



CloverDX

The Buyers' Guide to Data Integration Software

Contents

Roadmap	3
Introduction	4
Before you start shopping	5
AI capabilities to ask about	6
Choosing a provider	11
Evaluating vendor software	12
Five criteria to test any platform	14
After you buy: Implementation	15
Conclusion	16



The Buyers' Guide to Data Integration Software

Before you speak to anyone establish your needs



What are you trying to do?

- Data ingest
- Data for BI & Analytics
- API integration
- Data migration
- Data error management
- Data transformation & movement



General considerations:

- Budget?
- Number of data sources?
- Where is your data mostly located?
- Where does your data need to be?
- Future business needs?
- Data quality?



Technical considerations:

- Open source vs commercial software?
- Cloud vs on-premise?
- AI capabilities vs where they run?
- Visual interface vs developer tool?
- Admin & hardware resources required?

Creating your vendor shortlist



How to identify appropriate vendors:

- Ask lots of questions
- What's included/not included in the price?
- Have you decided on an appropriate licensing model? What support is on offer?
- Will the vendor help with implementation and set-up?



Dos & don'ts:

- Do consider a proof of concept
- Do consider more than one vendor
- Don't take solution promises on face value
- Do think about current and future needs
- Don't assume all solutions/vendors are alike



Demos & evaluations:

- Do both for each shortlisted vendor
- Ensure any evaluation is a full version of the software
- Don't accept canned demos
- Test software against your specific use case
- Test the vendor, as well as the software

Who will be using the software?



User profile:

- Will the software be used by technical or business teams, or both?
- Do you want an easy to use visual tool or something for developers to code in, or both?
- Is it for internal users, external consultants or a combination of both?
- Do you have experienced staff or is training required?
- Have you considered cultural factors, change management and new ways of working?

Implementation



Challenges of working with a new technical solution:

- Expect trouble
- Lean on the vendor to mitigate the risk of your project going wrong – ensure help when you need it!
- A good vendor will help design your data architecture and provide consultancy services if required
- Make sure your vendor of choice provides adequate support and documentation
- Set clear goals and expectations, establishing KPIs so you can measure success

Introduction

If your business is like most others these days, you're probably striving to wring value out of ever-expanding volumes of data. That means data of different kinds, different formats, and from many different sources. Maybe you can't access the information you need when you need it most. Maybe your data needs a lot of cleaning before it's usable. Or maybe you're tired of asking your developers – or costly outside consultants – to update data processes every time another problem arises.

Poor data quality is expensive. IBM estimates that US businesses lose \$3.1 trillion a year to the cost of poor data quality^[1], and that figure only grows as AI moves more of that data into decisions your business can't easily unwind.

Poor data integration can come with steep costs too. Companies in the US and UK alone lose \$140 billion each year due to disconnected data, according to a study by the tech research firm Vanson Bourne. And 47 percent of those surveyed believe that's hurting their ability to innovate and develop new products and services.

Data integration provides a way to consolidate data from multiple sources, gain new insights from different types of data, and bring your organization into the digital business era. And if you've decided against building such a solution in-house – or have concluded that the one you built no longer serves your needs – you're probably thinking about shopping around for a data integration software provider.

So how do you decide which data integration software is right for you?

This guide is designed to walk you through that decision-making process, step by step. Whatever your role – whether you're an IT team leader who'll be working with data integration demands on a daily basis, or a CEO looking for a solution that aligns with your company's current business needs and long-term digital strategies – you'll learn what to consider and what questions to ask before committing to a data integration solution.

You'll also find a roadmap with clear dos and don'ts to help you get on the fast track to finding the right vendor, plus a dedicated section on what to ask about AI capabilities.

