



Joyful  
Craftsmen

# WHY AUTOMATED DATA TESTING

[WWW.JOYFULCRAFTSMEN.COM](http://WWW.JOYFULCRAFTSMEN.COM)



At Joyful Craftsmen our primary focus is  
Microsoft Data Platform,  
and everything that goes with that.

We boost companies' data competencies  
through development of smart data  
solutions, smooth data operation  
and automated data testing.



Automated data testing itself is magical.

You'd never imagine that turning from manual  
data verification to automated tests can save you  
so much money and unlock so much potential.

The savings and outcome enable you and your business  
to walk an authentic journey to data-driven culture.

This white-paper provides 5 reasons why you should care,  
describing the 5 major pain points you could overcome,  
if automated data testing is applied smartly.

#JOYFULBI





“  
Data matters.  
But only TRUST  
determines  
its value.  
”



# #1. TRUST

Trust in your data is essential. The business world claims to be moving towards data-driven cultures. However, more than 50 percent of companies base at least half of their regular business decisions on gut feeling or experience rather than data and information.

## Is it really that bad?

According to KPMG, only a third of CEOs have a high level of trust in the accuracy of their data and analytics, and one out of five have limited trust for nearly every aspect of the way their organization uses data and analytics.

As perception of trust is a highly subjective area, you need to be able to provide strong evidence that the data being used for everyday decision making in business is accurate and under sustainable control.

You need to ensure the data utilized for your financial statements, your operational systems, or highly important key management decisions, as well as the data behind all those analytics, statistics and reports, is extraordinary.

## So What?

Our vision is to change those statistics. We improve the perception of trust in your own data by an order of magnitude.

Automated data testing provides controlled, repeatable processes, standardization, and stewardship around how your data is handled and perceived.

It's an enabler on the path to data-driven culture!



## CLIENT EXAMPLE

A multinational company needed to increase confidence in the data regarding exchange rates. As a result of frequent changes, it was necessary to implement a set of tests that cover critical financial data marts. Those checks are executed with each exchange rate update.

## THE TAKEAWAY

With automated testing, the customer has immediate insight into exchange rates and the quality of related data marts. She knows she can trust the data and can thus make business decisions without worry.



# #2. EVIDENCE

Testing is nothing else than the process of improving your confidence through evidence! All your data stakeholders are fundamentally interested in an overall view of how your data is doing - what its actual quality, reliability and availability level is, whether it is on track or due to be ready when expected.

## How About You?

When your stakeholders make a decision, how much do they rely on available data as opposed to their gut feeling? And when using data, how much confidence in it do they have?

Without having a reliable source of evidence about data quality and freshness, it is very hard to solidify trust in your data audience.

Because there's little trust or confidence without evidence. You have to be able to bring clear and reliable information about how your data is doing into the hands of your stakeholders.

## So What?

The natural outcome of automated data testing is all sorts of useful and interesting data (about data) that can indicate the success or failure of the test efforts and thus describe overall data quality.

You can easily track what has happened in the past, what's happening now, and allow yourself to look at the data to make an educated guess about the future.

Bringing all this data in a friendly format to your business is precious added value which sets you on the way of improving the trust in your own data.



## CLIENT EXAMPLE

A client was not sure in which parts of the developed application errors occurred and was threatened with not being able to deliver by the agreed deadline. Our automated tests suite, which also included a configurable report suite with test results, overall execution summary and trends provided an immediate overview of the current status of the project.

## THE TAKEAWAY

Thanks to well-arranged reports, the client identified critical areas of the project, and could thus prioritize accordingly.



# KNOWLEDGE

“ Evidence is knowledge, knowledge is power ”





# Agile or Irrelevant

KPMG 2019

## #3. INSTANT

Receiving feedback is one of the most vital elements of making progress. The more you automate, the more feedback you get. The more feedback you get, the more you can improve. The more you improve the more resilient your business becomes.

### Continuous Feedback

In today's world instant feedback (data) is a necessity and your data solutions are no exception.

Instant feedback on the status of your data, gives you a continuous feedback loop on how your data is doing. Such feedback gives you valuable insight for future actions and further reactions.

### So What?

Automated data testing solutions become your reliable foundation for providing instant feedback on your data.

You get a wealth of insights for discovering bottlenecks and problematic areas of your data solutions. If used smartly, you can dramatically shorten learning cycles, boost your business agility, and thus accelerate innovation and value delivery.



### CLIENT EXAMPLE

A client was developing a data solution that contained more than 120 measures. It was crucial the metrics be reviewed each time the app changed. With nearly 500 automated tests the app was verified over 60,000 times in 50 days.

