

**RPA GUIDE**

# **HOW TO PREPARE DATA TO AVOID PITFALLS**

**minit**



## **RPA GUIDE**

### **HOW TO PREPARE DATA TO AVOID PITFALLS**

Published in 2019 by  
Minit, j. s. a.  
Headquarters office  
Kosicka 56  
821 08 Bratislava  
Slovak Republic  
Tel: +421 2 555 615 89  
E-mail: [sales@minit.io](mailto:sales@minit.io)  
**[www.minit.io](http://www.minit.io)**

© 2019 Minit, j. s. a.  
Written in 2019 by  
Michal Rosik

All rights reserved. No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photography, recording or any information storage and retrieval system, without crediting a publisher.

Author: Michal Rosik, CPO at Minit  
Book and cover design: Robert Lively  
Senior Editor: David Miller  
Photography: Unsplash

## Introduction to RPA

2020 Is the Year of RPA

### Chapter 01

## How to Set Up Data for RPA Success

Set, Document, and Publicize Project Goals

Loop In All Relevant Stakeholders

Work Iteratively

### Chapter 02

## Documentation is Crucial

The Data Story Solution

Context Is Key

### Chapter 03

## 3 Primary Ways Data Impacts RPA Deployments

Data Input

Data Integrity Relies on Process Integrity

Information Security and IT Processes Are Key

Process Mining for Visibility, Accountability, and Monitoring

Data Output

Reminder: Data Security

### Chapter 04

## The Road to RPA Success? Process Mining



# Introduction to RPA

Robotic Process Automation is data-driven and more accurately, data-dependant. In RPA, moving fast and breaking things is not the way forward. An unplanned RPA transformation puts an organization at risk for data breaches, regulatory missteps, mass employee exodus, and creates a metaphorical bull in a china shop.

Yes, act with fever, act with speed, act with agility. But above all, RPA implementation requires you to act according to plan. If you move too fast and break things, you're doomed. If you plan and measure, you're poised for success.

This guide was created for decision makers and leaders responsible for making RPA a success.

# 2020 Is the Year of RPA

Up to this point, RPA was something for early adapters. Those brave leaders who seize new technology before it reaches critical mass. Not anymore.

Gartner predicts the [RPA market will top \\$1 billion by 2020](#) , and with a 63% growth rate from 2018 to 2019, RPA is currently the fastest-growing segment of the global enterprise software market.

*As we move into 2020,  
40 percent of large enterprises  
will have integrated robots into  
their business processes.*

[\[Source\]](#)

The problem, however, is that nearly half of all businesses that employ RPA fail the first time. An [Ernst & Young study](#) estimated that 30-50% of all first time RPA initiatives don't succeed.

Why such a high failure rate?