[1] IBM, cited in CloverDX's white paper "Architecting Systems for Effective Control of Bad Data," <https://www.cloverdx.com/gc/lp/white-paper/architecting-systems-effective-control-bad-data>

Before you start shopping

Key questions to ask

Finding the right data integration provider first requires you to understand your needs. Start by asking, “Why does my business need data integration? Is it for a one-off project, e.g., a migration? Or is it for ongoing data integration demands, such as automating data jobs?” To answer these questions, you need to be able to envision what a successful data integration deployment looks like. Does it involve moving data job responsibilities to non-developers? Are you looking for faster turnaround in changing data flows or onboarding new data sources or customers?

“Data quality, or the lack of it, is the top complaint among business analysts, migration teams and other data users.”

Is this an isolated project, or is your company completely restructuring to become more data savvy? It’s also important to consider how you need your solution to work operationally. Who’s going to maintain it? And what information are they going to need access to within the system in order to keep everything running as it should?

You also need to think about the number, types and nature of data sources you’re looking to connect. Are you working with databases? APIs? Cloud-based data? Fast data or slow? The nature of the data itself also matters, so consider volume, quality, complexity and structure.

All of this means asking questions like, “How many data sources do I want this solution to cover? What do they look like today? Where are they located? And how might my data architecture change in the future?” The data integration software you choose should meet your current needs but also be able to accommodate future requirements. (This also means examining your company’s long-term plans to develop people’s data skills.)

Location of your data in particular matters, as your ideal solution may vary depending on whether you’re working with a single, giant database or with multiple data sources around the world that need to be consolidated. Your options will also be different if you’re moving on-premises data to the cloud, rather than keeping data in-house to maintain control and security, and to avoid vendor lock-in.

Another thing to consider is data quality: “How good is my current data? What’s needed to clean, standardize and optimize it? And what insights could my company potentially draw from well-integrated data?”

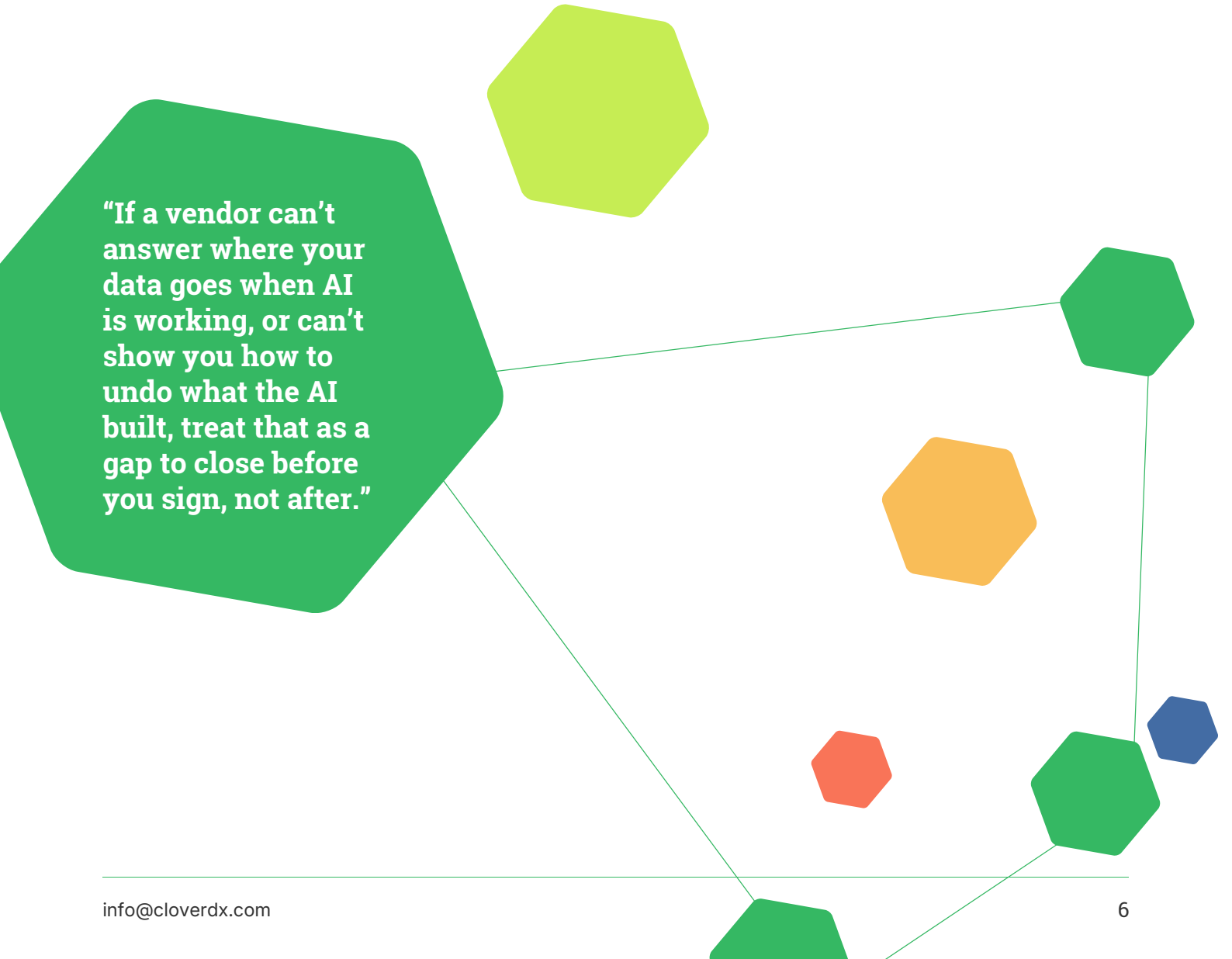
Data quality – or, rather, the lack of it – is the top complaint among business analysts, migration teams and other data users. Cleaning data, and consolidating information once kept in many different siloes, isn’t easy. But when it’s done well, the benefits and competitive advantages delivered can be huge. In fact, those results are a big reason to make the case for data integration.

