How to Evaluate Scrum Team Performance

A practical guide for managers to effectively evaluate the performance of their Scrum Teams + Case Study.

Mehdi Hoseini

Certified Trainer with Scrum.org (PST)

Scrum Team Performance is NOT a simple one-aspect, but a multi-aspect concept.

Imagine you want to compare two cars, saying which one is better.

To do so, you need to consider multiple parameters:

1. Safety
2. Fuel consumption
3. Comfort
4. Brand reputation
5. Durability
6. Environmental impact
7. Cost of maintenance
8. Build quality
9. Design & Aesthetics
10. And many more ...

Be aware that your context and goals have a crucial role in the evaluation. For example, Americans generally prefer large cars, while Europeans prefer small cars.

Performance Evaluation

Is

- ✓ Focused on outputs & impacts
- ✓ About triggering quality conversations
- ✓ About initiating right improvements
- ✓ For collective interpretation
- ✓ For having meaningful insights

Is NOT

- ✓ Focused on utilization and efficiency
- ✓ Used for blaming people
- ✓ About comparing multiple Scrum Teams

Categories of Evaluation Criteria

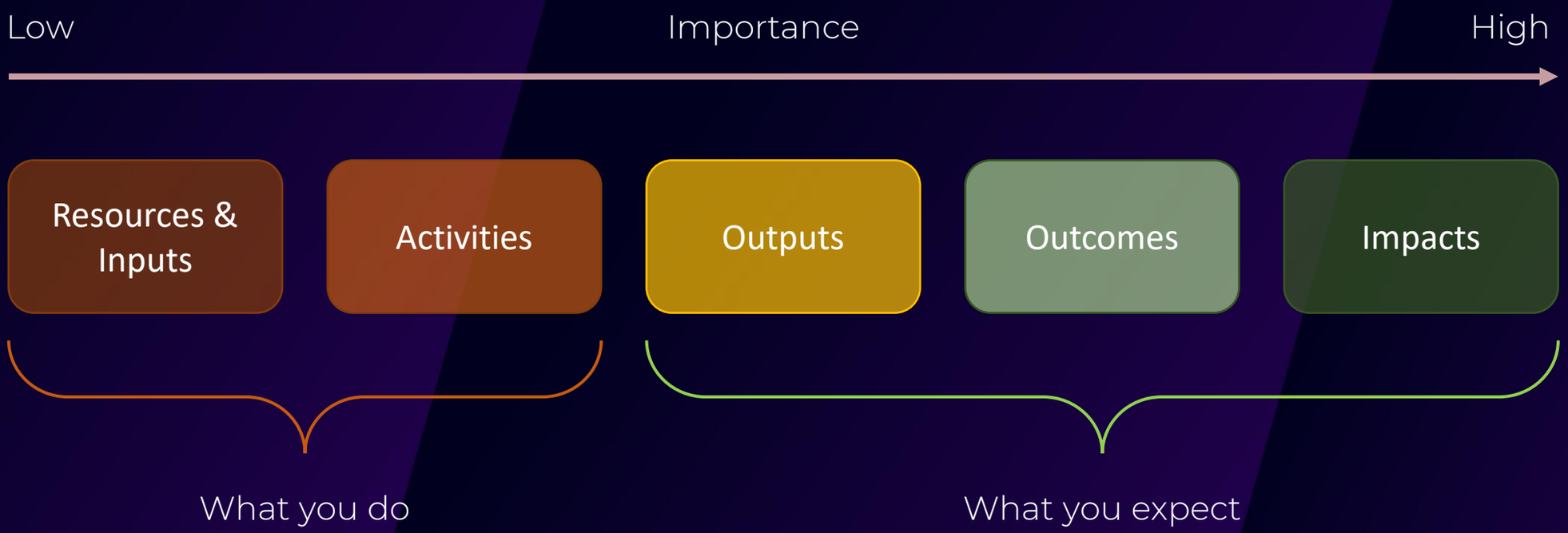
Performance evaluation needs two main categories of criteria

