rewind

Backups and Disaster Recovery for SaaS

Intro: on-prem vs cloud

There's an old mainstay in IT that there are two types of people: those who have backups, and those who are about to. The concept of a backup and recovery system for your business-critical data is nothing new, but the transition from on-prem data centers to cloud-based systems and SaaS apps requires a new approach to backing up data.

If you're storing business-critical data in a SaaS or cloud-based instance (over 55% of business data is stored in the cloud), then your disaster recovery planning can no longer ignore the security risks and realities inherent with cloud-based software.

As a SaaS-enabled company that provides backups of many of those SaaS products and has backed up the SaaS data of over 100,000 similar organizations, Rewind's data security experts explore what disaster recovery planning means to SaaS and SaaS-enabled organizations.

James Ciesielski, CTO

James is the co-founder and CTO of Rewind, the leading data backup and recovery provider for cloud and SaaS data. After completing a Bachelor of Math, Computer Science/Software Engineering at the University of Waterloo, James has over 20 years of experience building highly scalable software and services in the fields of telecommunications, media, and financial technology in both enterprise and start-up environments. An experienced technical leader, James has successfully overseen the development and launch of a variety of software products, including Rewind's inaugural backup-as-a-service (BaaS) app, Rewind Backups for Shopify. In 2019, James was honoured as a member of the Ottawa 40 Under 40. When he isn't in front of his computer, James can typically be found running after his kids, cooking with his wife, and volunteering to be in net for every pick-up hockey game he can find.

Michelle Crane, PhD, Software Development Manager

Michelle Crane is an experienced engineering manager who enjoys working with development teams focused on designing, creating, and delivering complex software products that solve real-world customer problems. Her personal development background includes Java development, as well as testing, release process, and continuous delivery, all with a heavy emphasis on automation. Her first career was as an officer in the Canadian Air Force, where her speciality was Logistics, specifically fourth-line air/sea movement control. She holds a PhD in Computer Science from Queen's University with a focus on formalizing the UML Action Language. Michelle's other interests include reading, puzzles, and a new-found interest in Star Wars lego kits.

Dave North, VP Cloud Operations

Dave North has been a versatile member of the Ottawa technology sector for more than 25 years. Dave is currently working at Rewind, leading the technical operations group. Prior to Rewind, Dave was a long time member of Signiant, holding many roles in the organization including sales engineer, pro services, technical support manager, product owner, and devops director. A proven leader and innovator, Dave holds 5 US patents and helped drive Signiant's move to a cloud SaaS business model with the award-winning Media Shuttle project. Prior to Signiant, Dave held several roles at Nortel, Bay Networks, and ISOTRO Network Management working on the NetID product suite. Dave is fanatical about cloud computing, automation, gadgets and Formula 1 racing.

Megan Dean, Information Security and Risk Compliance Manager

Megan Dean is an experienced Information Security Professional with a focus on governance, risk, and compliance. She is a Certified Information Systems Security Professional (CISSP) and a Systems Security Certified Practitioner (SSCP). She currently serves as Rewind's Information Security and Risk Compliance Manager. In her spare time, Megan can usually be found watching documentaries, playing video games, or reading the latest cybersecurity news.

Back to basics: 3-2-1 in the cloud

While the method may be new, the same principle of the 3-2-1 backup rule still applies: three total copies of your data, on two separate mediums, with one stored off-site. Cloud backups help support this principle by allowing for efficient copying and storage of your business-critical data.

For example, automatic repository backups can also be automatically synced to "off-site" or separate cloud storage containers, achieving the "off-site" and "separate medium" criteria at once.

Cloud backups are also easily tested and configured with event alerts, such as a backup failure notification. Backup testing is an essential part of your DRP. Regular tests of backup and recovery procedures are often required for industry-related security compliances such as ISO 27001, SOC 2, HIPAA, and more.

While backup testing is key, it's only half the task. The real thing you must test to ensure is functioning properly is your data backup and recovery procedures. In other words, a backup of your data isn't helpful if you can't restore said data back into your systems. This is why a CSV or JSON-based backup file of many SaaS environments isn't useful in an emergency, as the data can't be directly and easily imported back into the necessary SaaS tool.