### THE TAKEAWAY

With automated tests that are performed on a regular basis, the customer is informed daily about the accuracy of the developed data app, and can thus refocus on a daily basis, too.

In addition, more than 600 days were saved on manual testing!



# #4. PERSISTENT PREDICTABLE

Testing is a critical part of any software development project. Each individual solution has a set of test cases to specifically evaluate its unique features, functionality and requirements. Running these tests is not a one-and-done affair; we have to ensure that the program will be evaluated throughout its lifetime, necessitating easy access and repeatability in its test set.



## Repeatability!

There will always be defects! Repeatedly executable testing suite helps you avoid these without interruption. Regardless if the defects are caused by a new feature, some other defect fix, or any undesired intervention, you have continuous feedback available.

Persistent data testing set can be applied on various environments, including production! This gives us the ability to uncover undesirable effects caused by different operating environments.

Regression testing also has the power to hold the product team accountable in maintaining performance standards, upholding quality expectations and ensuring the project is primed for success.

## So What?

Repeatable testing suite gives you continuous feedback on how your data is doing. This gives you the ability to predict which problems could possibly arise, and to what extent existing features have been compromised.

All these continuous insights bolster your certainty, allowing you to release your data solution to users, confident in providing data at the level of quality your users have come to expect.

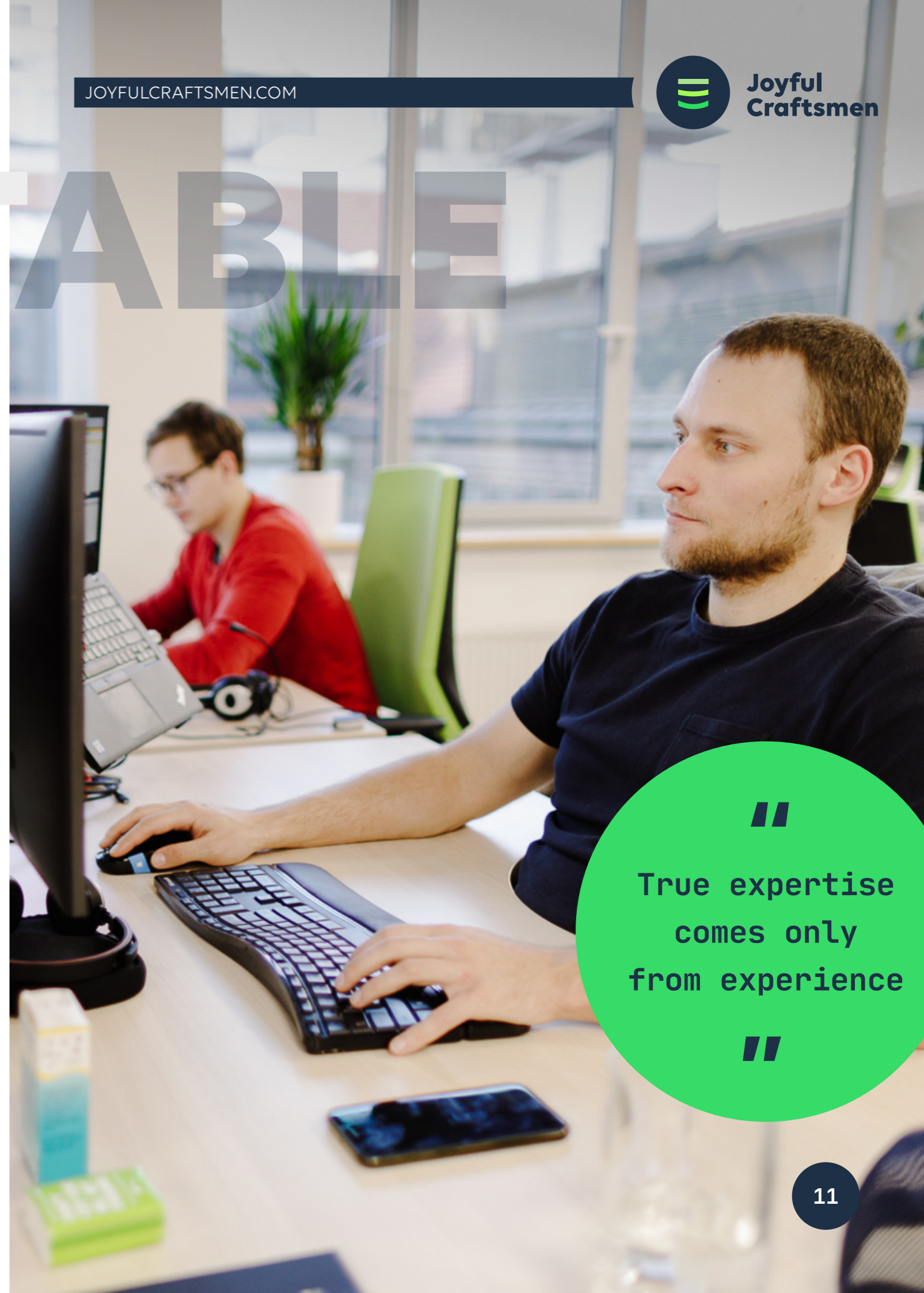
## CLIENT EXAMPLE

A customer had dozens of tests that needed to be run manually after each release. We delivered a set of automated tests that can be easily repeated every day, plus applied to various development environments.

## THE TAKEAWAY

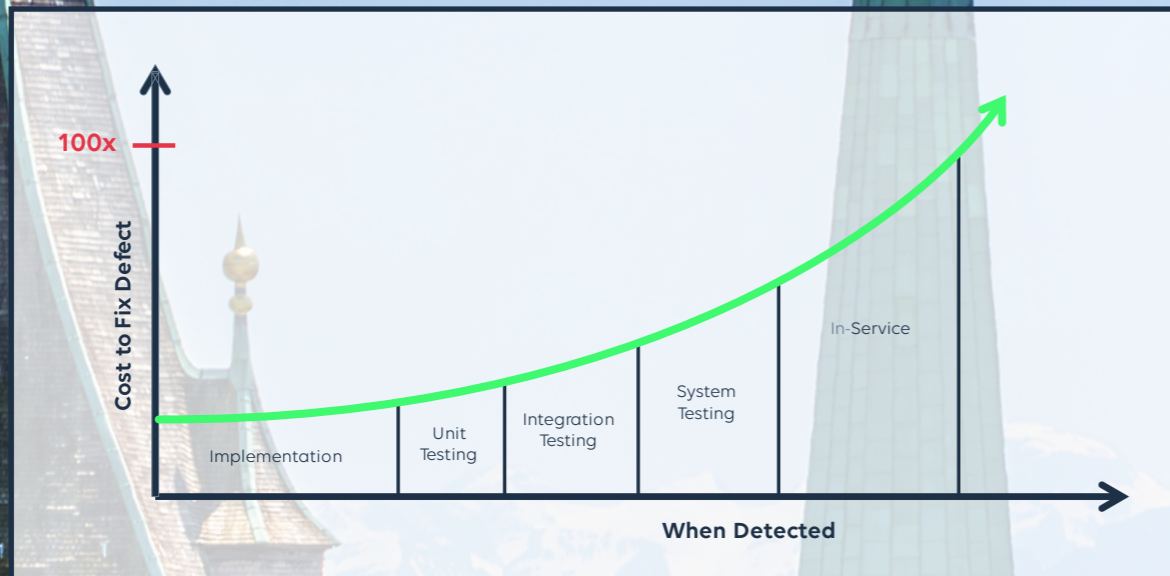
