



Recent outages at KnowledgeOwl

Last Modified on 01/11/2024 1:40 pm EST

We've received some questions about the number of outage incidents we've had recently. I wanted to take a break from our normal release notes to talk more about this.

KnowledgeOwl has historically been extremely stable. We've had very few outages, and until February, we'd never had a month where we violated our up-time SLAs and had to issue credits to customers. You can see an overview of the major incidents we've had recently on our status page [here](#).

If you page through, it's clear that the last two months have been an outlier for our uptime; we've had more major incidents in this time than for several years prior.

So these last six weeks have been very out of character for what you've come to expect from us, and we've anticipated that y'all would have questions. Single, isolated incidents can be brushed aside, but multiple significant incidents can feel like it's a larger pattern that you should be concerned about.

Saying "we're sorry" feels so inadequate here, but it is the truth: we are sorry for these issues. Outages and slowness are a terrible experience for you, your authors, and your readers, and they undermine people's trust and confidence in the knowledge you're sharing. They are the exact opposite of what we want your experience using KnowledgeOwl and interacting with our team to be.

First, let's walk through the 4 incidents we've had (though only 3 technically impacted our uptime). This will summarize our incident postmortems and provide some updates on our progress toward the identified next steps for each. Then we can talk a bit more about what all that means for us and, possibly, you.

February 7th

Underlying cause: System automation failure; alert notification failure; lack of access for available staff

Description: Outage. A piece of our traffic-management system went down and did not automatically restart, so it required a manual restart. However, our internal alerting system didn't surface this quickly and we didn't have any staff online who had appropriate access to complete the manual restart. The app was brought back online within 15 minutes of having appropriate staff online.

Identified next steps:

1. Resolve issues with internal alerting system to alert more members of our staff: in progress; first round of changes is complete but we're examining larger/more robust changes
2. Document the systems involved and manual restart processes: in progress, early stages

3. Grant additional access & train more staff in troubleshooting these issues: happening in tandem with the documentation

February 16th

Underlying cause: Inadequate testing procedures for changes to tables of contents/QA failure led to bad release

Description: Slow performance, ultimately resulting in a full outage. This incident was human error on our part. We released some updates to knowledge base tables of contents that had issues, and those issues caused larger performance problems. Once we rolled back the release, the app stabilized.

Identified next steps:

1. Update testing procedures for table of contents changes: complete
2. This combined with the previous February outage violated our Enterprise uptime SLA, so in accordance with that SLA we're issuing credits for the downtime: in progress; credits are going out now—we waited for month's end to ensure we had the correct amounts
3. Improvements to our load-testing processes: in progress; phase one is complete and is currently being used for any releases with possible load issues; we're continuing to evaluate if further changes are needed

March 14th

Underlying cause: Inefficient processes being run on very large knowledge bases.

Description: Slowness, not full outage. This incident didn't result in downtime, but it did cause degraded performance. We believe this was a combination of factors.

Identified next steps:

1. Rework two different costly processes to run more efficiently: in progress, hopefully coming to knowledge bases near you within the next couple weeks

March 24th

Underlying cause: Suspected DDoS attack on www.knowledgeowl.com, which shared infrastructure and resources with app.knowledgeowl.com and your knowledge bases.

Description: Outage, suspected DDoS attack. [Technically several shorter outages, all with the same root cause.]

Identified next steps:

1. Remove www from some elements of shared infrastructure: complete
2. Implement stronger firewall and proxy rules on www to be more consistent with the protections we have for

the app and KBs: phase one complete; still evaluating to determine if further changes are needed

3. Incorporate knowledge gained from this potential attack into incident monitoring and response processes: in progress
4. Examine larger architectural improvements to identify bigger-picture changes: not yet begun; we're still waiting for the dust to settle on the other changes here

To summarize

As you may have noticed from the summary above, we are "in progress" on the majority of the identified next steps. In progress means lots of things:

- It means that we have shifted a lot of our internal resources and priorities to address as many of these as we can.
- It means that we're not cutting corners or doing something slapdash just to be able to say it's complete.
- It means that in many cases, we've rolled out an initial set of changes but we are continuing to monitor those changes to be sure there isn't still room for further improvement.

We don't want to call most of these complete until we've had several weeks—if not months—to continue monitoring and verifying that all of the solutions we've put into place have fully remediated things.

The elephant in the room

If you look through the root causes carefully, you'll notice something: none of these incidents was caused by the same thing.

On the one hand, this is good news: it means that we don't have a consistent, persistent problem that is going unaddressed.

On the other hand, it can make it seem like things here at KnowledgeOwl are a bit out of control.

Each of these incidents alone would have been not great, but cumulatively they do look worrisome. And in that sense, I'm not sure there's anything I can say that would fully reassure you that all is well.

Truthfully, I think this series of incidents is a perverse form of bad luck. Having a bad release happens to all software companies. Getting a suspiciously high volume of traffic to a public website also happens, as do automated alert failures. The one incident here that we believe was due to a flaw in the software itself didn't even cause a full outage.

Having all four of those things in a six-week period? It's like we decided to save up 3 years' worth of incidents and have them all at once.

The thing is...I can't promise that we won't have another incident this month, or this year, or the next five years. No matter how many systems and fail-safes we put into place, the nature of SaaS is that there's always room for something to go wrong, or for a malicious agent to make every effort to make something go wrong. We can reduce the risks, but not eliminate them completely.

What I can promise you is that we are doing everything we can to avoid further incidents.

We take these outages very seriously. We know how integral and essential knowledge bases are to how you work or what your customers' experience of your product/service is like. That is a huge amount of trust you place in us, every day.

We are doing everything we can to make sure you keep trusting us to solve this problem. And we're trying to do it intentionally, and the right way, with an eye toward future maintainability and stability. That can mean that it takes more time, but we feel that is time well spent.

We're continuously working to improve our infrastructure, our software, and our team. We understand if these last few weeks have undermined your confidence in us, but we're not going anywhere. We'll still be here next month, next year, and years from now, still trying to make KnowledgeOwl the best it can be. We hope you'll stick with us to see what that looks like.

~Kate Mueller, Chief Product Owl & Resident Cheesemonger, on behalf of the entire KnowledgeOwl team

p.s. And, as always, if you have additional questions on any of this, we encourage you to [contact us](#) for more details.