Business-critical data backup and recovery procedures include the following considerations:

- What precisely is your business-critical data? What data dependencies exist? For example, metadata often contains key information about development progress, bugs, commits, etc. To know if you have a backup and recovery system for all your critical data, you have to know precisely what your critical data consists of and where it's being stored, accessed, or archived.
- Once you have identified everything that needs to be backed up, you can begin to define and implement backup and recovery procedures. Tools that regularly schedule and perform backups (also known as BaaS, backup-as-a-service) are helpful for automating this process and reducing developer time needed to maintain it.
- Now, begin to build your recovery procedures. A tabletop exercise can be useful for fully imagining and mapping out your data recovery process. Pick a disaster scenario, and play it out (similar to a tabletop roleplaying game). If a hurricane/cyberattack/alien invasion breached a specific data center or warehouse, what would you do? Begin listing the steps and exploring all possible setbacks you could encounter.

An Introduction to cloud backups (and how to test them)

Just like any other backup, backups of cloud or SaaS data also need to be tested.

Finding out your backups are incomplete or un-restorable during a five-alarm fire is not fun for anyone. Scheduling regular backup testing helps ensure your backups are always functioning properly and can help with audit compliance.

> Rewind's VP of Cloud Operations, Dave North, reminds us: "Anyone who isn't testing their backups on some cadence is on borrowed time. I look at it like this: You can deal with the stress of a backup not restoring during a normal day and fix whatever the problem is. Or, you can deal with the stress of a backup not restoring during a five-alarm fire at 2 am."

Regularly testing backups can have other operational efficiency benefits as well. Just like the table-top exercise, every failure is an opportunity to learn.

> Michelle Crane, PhD, Software Development Manager at Rewind, remembers an experience from earlier in her career: "we were also SOC 2 compliant, and one of our controls was around disaster recovery and backups. Every six months, we were supposed to test backups and restore. Basically, are our backups automated and working well? Yes, those were always fine. Restores, on the other hand.... I did at least two of these restore tasks, each focusing on a slightly different aspect. In no situation was I able to actually perform a restore properly; I did, however, find a lot of areas for improvement. Although we weren't successful at the automated restore test, the knowledge we gained there was actually quite valuable during one of the AWS outages that happened late last year."

Getting ahead of any potential issues with regular backup testing helps future-proof your organization against unexpected issues such as outages, cyberattacks, or simply human error.

How often should I test my backups?

Like many security best practices, how often backup testing should be performed is different for every organization. If you are only updating the data being backed up monthly, weekly backups are probably irrelevant. However, if your organization makes frequent changes to the data (a much more likely scenario), daily backup snapshots are recommended. It's also a best practice to run a full backup before implementing any risky or new procedures, code, apps, or anything else that could have unintended consequences for your data and data structure.

How to make the case for backups: balancing cost and benefits

Part of being an engineering leader is communicating the needs of the engineering team to non-technical stakeholders.

Take backups, for example. Backups cost money to store and take time to create. Theoretically, if everyone does their job perfectly, they shouldn't be necessary, but experienced engineering leaders know to prepare for failure.

Having proper backups is an investment, but the reduced business risk often provides a significant return. Let's look at a few situations where backups can make a difference. Hopefully, you can use these examples to justify putting a rock-solid backup and restore plan in place at your organization.

Account Compromises

In July of 2019, Ubuntu Security reported that the credentials for a company-owned GitHub account were compromised. These compromised credentials were used to create repositories, issues, and more.

> "We can confirm that on 2019-07-06 there was a Canonical-owned account on GitHub whose credentials were compromised and used to create repositories and issues, among other activities. Canonical has removed the compromised account from the Canonical organization in GitHub and is still investigating the extent of the breach, but there is no indication at this point that any source code or PII was affected," says Ubuntu Security.

In this case, it seems that critical infrastructure was decoupled from GitHub, and the breach wasn't allowed to spread. However, Canonical (the publisher of Ubuntu) had to restore various repositories and issue trackers to their previous state. When rolling back from an account compromise like this, backups are infinitely helpful, as they give you a previously-good state to compare with the current state.

Additionally, attackers often leave backdoors in compromised codebases. This can allow them to gain greater access once the initial discovery and remediation process is completed. If you only have your infected codebase, it can be challenging to uncover all the infected files or possible vectors for future attacks.