People make mistakes, automatic solutions do not. The customer has the same feedback every day and knows she can rely on it.

Btw the full test suite was also executed tonight.



“ True expertise comes only from experience ”





## #5. IMMEDIATE

The impact of a data issue on a production system can be immeasurable. Thus the sooner we are able to detect defects the fewer resources needed to fix it. The detection needs to be quick. It needs to be immediate. As well as the remediation.

### Time is Money

How many times were you in a situation when data issues caused you serious costs?

It's a well known truth that the later a software defect is found, the more needs to be spent on its mitigation. If such issues make it through to the end users, we need to count on product support calls, bug analysis and tracking, related management, possibly fixing, retesting, releasing the software, and - in a worst-case scenario - a product recall or lawsuit.

The more time that's passed since a product's release, the more additional time is required to fix issues. Based on our experience, the multiplication factor between fixing an issue the same day or only few weeks later is between 100-1000x.

### So What?

Stomp out issues early, and your product will cost less than it otherwise would.

Automated data testing can give you feedback that your solution still does what it's expected to do in matter of minutes. The team is then able to reach new levels of productivity, regardless if it's a green field, complicated or legacy data solution.



### CLIENT EXAMPLE

As part of a development of BI solution, we identified the need to reduce the number of production errors found in order to save time spent by the development team on support. We implemented a set of automated tests that is executed immediately upon every change made to the solution. The feedback is available in 20 minutes!

### THE TAKEAWAY

Thanks to these tests, the development team can immediately identify and thus immediately fix such errors. The number of errors found in production and related cost was reduced significantly.