“Data doesn’t exist in a vacuum: it’s put to work by people to help solve specific business problems.”

AI capabilities to ask about

AI now touches nearly every stage of data work, so it deserves its own line of questions on your evaluation checklist. A few worth asking any vendor:

- **Where does the AI run, and whose infrastructure does your data touch while it works?** Some platforms send your data to a vendor-hosted AI service. Others let you bring your own provider or model (BYOK/BYOM), so your data goes only where you choose to send it.
- **Can developers use AI to build pipelines while keeping full visibility into what it produced?** You should be able to open, read, and change anything AI builds, the same way you'd review a colleague's work.
- **Is AI an addition to a mature platform, or is it the whole product?** A platform with years of production use behind it can add AI as an accelerator. A brand-new AI tool has no track record to lean on.
- **Can business users get AI assistance too, or is it developer-only?** If your organization wants to involve more people in data work, ask whether AI support extends beyond the engineering team.
- **What happens when the AI gets something wrong?** Ask how you'd spot it, roll it back, and keep the pipeline auditable, regardless of who or what built it.



"If a vendor can't answer where your data goes when AI is working, or can't show you how to undo what the AI built, treat that as a gap to close before you sign, not after."

The Buyers' Guide to Data Integration Software



Other issues to consider:

What is my budget?

How much am I willing to spend to find the right solution? Consider more than just license costs. So-called 'free', open-source solutions, for example, will likely require you to spend money on consultants or pay your own developers to work on issues. And some offerings' lack of direct support, or limited support via community forums, might work against you if you're bound by SLAs or demanding users who require quick turnaround of critical issues. Look for vendors whose price includes not only the software, but the levels of maintenance and support you require as well, including product upgrades or a la carte options. And ask for ROI data and other information that shows how a solution could help your bottom line.

Also ask how the vendor prices its software. A per-source or per-connector fee can climb quickly as your data grows. A usage-based, metered model means your bill moves every time your business does. A flat, capacity-based price gives you a number you can plan around.

How much self-service is my organization prepared to take on?