Category 1

Quantitative Criteria

Five Categories of Quantitative Metrics

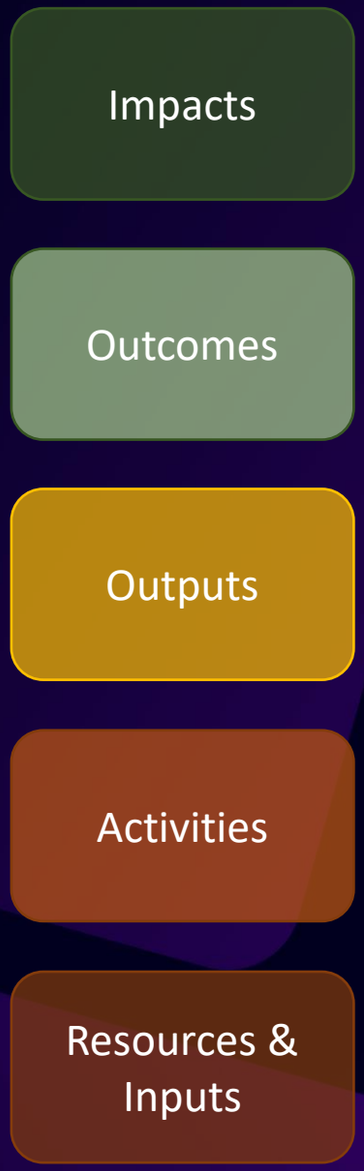
In product development and delivery, you start with doing something to achieve something.



Scrum Team success is defined by meeting the expectations, not doing more.
So, for performance evaluation, your focus should be on the right side of the spectrum.

Metrics Importance

Each priority overpowers the next priority.



Priority 1

The ultimate effect on the lives of customers and the business.

Priority 2

The value perceived by the customers includes achieving dream gains or getting rid of pains.

Priority 3

What is created and delivered (Increments)

Priority 4

What you do to create Increments (tasks, ...)

Priority 5

What you use to do the activities (time, effort, ...)

Metrics Specifications

Let's check how you should treat each category of metrics.

Impacts

Is a lagging indicator and a long-term concept. So, you cannot evaluate it right after each Sprint.

Evaluation interval: Every 6 months, year, or longer periods.

Your regular focus should be here based on their priorities.

Outcomes

Is a lagging indicator and a short-term to mid-term concept.

Evaluation interval: Every Sprint, month, quarter, or longer periods.

Outputs

Is a short-term concept.

Evaluation interval: Every week, Sprint, or longer periods.

Activities

Is a short-term concept.

Evaluation interval: Every week, Sprint, or longer periods.

Resources & Inputs

Don't use this category for evaluation. They are poor indicators of value. So, don't use destructive metrics like log work, spent time, velocity, ...

Case Study

MetaLearn

A double-sided learning platform that connects trainers with learners. Let’s define quantitative metrics for it.

Impacts

The ultimate effect on the lives of customers and the business.

Trainer

Impact:
Building a sustainable training business.

Learner

Impact:
Growth over their professional life.