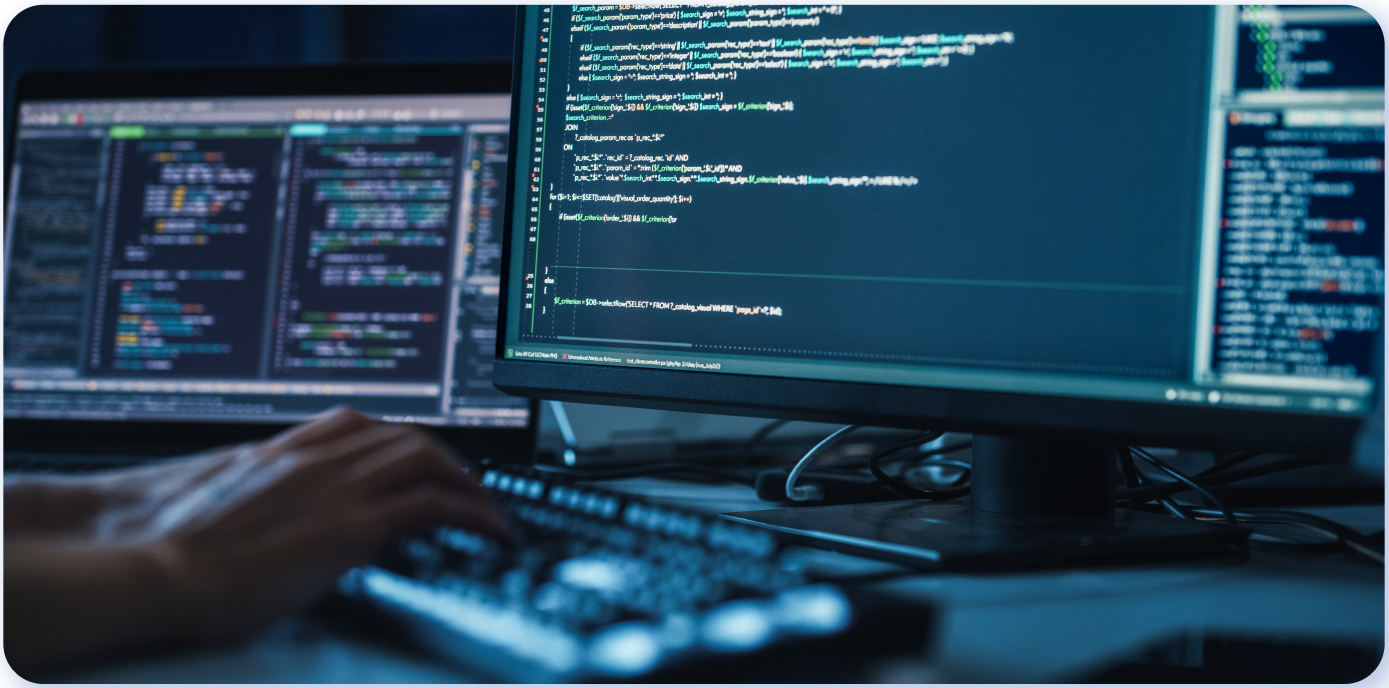
Ransomware Attacks

Ransomware is the act of taking control of a codebase or set of infrastructure and encrypting it so that only the attacker can unlock it. In exchange for returning access to the victim, the attacker demands a ransom be paid, usually in cryptocurrency. If the ransom is not paid in a certain timeframe, the attackers threaten to delete the files, rendering them irrecoverable.

This is exactly what happened in May of 2019 when ZDNet reported that, "Hundreds of developers have had Git source code repositories wiped and replaced with a ransom demand." The hackers had modified the git histories to the point where the repositories were unusable, and they demanded payment within ten days to reverse the changes.

If a compromised repository is part of your organization's core codebase, an attack like this will cause a severe disruption in operations. Developers will be unable to commit code, creating a complete stoppage in new feature development. Bug fixes and even support tickets could be affected. However, if you were hit with an attack like this and had a backup of your data, you could restore from this backup and continue working on day-to-day tasks while the issue was resolved.

Ransomware attacks are most effective when the victim has no other option for accessing their data besides paying the ransom. Having a complete backup—even one that is a few hours old—ensures that you're not helpless in the face of a ransomware attack.



Service Downtime

Many companies have moved parts of their codebases to 3rd party providers. The reasons for this are many, but if code storage is not your core business, you can save time and money by relying on an outside service.

However, depending on a 3rd party always carries some risk. No service can deliver 100% uptime, but when your entire business (or codebase) depends on a SaaS platform's availability, you might want to mitigate that risk by having your own backups.