Do we have the resources to handle building and maintenance requirements in-house? For example, employees outside IT might not be able to handle raw data, but might want to dig deeper into data themselves if they're provided with curated data sources. How do you envision business users working with data? Self-service might be tempting, but watch for potential gaps between the raw data you have and the information your people need. Who is going to maintain your solution? Do they have the skills they need to keep your data flowing seamlessly?

Do we need an on-premises solution or cloud-based software?

What types of licensing or subscription models work best for my business and budget? How are fees calculated, and do they make sense if new use cases arise over time? Think about how licensing works, and ask how costs might change as you add more projects or users, or as your data volumes grow.

For many organizations, this question is really about control: who has access to your data, where does it sit, and what happens to your costs if a cloud vendor changes its pricing? If data residency or vendor lock-in are concerns, self-hosted or hybrid deployment keeps that decision in your hands.

What sort of security do I need?

Do I need simple user and group management, or a completely isolated multi-tenancy solution? Also consider your security needs for data storage and processing. Should records be encrypted in-transit? Or are the performance advantages gained by not encrypting more important?

How will this solution affect my organization's business processes, both now and into the future?

What potential impacts do I need to consider? With better-organized and faster-flowing data, companies can see rippling benefits such as shorter processing times and more efficient change management. You need to account for such developments in your mid- to long-term business plans.

"When getting ready to try out demos and evaluations, think not only about today's needs but tomorrow's."

The Buyers' Guide to Data Integration Software

Think about your users

Data doesn't exist in a vacuum: it's put to work by people to help solve specific business problems. That means any data integration software must make sense for the people who'll be using it. So look at the solution's features and tools, and consider their use in your organization. For example, if your IT and tech teams will be the primary users, look for software that's developer-friendly. On the other hand, if your business users will be involved in data preparation, you'll want a solution with a more user-friendly interface, yet likely to be constrained in depth of features.

Your ideal vendor will be different if your end users are people in marketing, sales, finance or legal departments who typically aren't comfortable with anything more than Word or Excel, rather than in-house developers who expect the ability to fine-tune code and create custom functions and modules. Data integration moves and prepares data for further use, so think about both users and, ideally, clear separations of their duties.

The technical ability of the people using the software will help determine, for instance, whether you'll need something simple to manage, with a visual interface only, or something that offers administrators and developers an open architecture, allowing them to customize the solution and integrate it into their own IT ecosystem.

You also need to consider not only who's going to be working with the front end of your data integration software, but also the people who are going to be maintaining the infrastructure. If your developers are going to be handing off jobs to a support team, that support team needs to understand and be able to work with the system architecture. It's important to involve these teams early in the process, so they can establish how best to integrate the new tool with existing systems and how they will manage ongoing maintenance.

Beyond these considerations, you also need to think about who's going to be evaluating any proposed solution, and who'll be helping to implement it. Some systems are easier to set up and use than others, and require less-intensive technical resources to run. Once you've thought about everything above, you'll be able to come up with a shortlist of solution providers to examine more closely. Let's look at this evaluation process next.

"Implementing any data integration initiative is guaranteed to cause a degree of pain."

Demos and evaluations

While product demos can provide a taste of a solution's capabilities, a standard vendor demo may not be enough to help you decide if the software is right for you. Look for vendors that will tailor a demo to your specific use case, maybe even using your own test data. Do everything to steer the demo from what the vendor wants to show you to what you really need to see.

Meanwhile, a full evaluation requires a deeper dive by the people who'll actually be using the software. They should also test the software to see how it meets their real-world task and performance needs. Both the demo and the evaluation are important in helping you assess an offering's technical capabilities, as well as the vendor's attitude and response times.

Don't automatically assume rich functionality behind slick user interfaces. Take time and try to actually implement a part of your solution, don't just 'play' with the tools. Ideally, an evaluation will approximate your real-life use cases as closely as possible. Your evaluation should also include testing any server-side processes, so you can see where any errors occur and whether notifications and error handling processes are sufficient for your requirements.