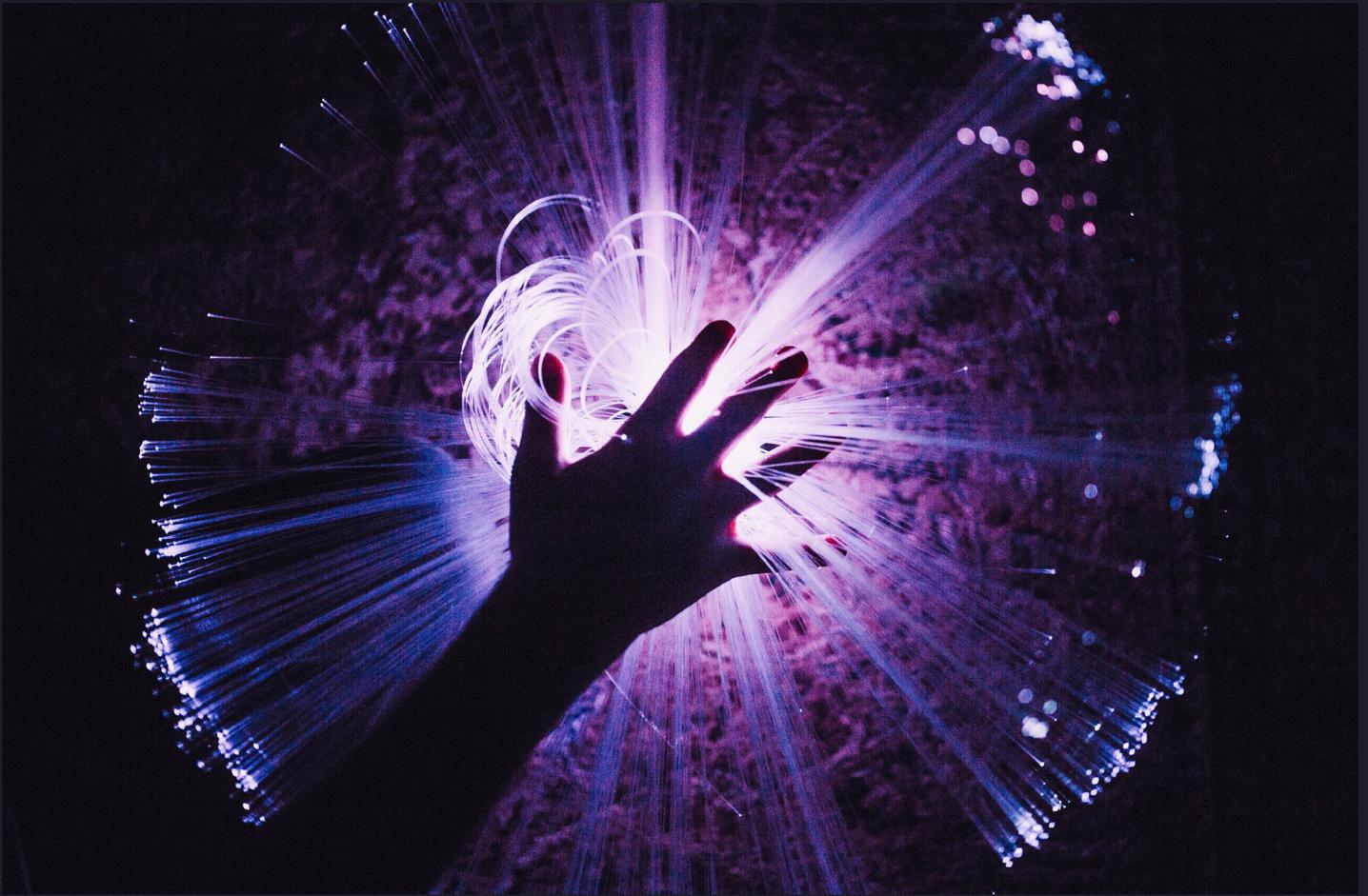
For starters, data. Data is messy, sensitive, and scattered. [Forrester reports](#) that 60% of decision makers at organizations adopting automation technologies like RPA, cite “data quality as either challenging or very challenging — it’s their top challenge when trying to deliver AI capabilities.”

And then there’s that metaphorical bull. And the things the bull breaks are your most precious resources — data.

Stoked by the fear of missing out, some organizations run head first into RPA without considering the ramifications (both good and bad) that are inseparably coupled with RPA.

That’s why you’re reading this ebook: to help guide and plan your organization’s RPA transformation and ensure success up front, while also avoiding common pitfalls.

2020 is the year RPA reaches critical mass. Here’s how to avoid getting trampled.



Chapter — 01

# How to Set Up Data for RPA Success

Insight  
Selection

Set, Document, and Publicize Project Goals  
Loop In All Relevant Stakeholders  
Work Iteratively

Imagine process data as a vast river, each tributary spring contributing its share of water. In the case of RPA, the tributary springs are various departments which add data (water) to the process (river).

Some springs contribute more water than others, but each play an important role in creating the river. A trickling spring at one point may turn into rapids at others. Setting up an RPA project for success means ensuring each spring is treated with respect.

**Don't neglect the small springs:** RPA requires full data access which will come from various departments, team members, systems, and data sources

**Don't build dams:** keep communication lines open and remain transparent

**Don't poison the water:** ensure data integrity and security throughout the entire process

Take the 3 following steps in preparation for an RPA project.

## 01 **Set, Document, and Publicize Project Goals**

Define project goals early on by way of comprehensive documentation. This not only means to record and communicate SMART goals, but also to publicize the location of such documentation.