# TYPICAL SCENARIOS

## Data Migrations

Typically you migrate from one system to another and barely have a clue how successful the migration is. How much has been lost during the data transfer from one system to another? You are usually very limited in the number of retries, as you are most likely heavily into retesting those migrations. It usually takes a week(s) to gain confidence, so you can release your business people to test it.

## Data Access

Data must never get compromised. This is rule number one in any serious data solution. In this case, you need to ensure and be sure, that a user (or set of users) have access to exactly that data they should. But more importantly they must not see the data they shouldn't. Automated checks validating user role data access is a great booster in data security.

## Data Consistency

Data consistency is essential. Because all your applications work with data, it must always be correct, otherwise you only amass damages. It's very common that your data get broken as soon as a new change is applied, regardless if it's caused by the system itself or any externally dependent application. Covering your data with consistency checks permanently ensures the data are fine and were not broken by any internal or external modification.

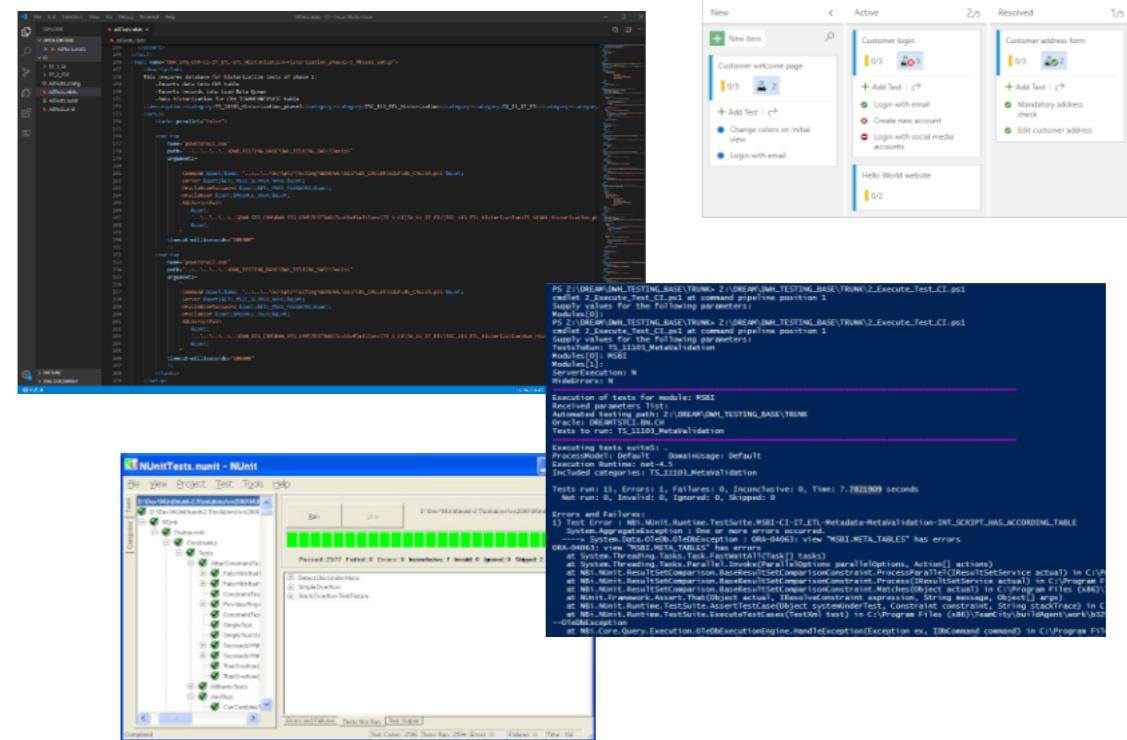
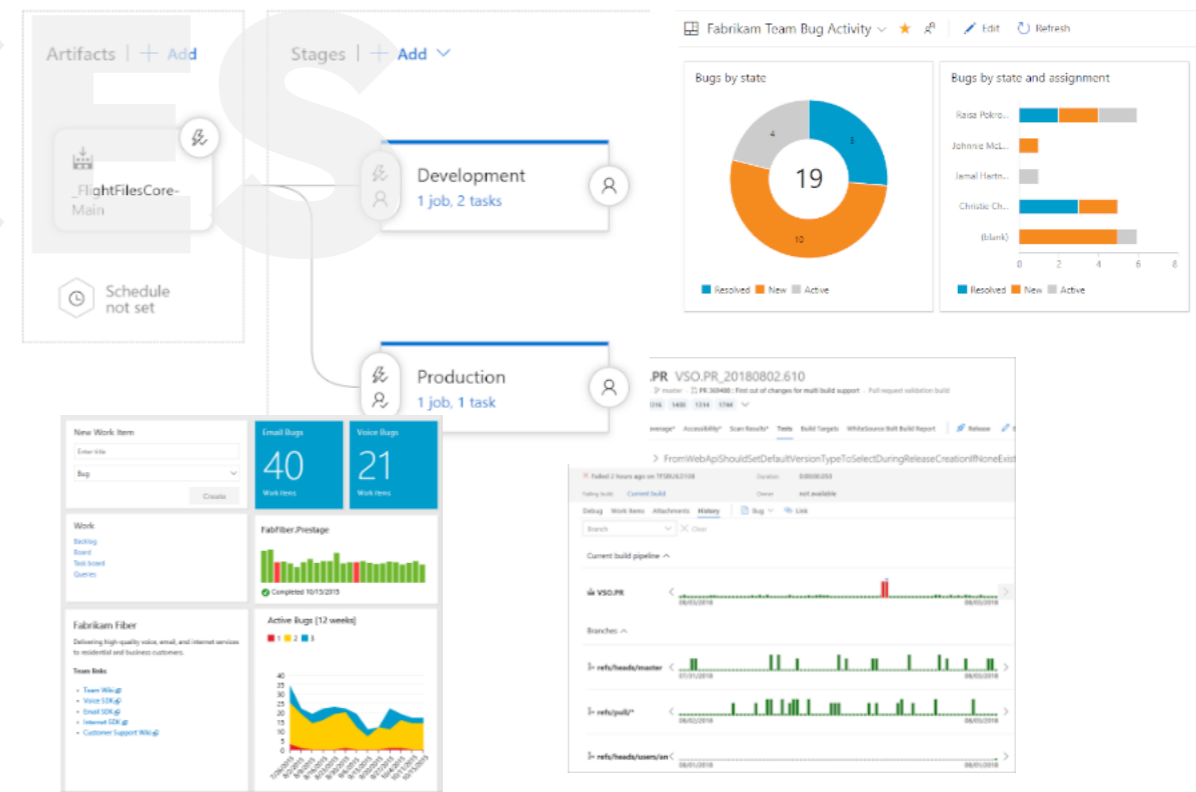
## Data Architecture