This will help you visualize how the software will address your organization's specific needs and solve its specific problems. It should also give you a sense of what your support and maintenance requirements might be, and how you'll handle troubleshooting. This will enable you to know what to look for in a vendor's services as well as software.

For example, does a vendor provide all of its services in-house? Or does it work with an ecosystem of partners for support? How large is that ecosystem and what quality can you expect from those partners? If you prefer a close, personal vendor relationship, you might want to weed out the candidates that rely on third-party partners for services and support.

Finally, when getting ready to try out demos and evaluations, think not only about today's needs but tomorrow's. Which features and capabilities might be important two to five years down the road? How are your own resources and capabilities likely to change? And what might your future business needs be?

Dos and Don'ts

Do be realistic about your organization's resources and capabilities.

Weed out vendors whose services would leave you overextended or unable to cope.

Don't be afraid to acknowledge when a solution isn't right for you.

Expect the vendor to address your suitability concerns, and watch closely to see how honest they're being with you. Good vendors should be clear about what their solutions can offer, as well as what caveats to expect.

Do recognize the importance of good data.

Data integration software can't magically solve the 'garbage in' part of GIGO. But a good solution^[1] can help you sort out data quality issues, provide you with error management capabilities and help you fix things when they go wrong.

Don't be afraid to make vendors work during the demo and evaluation stages.

Make sure they can answer all of your questions, provide you with proofs of concept and walk you clearly through every part of their solution.

Do ask for a prolonged trial period to test drive the platform thoroughly.

Don't commit to any decisions before you understand completely how a solution will affect your business and business processes. This might use up some of your resources now, but will pay off greatly over the long term.

Do ask to see AI features work on your own data.

A demo that only shows a pre-built AI workflow tells you little about how it performs on your problem.

[1] For more on this subject, see our related paper, 'Architecting Systems For Effective Control of Bad Data,' <https://www.cloverdx.com/gc/lp/white-paper/architecting-systems-effective-control-bad-data>

Choosing a provider

Evaluating vendors

Now that you've identified your key considerations and needs, you're ready to begin weighing your vendor options. Start by asking shortlisted candidates not only about their software, but about their business, service models, support and more.

Questions to ask the vendor include:

- What can you tell me about your company? How many customers do you have? Which markets are you most active in? How many businesses and/or use cases like mine have you worked with? Can you provide reference customers I can speak with?
- What is your onboarding process for new customers? How does it work? What are the typical timeframes?
- How much support will we realistically need? What might our SLAs need to look like? Do you provide hands-on training?
- How will your kick-starting or consultancy services affect my costs? Who does the work: your people or contractors/partners?
- What's covered under a commercial agreement? What isn't?
- Do you have online communities, knowledge bases, FAQs and other sources of support? If so, are they available for no extra cost? Are they in addition to your support services, or there instead of personal support?

You'll develop a better sense of the vendor's capabilities during the trial period, so weigh your impressions carefully. Does this company seem likely to value your business? Has it shown it will respond promptly when you need it? Does it seem like the kind of business you want to have a relationship with? Or are there nagging doubts the vendor might forget you once the order is signed?

"Choosing the right data integration vendor is like choosing a personal trainer"

Evaluating vendor software

Before anything else, make sure you're evaluating the full version of the vendor's software, not a light version with limited functionality. And don't rush – the vendor should give you enough time to make a full assessment.

Develop a shortlist of your top software choices by asking:

- How easy is the software to set up, administer and maintain? What are the hardware requirements?
- How scalable is your software? Can it keep up with my data integration needs if my volume of data grows 10-fold... or 100-fold?
- What about performance? What metrics do you use to measure software speeds and capabilities? Be realistic here: don't pursue performance for performance's sake alone, but have a clear purpose in mind. Know when you need real-time responses, and when processing times of a few hours might be just as effective.
- Can you help me design a data architecture that's fit for my business?
- Do you support all types of data sources? How effectively does your software work with each one?
- How do you test and verify the systems you deploy? How do you ensure data quality?
- What is your support like? Will you be there to help me when things go wrong?
- How do you determine fees? What is included and what is extra? What is the total cost of ownership (TCO), and how do you calculate that?