**FOR FULL TRANSPARENCY,  
BE SURE TO INCLUDE THE FOLLOWING:**

- > **A complete project outline, including a timeline**
- > **Details on what information is being collected**
- > **An outline of how information is going to be used**
- > **Details on how information will be handled and stored**

By maintaining this documentation throughout the project, you also keep everyone on the same page. Stakeholders will know what data you're collecting, what you're doing with it, and reiterate the overall goal of process improvement — not job displacement.

## 02 **Loop In All Relevant Stakeholders**

A successful RPA project is collaborative. Include all decision makers in the developmental phase in order to eliminate holdouts, increase buy-in,

and ensure cross-departmental cooperation.

The process discovery phase of RPA often uncovers inconsistencies and process gaps which are traceable to individuals. The allies you establish early on can help resolve problems in their respective teams, and quell egos which may feel targeted by unsightly findings.

## **Work Iteratively**

Remember the bit about moving fast and breaking things? Working in stages, one process at a time, helps avoid broken glass and establish evidence as you go.

Establish a timeline which allows time for open communication of gains and setbacks. By seeing and hearing about RPA in other departments, employees can prepare themselves for the eventual arrival to their own.



Chapter — 02

# Documentation Is Crucial

Insight  
Selection

The Data Story Solution  
Context Is Key

You need your bots to be an active part of telling your company's data story. And the only way to ensure they're configured to play that role accurately is to provide the input they need to succeed.

The automated process discovery, possible with Process Mining solutions, guarantees the process maps you're using to configure your bots are complete and accurate.

## **The Data Story Solution**

Enter the source of truth (SOT). Establishing a source of truth for your company's data is the best way to tell the story you want, and ensures that story meshes with stakeholders' expectations.

So, how do we tell a data story rooted in the SOT methodology?

- > Consolidate data from each disparate system
- > Set policy dictating input standards
- > Unify output formats
- > Document everything

## **Context Is Key**

SOTs and the data they contain do not exist in a vacuum. As opposed to a SINGLE source of truth (SSOT), SOT maintains context for the data.

For example, if you know the data came from Github, then you won't be surprised when it's in various programming languages. If you know it's from Trello, you can expect it to relate to project management.

In an SSOT, all you may see is pages of unintelligible code, followed by dates and times, making it difficult to point your bots at the proper input files.

It's not uncommon for an organization to set about creating a single source of truth (SSOT). This leads to the discovery that can strip that all-important context, rendering some data useless to other departments. SSOTs bring their own set of possible issues along with all the benefits.

#### **POSITIVE ASPECTS OF SSOT**

One place to look, no matter what data you need. Be it an external audit, tax return information, or order status—it's all in the same repository.

Eliminates data silos, because often times multiple teams work on the same issue.

#### **NEGATIVE ASPECTS OF SSOT**

Can be incredibly time consuming to create. Pulling in data from systems as disparate as an ERP, CRM, and ITSM ticketing takes a lot of collaboration across the entire organization.

Removes context from the data. By compiling all company data in one central repository, all context is stripped from it, making inter-team collaboration difficult.



Chapter — 01

# 3 Primary Ways Data Impacts RPA Deployments

Insight  
Selection

Data Input

Data Integrity Relies on Process Integrity  
Informational Security and IT Processes Are Key

With project plans documented, input sources cleaned up, and cross-department stakeholder buy-in solidified, it's time to organize the bots.

**THERE ARE THREE DATA QUESTIONS YOU NEED TO ANSWER BEFORE PROCEEDING**

**1**  
**How clean is  
the input?**

**2**  
**What format is  
the output?**

**3**  
**How  
secure is it?**

## **Data Input**

Junk in, junk out ... at best. RPA bots require clean data for optimal functionality. Automating a bad process will only exacerbate a problem, not solve it.