For example, in June 2020, GitHub had a major outage that lasted for hours before stability returned. If you relied on GitHub to store your code, this meant much of the development work for the day was on hold until they resolved the issue. These outages may impact developer productivity and project timelines if they occur during a crucial launch window.

Like the ransomware scenario described above, the best way to mitigate service downtime is to have a plan in place and a backup of any repositories and associated metadata. While you can create these backups on your own, using a service like Rewind will make it significantly easier. With a proper backup and restore plan in place, what could potentially be a work-stopping outage can instead be reduced to just an alert and switchover.

Reducing Platform Dependence

One of the downsides to using a 3rd party provider for something that's not your core business is that your business depends on that provider. Those providers are businesses in their own right and may face financial or regulatory pressures that limit your ability to use their platform.

For example, in the summer of 2019, GitHub was forced to comply with US export law and had to prevent users in Iran, Syria, Crimea, and other sanctioned nations from accessing their service. Anyone in the affected countries was cut off and forced to find a different repository host.

This is another case where backups would have been invaluable. With a backup, your repository can be restored and pushed to another provider, or you can maintain a self-hosted version of your repository. Events like these are rare, but having a backup is a small price to pay for maintaining access to your code.

Offloading the technical cost: outsourcing backups

So, you know you need a backup solution you can test and rely upon in a disaster recovery scenario. But it is better to DIY or outsource your backups? Let's dive into the eternal question of "build vs. buy."

Pros of a Bespoke Solution

As with any first-party solution, most benefits boil down to increased control. Companies and internal teams can design their backup solutions as they see fit, implementing their own controls, preferred technology stacks, and coding implementations.

Accordingly, developers can marry backend and frontend designs however they wish. The opportunity for customization is nearly limitless. It's relatively easy to determine what to back up, which mechanisms that can happen through, and how internal resources may tie into resulting backups.

Additionally, you can back up as frequently as you wish with a bespoke solution. While third-party tools may only permit backups—often called snapshots—at certain intervals, you can calibrate internal solutions to back up daily. This is even possible multiple times daily. If that ambitious backup schedule is essential, then an in-house approach may be ideal.

Finally, bespoke backups can point to whichever storage solution you prefer. This is especially useful when weighing multiple cloud-based storage options against one another without introducing vendor lock-in or dependency into the equation. Companies can also point backups toward on-premises drives for extra peace of mind.

We've previously covered how to write a bespoke backup script for GitHub and Jira.

Cons of a Bespoke Solution

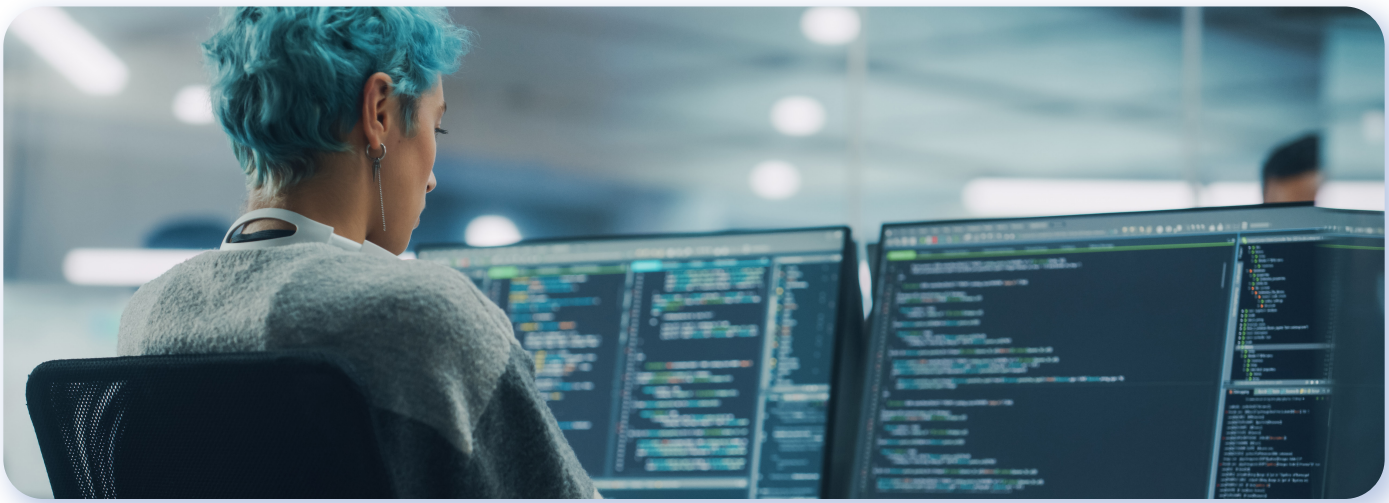
A bespoke backup procedure may sound appealing, yet no system is perfect. Making your own backup solution means incurring a lot of expense to get it right—from inception to development to maintenance. Will you have enough team members to oversee this solution throughout its lifetime effectively?