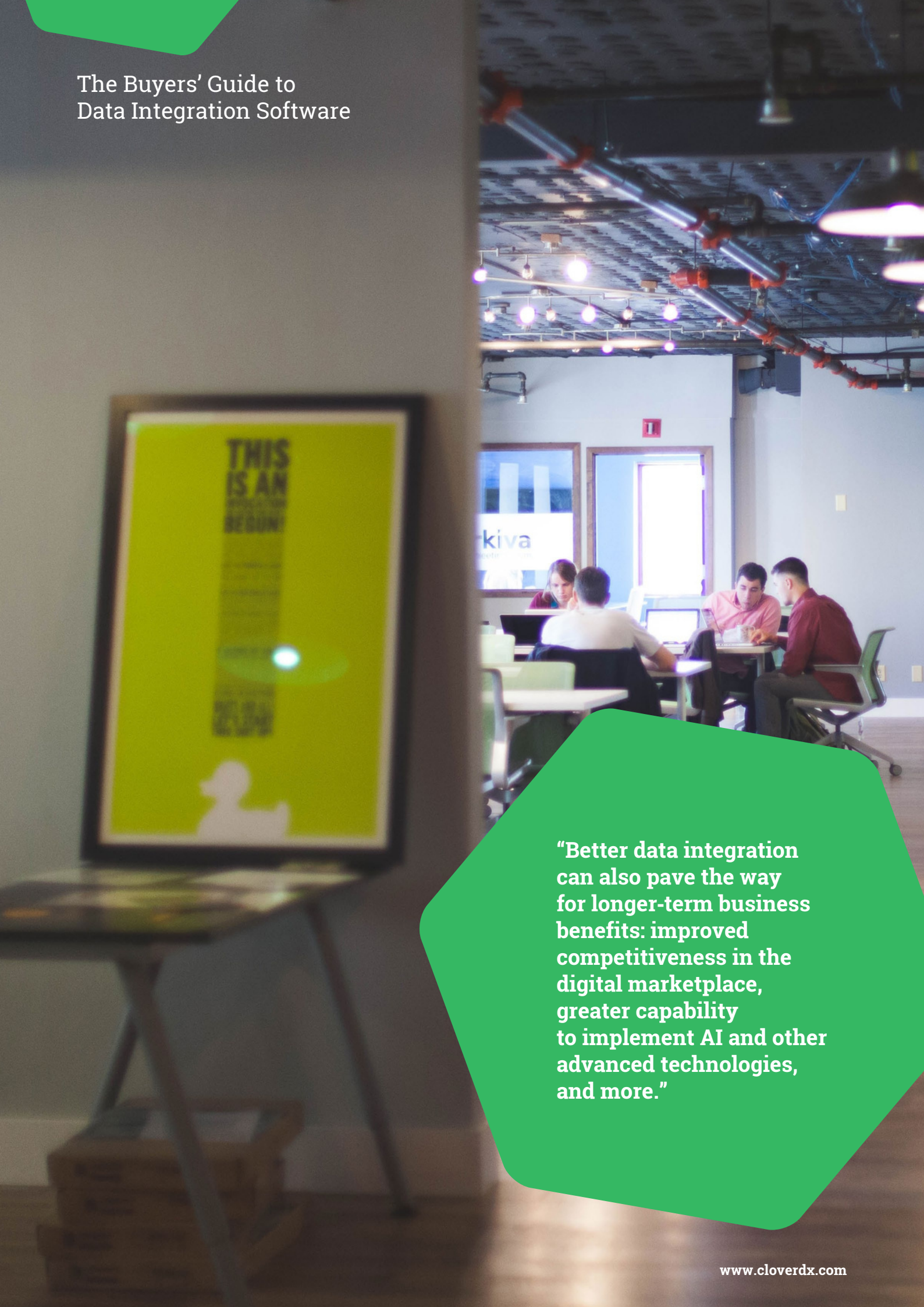
Licensing is a major consideration in any software decision. Ask what your options are, and which licensing model is best for your particular use case. Be sure to also ask about any setup costs or 'hidden' fees.