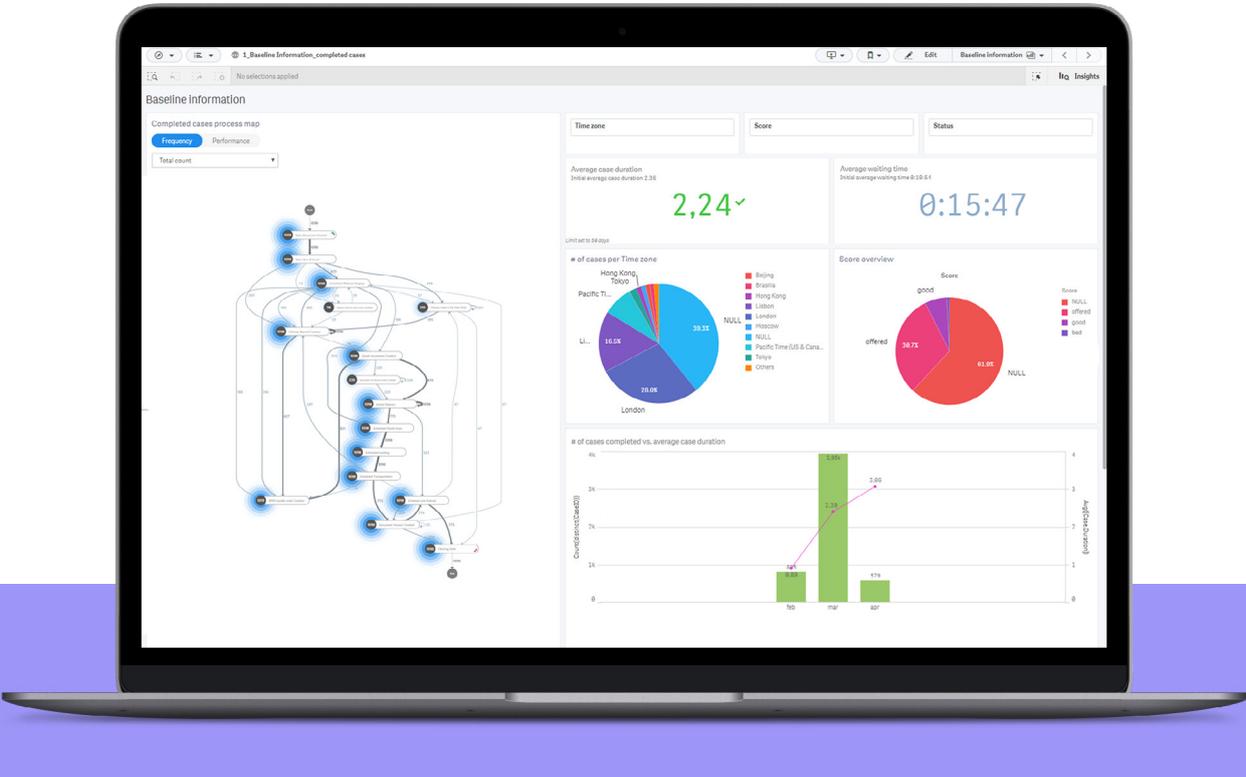
From an account representative manually inputting client information to an automatically generated stock level report from the warehouse, processes dictate what information is fed into business systems. This ultimately being data input for bots.

Many obstacles can derail data integrity projects:

- > Unreliable data source
- > Unreadable data source
- > Conflicting data from multiple sources
- > Lack of data / no data at all

Root causes of data integrity issues vary, but all hold at least one thing in common — a business process is at the foundation.

Data can be corrupt due to non-standard processes, or there may be multiple systems in use that create conflicting data trails. It's also possible that undocumented workarounds or other inconsistencies affect the reliability of data coming out of the relevant processes.



Minit Process Mining dashboard with interactive process map

This is another place Process Mining can save time and headaches. By using the interactive process map, you are assured of having a complete picture of the end-to-end process at the heart of your RPA project.

## **Data Integrity Relies on Process Integrity**

The path to data integrity starts with how information makes its way into business systems. If there is no standardization to data input, there won't be any standardization to data output.

For example, consider a lead generation form for potential customers. If the Country or State field is unrestricted (no drop-down menu), you will end up with dozens of misspelled place names that bots won't know how to handle. Any process being automated by RPA will kick these process runs out to a human, assuming your bots were programmed to take this possibility into account.

## **Information Security and IT Processes Are Key**

Information Security (IS) and IT are the two departments most associated with data integrity practices.

IS as a department has been shifting towards better integration with core business strategies. This transformation in recent years has elevated the importance and visibility of IS practices and processes.

IT departments, on the other hand, are in charge of activities relating to data while it is held within the organization's computer systems. Any process revolving around IT systems has the potential to impact the health and accuracy of information stored on company servers and databases. In addition, it controls how that data is accessed, and by whom.