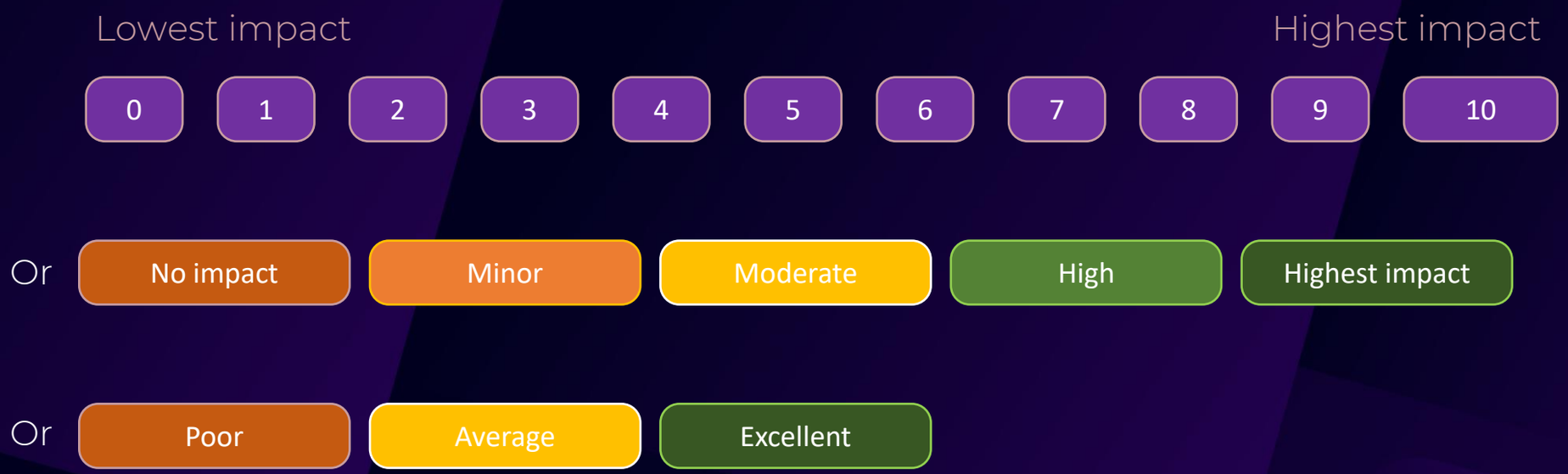
MetaLearn Platform

Impact:
Becoming a thriving business that circulates the world's knowledge.

Impacts

A great way to measure impacts is to run a survey asking customers how their lives have been impacted by using your service.

As an example, use a score-based question similar to NPS (Net Promoter Score).



Outcomes

This is the most important operational level of measurement.

You need to sit with the Scrum Team and stakeholders to define what your customers value most and create metrics around customers' perceived value + define the evaluation intervals. For example, for trainers, it is crucial to see that by using your service, their monthly income is increasing. They deeply value it.