Data solution architecture always consists of many rules, various concepts and tons of patterns. It defines how data should be extracted, loaded, transformed, checked, flagged, removed, enriched, stored, reported etc. The usual goal here is to build a sustainable solution providing business value incrementally. Automated architectural cross-checks ensure the development team isn't compromising common and agreed design patterns and therefore is not compromising the future value.

## Data Quality

Data quality is everything. As was already mentioned there's no value in data without trust. Data quality checks, usually focusing on data through the lens of business, ensure the data fulfill the required business criteria. It's continuous feedback to all stakeholders that data is of a quality they all expect.

Automated data testing is a priceless approach that gradually improves your trust and the confidence that the data you use daily is fine.





# AUTOMATION IS ESSENTIAL

How long does it take when you manually run just a few hundred test cases? How often you can rerun those? What else could you do instead of manually executing all those test cases? When using the proper tools smartly, you can easily reduce the time for test execution by an order of magnitude. Eventually you can invest the time you've saved into more value-adding activities...

Automation does not come out of nowhere. To be able to automate full application development life-cycle (ALM), which means having a central repository of your code, automated build and test procedures, and up-to one click deployments to any of your data environment (including production), substantial knowledge and experience are needed.

Luckily today, there are plenty of tools you can easily orchestrate into a beautiful concert of automated processes that gives a real face to the buzzword "DevOps".

**Contact us for more details  
if data DevOps (aka DataOps)  
is something you'd like to rule, too.**

Microsoft  
Partner



Gold Data Analytics  
Gold Datacenter  
Gold DevOps



**We are friendly, skilled, joyful.**

Don't be afraid to get in touch with us!  
Let's discuss your data testing ideas, visions or problems.  
We'll find the smart solution together. Guaranteed!

**TALK TO US TODAY!**

**YOUR DATA DESERVE IT!**



**CZECH REPUBLIC**

Bredovsky Dvur Centre  
Olivova 4  
Prague  
110 00

**+420 772 001 100**  
sales@joyfulcraftsmen.com

**+420 734 346 168**  
hello@joyfulcraftsmen.com

**WWW.JOYFULCRAFTSMEN.COM**



**Joyful  
Craftsmen**

**Microsoft  
Partner**



Gold Data Analytics  
Gold Datacenter  
Gold DevOps