Both IS and IT need to have clearly documented processes, and a way to monitor these processes to ensure compliance.

C-level executives fundamentally don't trust data. It seems to stem from a lack of understanding of how these back-end processes are run. These doubts are often due to poor visibility into those processes, naturally.

*The way in which IS responds to breaches and asset loss or theft, underpins the security of a company's data assets.*

## **Process Mining for Visibility, Accountability, and Monitoring**

Process standardization is priority #1 when it comes to ensuring the integrity of your company's data.

With standards in place, the output of a process is more likely to be compliant. Process Mining is the best tool to use when it's time to get to work, as the automated process discovery phase will be faster and more accurate than manual process discovery.

The visibility provided by Process Mining is unparalleled. By combing IT records and event logs, mining software compiles an interactive process map that shows the as-is state of any business process. This process map gives the transparency necessary to identify the who, what, when, where, and how of these processes, allowing automation to proceed with clean data assured.

Once a business has its processes standardized, they must be monitored to ensure future data integrity.

Process Mining operations can be set up to generate new maps on a regular basis to red flag process variations or discrepancies.

Once a process is automated, the output will still need to be monitored. Why? Discrepancies may indicate a glitch in the configuration of the bots, or a piece of software has been updated and the bots are unable to conduct the process correctly.

## **Data Output**

This brings us back to the subject of SSOT. Establishing an SSOT can be a tricky proposition considering all of the systems, locations, networks, and so on.

The first step has to be a comprehensive process discovery project. It will locate each current data store, identify which systems are compatible, and determine exactly which system is outputting what data to each location. Only then can you program the bots accordingly.

Process Mining is a logical choice for this phase. However, remember that not all data is going to be stored on minable systems.

While the automated mining operation will produce much of what you need to get started with an SSOT project, manual discovery needs to take place to root out any additional data stores. You never know when thumb drives or Excel spreadsheets stored on CD-ROM may be in use without anyone realizing it.

The next step is to work out precisely where and how to host your SSOT.

This is all based on what the data is and who needs access to it. Various data warehouse or repository tools exist, ranging from self-hosted options your IT department can install on an on-site server, to cloud-hosted solutions that take advantage of distributed networking and redundant server locations.

Or you can go with another option, known as a Distributed SSOT. This is where a company decides that they want each type of data to have its own SSOT, rather than every system dumping its data into one centralized location.