These are uncertainties worth addressing before making the decision to go with a bespoke backup solution. Open-source resources can indeed dampen initial costs, but long-term expenses will be noteworthy.

Accordingly, there's also a fair amount of effort required to devise such a solution. Many man-hours are needed to make something functional and reliable. Essentially, you're creating a backend to serve multiple purposes by leveraging APIs. Development time increases if you wrap a GUI around everything. There are very few out-of-the-box functionalities to be had. These APIs must be kept current, and validation testing takes a lot of time to really nail. That's doubly true for companies without optimized CI/CD and DevOps pipelines.

It then becomes much more essential to amass expertise within your organization. Experienced professionals are needed to create and understand how a bespoke backup solution works. What happens if these keepers of the keys leave the organization later on? Sound documentation becomes another priority. Your team is essentially becoming its own software vendor.

Finally, it's worth questioning whether you can fully trust your internal solution even in the best-case scenario. How will unforeseen development mistakes, testing errors, and backup failures impact your data? After all, data is critical to a company's workflow —impacting reputation, efficiency, and security. IP loss can even threaten your company's existence if the fallout is grand enough. Can third-party solutions do a better job?



Pros of Purchasing a Backup Solution

When we assess some core negatives with the bespoke route, the inverse soon becomes true for external solutions.

Your development team saves time that would've been allocated to backup creation and can harness it for other projects. Building and maintenance are already handled by the vendor. Similarly, ongoing updates and security fixes are pushed without involvement from your team.

While some initial setup is required, it's possible to lock in your settings and forget. Blissfully watch as your tool handles recurring backups automatically. Additionally, no special code or experience is needed to set up these turnkey programs. Third-party backup solutions, like Rewind, come with batteries included:

- Automated daily backup snapshots
- Cloud sync features
- Item-level data restores, including metadata
- Data can be restored in minutes via the self-serve portal, with no specialized knowledge or coding required.
- Choice of data storage location
- Centralized audit logs

> "If you build a backup solution yourself, then you take on all of the quality, reliability, financial, and security risks that come with it. If someone else has addressed those risks, why would you want to reinvent the wheel and take time away from delivering value to your customers?" explains James Ciesielski, CTO and co-founder of Rewind.

Why are audit logs necessary?

Generally, logs are readily available within solutions like these, and the tool's developers work to ensure compliance for users across multiple industries. The backup tool acts as the single source of truth, within which a centralized, current backup is preserved at all times. You won't have to go hunting for copies of files, attachments, and codes if calamity strikes. These live in a remotely accessible cloud location. Authorized employees can retrieve data as needed, and DevOps teams can perform restoration work from anywhere.

Costs of third-party solutions

Note that costs may be expensive with these solutions. That may be a con as well, yet the pricing structures offered by many backup companies are predictable and consistent. You'll know what you're getting, what it'll cost, and how to budget accordingly. Additionally, most BaaS services include data storage costs, which can quickly escalate depending on the amount of data your organization creates.

Cons of Purchasing a Backup Solution

A clear disadvantage of going to a third-party provider centers on data storage.

Backup platforms include support for data storage solutions their developers choose to support—instead of letting you choose. This requires vetting before buying. It may also introduce storage fragmentation within your organization, should you be married to one specific tool. That applies to geographical data locations (North America vs. Europe vs. APAC, for example) and storage vendors (S3, Google Drive, and others).

Finally, you have less control over backup frequencies. Some tools can only back up at longer intervals. However, frequent changes to your projects in the wake of agile development may require more frequent backups. For some organizations, even daily backups may be inadequate.

Summary

- Building a custom backup solution brings inherent risks due to maintenance, updating, and employee turnover. Yet, bespoke backup solutions can offer more customization for organizations with unique backup needs.
- Buying a BaaS (backup-as-a-service) product unloads the work of building, testing, and maintaining your backup systems. Your devs can focus on their work rather than spending cycles running and testing backup scripts. Additionally, third-party backups are often required for compliance purposes.