Before you commit to any data integration solution, it's also vital to take into account how your needs might change in the future. Can the vendor's software scale and support greater complexity? What future-proof tools and capabilities does it provide? And can you implement such improvements without having to pay prohibitively expensive fees, retrain your in-house developers or build new business processes from scratch? If your data processes scale up, will you have to change the way you maintain the system (e.g. by taking on new people with different skillsets to manage it)?

Different types of data integration solutions come with different pros and cons, so weigh those against your unique needs. For example, not all software can handle data in real or near-real time. And not every vendor supports a hybrid solution when you need some data on-premise and some in the cloud.

"Be sure to also ask about any setup costs or 'hidden' fees."

The Buyers' Guide to Data Integration Software



“Better data integration can also pave the way for longer-term business benefits: improved competitiveness in the digital marketplace, greater capability to implement AI and other advanced technologies, and more.”

Five criteria to test any platform

Whichever vendors you're comparing, it helps to group your evaluation into a small number of capabilities rather than a long, flat list of features. This is how we think about it at CloverDX, organized around five capabilities: Build, Run, Engage, Know, and Own. Use the framework to structure your own evaluation, whichever platform you end up choosing.

The data integration platform that puts you in control.

Build

Build visually, in code, or with AI — and move freely between all three.

The payoff: the speed of AI-assisted development, without code that nobody understands. The visual form also lets non-engineers follow what happens to their data.

Run

Run it on autopilot or on demand, or both in the same process.

The payoff: a disjointed chain of scripts, schedules and one person's memory becomes a single process. It is modern, the data owners still control it, and a failure shows up in minutes instead of weeks.

Engage

Engage more people in the pipeline itself — not in tickets about it.

The payoff: time to data drops, because the work goes to the person who can already do it instead of into a queue. People who are closer to their data also use more of it, and use it better.

Know

Know exactly what happens, and prove it.

The payoff: trust. People trust what they can see through, predict and control. Incidents are short, audits are cheap, and one person's resignation does not cost you the pipeline.

Own

Own it end to end — your infrastructure, your data, your experience.

The payoff: your data stays where your policy says it stays. Security stops being the reason a project cannot happen. Your platform cost does not rise because your business grew.

Score all potential vendors against these five criteria. A platform that's strong on Build and Run but weak on Know and Own will accelerate your team today and create risk later.

After you buy: implementation

The implementation stage of a data integration project is critical. And there's one lesson above all to keep in mind during this time: expect trouble!

Yes, you heard right. Even if you've made the best possible vendor choice, implementing any data integration initiative is guaranteed to cause a degree of pain, there's simply no way around it. Building any deeply technical solution is always a challenge, so recognize that from the start and avoid unpleasant surprises.

This is why it's important to know what the vendor can offer to mitigate challenges. Be sure the company can properly train your team(s) and provide you with the knowledge and best practices you need to succeed. Understand how its tech support works. And familiarize yourself with the vendor's other capabilities as well, so you know what it can do to help when you need more than incident-based tech support.

If you're paying for professional services as well as software, be clear about the deliverables and scope. For instance, does the vendor provide documentation of its work? If so, ask to see some past examples, does the documentation generally look helpful, or is the company just ticking boxes?

Also know how the vendor will work with you. That is, when and how does it take control of projects if necessary? How does it step in to either handle things to guide you through the processes necessary to get things right? Make certain the vendor will work according to your specific preferences on projects, big or small.

If you're bringing in a vendor to help with a data integration project, you should have a clear roadmap that defines the scope of the project, as well as milestones and timelines. The key project leaders, both on the vendor side and in your own organization, should also know what their roles and responsibilities are, who they'll be working with and what deadlines they'll be working toward.