#### FOR EXAMPLE

**CRM**

for customer data

**ERP**

for asset data

**PLM**

for product data

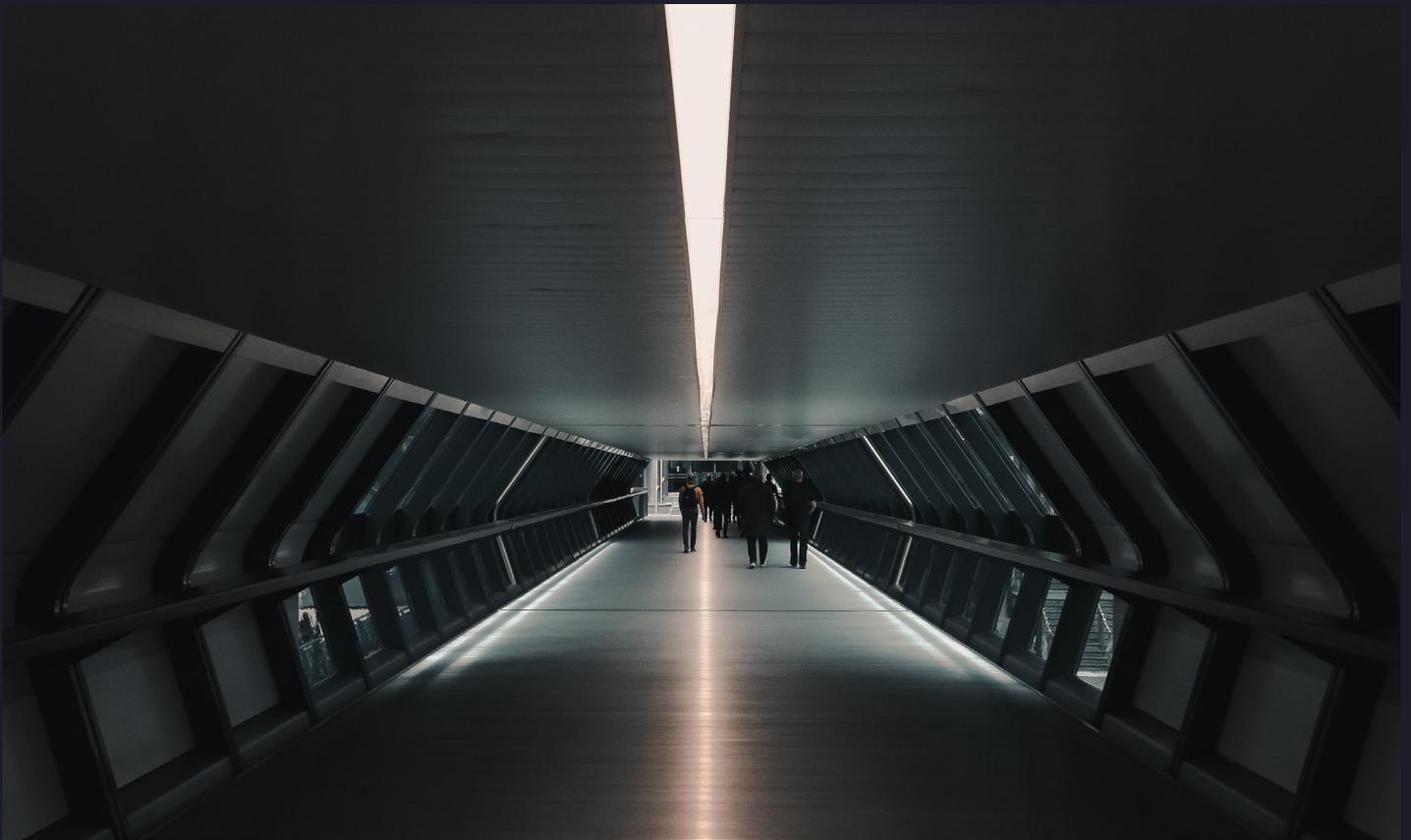
This model relies on company policy. Such a policy sets out instructions for employees to keep things neat and organized, since there won't be one data warehouse to monitor and set centralized rules.

## **Reminder: Data Security**

A large part of your project documentation should be a detailed accounting of what data you're using for what processes and how it will be handled. And part of that accounting should be information on data storage, security measures, and how long you will keep the data.

Data is a touchy subject. Major organizations have undergone highly public data breaches, and others have been caught in compromising situations with how they're using customer information.

The general public doesn't know the extent to which data powers the business world. That means it falls on you to be sure they know what you're doing when you start automating processes that include customer and/or employee data.



Chapter — 04

# The Road to RPA Success? Process Mining

When it comes to something as data-intensive as RPA, many legitimate concerns around the subject of data quality float around.

The data going into an RPA-powered process needs to be clean, and the data coming out of the process needs to be compliant and stored in a usable fashion.

These concerns are the basis for many documented [RPA failures](#), hence our desire to produce this ebook to act as a guide. Use it as an outline when drawing up your own RPA project. Only then will you stand a better than average chance of reaping the benefits of automating business processes.

Ready to set your company on a path towards RPA success?  
Contact one of our experts to [get started with Minit Process Mining](#).

## Minit Academy

### Join the Minit Process Mining Academy

Watch video demonstrations, and access feature-by-feature guides, available virtually, on-demand, any time, and any pace.

[Learn more >](#)

## Follow our Blog

### Follow our Blog

Weekly insight into Process Mining trends, how-to articles and real success stories from your industry.

[Learn more >](#)

## Resource Library

### Checkout our Resource Library

From checklists and how-to guides, to industry specific content, all our downloadable material is free. It's an enormous resources library for you and your colleagues.

[Get the Resources >](#)

## Try Minit

### Start a Free Trial of Minit

Interested in trying Minit software? Schedule a quick call with our representative to discuss your needs and type of Trial.

[Try Minit >](#)



Get in touch with our team to learn how it can help deliver effective business process improvement at your organization.