan, creating backups, and segretating storage meant that he managed to keep the bare minimum running and alive for fourteen years**, after the massive outage he experienced when he tried to kill baby Harry, and eventually he was able to fully restore from backup at the end of *the Goblet of Fire*.

Assessment

Voldemort's Security Method: 7.5/10

Good use of disaster recovery, redundancy planning, and infrastructure segregation. Monitoring and alerting need work, and personal details should not be used as keys to secure important information.

The Dark Wizard's Mindset

Why is Voldemort more thorough at planning defences than Dumbledore?

In the end, Harry Potter and the good guys end up winning against him, but the cost is huge and the losses are great. Part of why this happens is because Voldemort and his followers are driven by extremely different motivations, and have different standards of right and wrong than most other wizards do.

Dumbledore generally just wants to eat sherbet lemons and hang out with his pet phoenix and help young witches and wizards become their best selves.

Voldemort wants to start a race war and literally take over the world. He's got goals most wizards quite reasonably don't, so he's going to consider use cases and methods that most wizards wouldn't in order to get that stuff done. In attacking things, Voldemort tries all kinds of weird, out of the box, often unethical methods for achieving his ends.

If you're a wizard who generally wants to do good and make other people happy, getting into the frame of mind of someone who once tried to murder a baby in cold blood is going to take a bit of work, but **it is worth that mental exercise if you want to plan defences against the kinds of strange and disturbing things an opponent might actually try to do to get the information you have**, and if you want to understand the things that motivate them.

The other reason that Voldemort's defences are more thorough than Dumbledore's is because Voldemort is a paranoid person who trusts nobody and has no friends. He knows most people are against his plans, and that the only person who is going to help him out is himself, which is why he starts thinking about making Horcruxes while he is still at school.

Conversely, Dumbledore leans more towards creating groups of confidantes, building teams, and trusting others with important tasks. There's nothing wrong with wanting to do any of these things at all... the best places I have ever worked have been in strong teams that trust each other to do the right thing and give each other mutual support.

But for your app's outward facing defences, you will develop very strong ones if you **think that everybody is out to get you**. You'll plan even better defences if you can think of the types of groups that might be out to get you and what might motivate them to try to attack you. You'll do even, *even* better if you **assign likelihoods to each of these risks and prioritise defending against them accordingly**.

Failovers, backups, and disaster recovery plans are all the products of worst-case thinking, and we always hope we won't need them. But when we do, we're glad to have them, and we're glad we took the time to think about what to do in a horrible situation.

Web App Security

App security is a tricky space, and those of you tasked with defending your apps have an extra hard job trying to anticipate the work of attackers.

Here are the top six wizarding security tips for this introductory class:

Threat Modelling

Threat modelling is really handy for anticipating what an attacker might try to do to you, and countering against it.

Think about who is likely to attack you, what they're likely to try doing, and how you might stop them. Then build those things, and refine them as you learn more.

Defence in Depth

Well-planned defence in depth helps your app stay secure. Remember to make sure that you layer *different* kinds of defences over one another, that all of your defences are independently robust, and that they build on each others' strengths. Especially make sure that you review your defences regularly and keep them up to date.

Disaster Recovery

Have a disaster recovery plan. Things are going to go bad, and it helps to have thought about what you're going to do in those cases, and to have practiced for it. In security, thinking about the worst case scenario is our strength. Make sure you use it to your advantage, and plan for the absolute worst.

Mitigating Common Vulnerabilities

Stay on top of the very basic things that can let dark wizards past your defences. Patch your

software and your dependencies. Test for the OWASP Top 10✣ on a regular basis. And remember that you need to keep doing this often, because things can get out of date quickly, and dark wizards never rest.

Monitoring and Alerting

The wizard's downfall is never that they haven't planned defences, and almost always because they don't implement monitoring or alerting on these defences. The Philosopher's Stone, the Goblet of Fire, and Voldemort's Horcruxes all suffered from a lack of these two things.

Having good defences is great. Making sure you know when someone is trying to breach them is even better. It gives you an opportunity to respond, and more data to reconstruct the attack afterwards, so you can learn from the experience.

A Dark Wizard's Mindset

Finally, never forget:

It is a pentester's job to think as dark wizards do.

Notes

† OPSEC - Operations/Operational Security https://en.wikipedia.org/wiki/Operations_security

✣ OWASP Top 10: https://www.owasp.org/index.php/Category:OWASP_*Top*Ten_Project