Future considerations

In addition to enabling short-term improvements, better data integration can also pave the way for longer-term business benefits: improved competitiveness in the digital marketplace, greater capability to implement AI and other advanced technologies, and more. Stay alert to these opportunities as you put your new data integration processes to work.

Dos and Don'ts

Do have clear goals and expectations in place as you prepare to implement your software.

You should also think about metrics for measuring success. It's important that everyone involved – all stakeholders – agree to these. What's more, everyone at all levels, from your management to the vendor's developers, should be clear about the benefits you're aiming for. This will help to avoid nervousness and headaches during implementation.

Don't neglect people and business culture.

Give your employees the training, support and other tools they need to use your new solution from the moment it goes live. Also, don't be surprised if you encounter resistance from people who are used to doing things differently, for example, from dev teams accustomed to building everything from scratch with scripts and SQLs. Try to maintain a balance between efficiency and transparency throughout this process.

Do expect knock-on effects after implementation.

New insights will open up new opportunities for better working, cost savings and more.

Don't underestimate the potential for change

Better data integration can pave the way for future innovation and new technologies.

Do make sure the vendor, software and licensing structure can meet your needs.

As you put your new tools (and new expertise) to use to solve more and more data problems, you may be presented with opportunities to innovate with 'outside-of-the-box' thinking. Hopefully, your vendor will not penalize you with charges, when you find new and interesting use cases for the technology.

Don't be afraid to ask your vendor for advice.

Engage with your vendor to learn how to solve new data challenges that arise.

Conclusion

The bottom line is: choosing the right data integration vendor is like choosing a personal trainer when you want to get into good physical shape. In both cases, you should know your needs and goals, as well as your limitations. And you should look for a partner who's honest about the challenges you'll face, but has what it takes to get you through them.

Getting in good physical shape takes months and months of hard work, as well as the occasional setback. There's no magic pill.

The same is true for implementing a successful data integration solution.

Making the right choice for your business takes research and preparation, but those efforts can pay off in multiple ways. When you find the vendor and software that's best for your needs – your organization type and size, your types of data, your skillsets and more – you maximize your ability to reap both short- and longer-term benefits. By following the advice in this buyers' guide, we hope you can find the solution that's ideal for your company's unique requirements.

"Making the right choice for your business takes research and preparation, but those efforts can pay off in multiple ways."

If you have more questions, CloverDX is ready to help you find the answers you need.

Go to www.cloverdx.com or email info@cloverdx.com to find out more, [try a demo](#) or [download a free trial](#).

Popular use cases for data integration software:



Data Ingest

- Transformation & data cleansing capabilities
- Management of large number of simple transformations
- API support for cloud and app integration



BI & Analytics

- Automation for regular updates
- Processing performance
- Ad-hoc data prep capabilities
- Support for calculations & external libraries



Platform & Application Integration

- Connectors to well known cloud services
- Generic tooling for REST and SOAP-based services
- Ability to design own connectors
- Performance on small batches
- Monitoring and alerting capabilities
- Effective data pipeline management



Data Migration

- Data transformation capabilities
- Workflow orchestration
- Data discovery & documentation capabilities
- Transparency of the data process for users to see what's happening to their data
- Ability to export data in Excel for users to work on their data and then feed back to the process



Data Error Management & Data Quality

- Validation components & data quality reporting out of the box
- Formatting in & out of Excel
- Integration with 3rd party validation & enrichment services

The Buyers' Guide to **Data Integration Software**

The data integration platform that puts you in control.

CloverDX is a data integration platform with AI-assisted pipeline development, AI-enabled transformations, and self-service tools for business users, deployed on-prem or in the cloud, on your infrastructure and your terms.

Build pipelines visually, in code, or with AI. Run them on autopilot or on demand. Engage more of your people in the work itself. Know exactly what happens at every step, and prove it.

And own the whole thing: the infrastructure it runs on, the data inside it, and the experience your people use