> "Take the Lego example," Ciesielski continues. "You *could* 3D print your own Lego, but you are farther ahead if you simply *use* Lego! You eliminate all the problems that come with trying to manufacture your own bricks and can focus 100% of your energy on building."

DRP and Backup Testing for SaaS

And now, as promised, a handy checklist of things to consider when integrating backup testing into your compliance and disaster recovery planning:

- ☐ **Consider all mission-critical data.** Information stored in SaaS systems is often overlooked during data scoping, yet these apps and systems typically contain valuable information your team relies upon (and if they don't, perhaps it's also time to audit your SaaS spend).

- ☐ **Understand data dependencies.** An item of data is not just a single item: it's likely linked to other filetypes that help your SaaS system understand how to sort and organize the data. For example, code repositories contain much more than code. Vital metadata like comments, issues, and pull requests must be considered in your backup testing plan.

- ☐ **Create and document your backup testing plan.** Your backup testing plan is the complete procedure your team will take if they need to perform a restore. Documenting this process ensures that everyone will know precisely what to do (and when) during backup testing and if you need to restore. Plus, documented backup testing plans are often required or useful during audit compliance.

- ☐ **Make backup testing a regularly scheduled event.** Depending on your organization's size and regulatory requirements, backup testing may be required at a certain frequency, such as bi-annually. Regardless of how often you choose to test your backups, backup testing should be scheduled regularly and in advance. This helps make time for this essential task, even during a busy period.

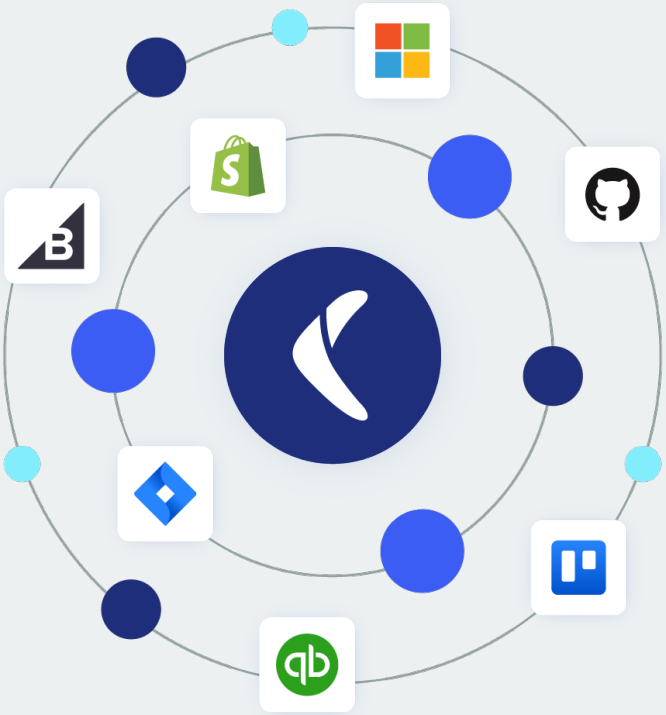
- ☐ **Backup testing is recovery testing.** A backup of your data that you can't restore is like having the spare key to your car locked in your trunk: it isn't very useful in an emergency. When testing your backups, ensure your team knows the exact steps to restore data back into your instance or workflow. The backup test isn't complete until all data has been restored back to its ideal state. Consider the length of time it will take to restore your organization's data and systems in case of an emergency. The backup and restore test is complete when your team has access to their data in its original state and can continue work as normal.

- ☐ **Location, location, location.** Be aware of any geographic requirements related to the storage of data. Certain regions, such as the EU, have specific rules related to data storage, i.e., GDPR. Look for data backup providers that can accommodate your geographic restrictions and ensure your backup recovery procedures comply with all relevant laws and regulations.

Conclusion

Regardless of whether you purchase or build, backups and recovery ensure your data is safe, secure, and available when your team needs it.

> "Backups are no different than an insurance policy. You hope you never need to use them—but when you do, you want the right insurance agent/company who will stand behind their backup tech," says Dave North, VP of Cloud Operations, Rewind.



Protect your data with Rewind

Get a quote or schedule a demo at Rewind.com. Reach out to sales@rewind.com for more information.

rewind