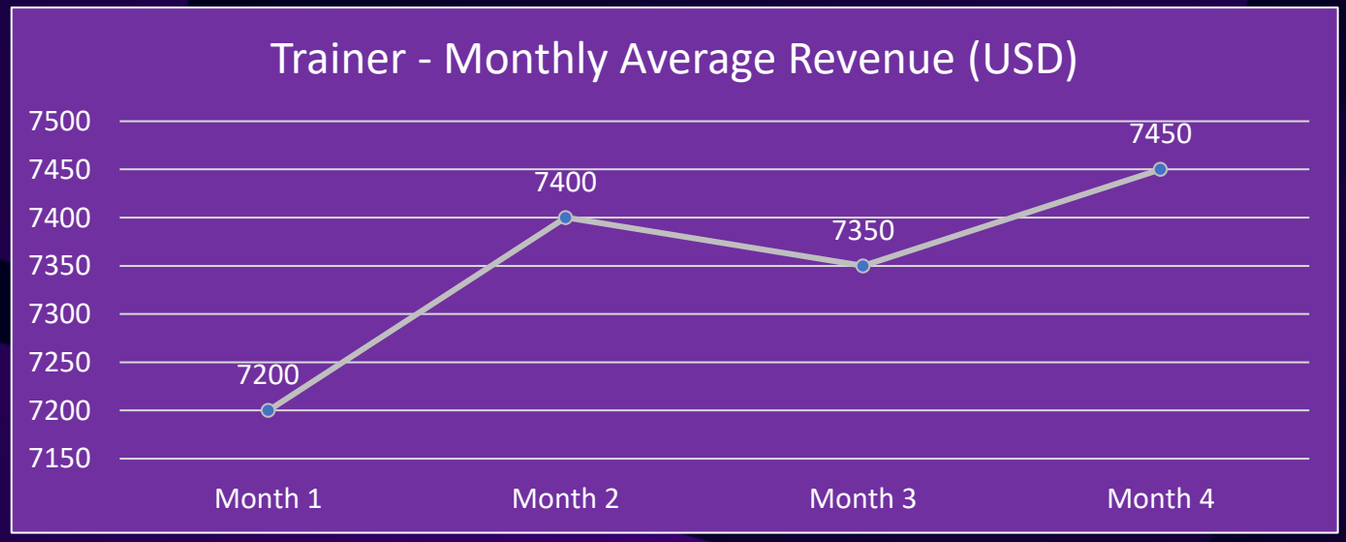
Trainer Side Metrics

- Average Monthly Revenue
- Average Learner Acquisition Cost
- Average Monthly Acquired Learners

Learner Side Metrics

- Average Monthly Active Learners
- Average Learning Time
- Search to Conversion Ratio

Example



Outputs

This is a lower priority than the outcomes, yet shows the efficiency of your delivery pipeline.

You need to sit with the Scrum Team to define metrics that show how efficient you are with building and delivering Done Increments. For the MetaLearn platform, we use the following metrics for outputs and measure them at the end of each Sprint:

- **Throughput**

  The number of Product Backlog Items finished per unit of time (usually Sprint).

  This indicates the team's ability to produce an acceptable amount of outputs per unit of time.

- **Customer Cycle Time**

  The amount of elapsed time between when the implementation of a Product Backlog Item starts and when a Product Backlog Item is released.

  This indicates how fast the Scrum Team is to implement a single Product Backlog Item.

- **Customer Lead Time**

  The amount of elapsed time between when a Product Backlog Item is requested and when a Product Backlog Item is released.

  This indicates how responsive the Scrum Team is to deliver a single Product Backlog Item.

- **Defects**

  Total number of open defects.

  This indicates the quality of the Increments that the Scrum Team produces.

Activities

This is the lowest priority compared to the previous categories.

You need to sit with the Scrum Team to define metrics that show the most important activities helping you create and deliver outputs. For the MetaLearn platform, we use the following metrics for activities and measure them at the end of each Sprint:

- **Deployment Frequency**

  The number of deployment per unit of time (usually Sprint).

  This indicates how frequently the Scrum Team delivers Increments to the customers.

- **Average Time to Restore Service**

  The average amount of time it takes for the Scrum Team to recover from a failure in production.

  Quick recovery indicates a focus on resilience and observability.

Tip: The Outputs and Activities categories are close to each other, and you may use these categories of metrics interchangeably. No worries. It's not a big problem. Just try to choose metrics that best suit your context.

MetaLearn Metrics Dashboard

Impacts

- Trainer Impacts Chart
- Learner Impacts Chart
- Platform Impacts Chart

Outcomes

Trainers

- Average Monthly Revenue
- Average Learner Acquisition Cost
- Average Monthly Acquired Learners

Learners

- Average Monthly Active Learners
- Average Learning Time
- Search to Conversion Ratio

Outputs

- Throughput
- Customer Cycle Time
- Customer Lead Time
- Defects

Activities

- Deployment Frequency
- Average Time to Restore Service

Category 2

Qualitative Criteria

Qualitative Criteria

This category is more like a health check for Scrum Team. At the end of each Sprint or month, you run a survey among the Scrum Team members. It is like a self-evaluation checklist as follows:

No | So So | Yes

Row	Criteria	Member 1	Member 2	Member 3	Member ...
1	Have we met the Sprint Goal?				
2	Have we created a Done releasable Increment?				
3	Do we have a clear Product Vision?				
4	Do we have a transparent Product Goal?				
5	Do we define a Sprint Goal for each Sprint?				
6	Do our Sprint Goals support achieving the Product Goal?				
7	Have we kept our Product Backlog actionable?				
8	Is our Product Backlog Refinement effective?				
9	Do we have Sprint Planning?				
10	Is our Sprint Planning effective?				
11	Do we have Daily Scrum?				
12	Are our Daily Scrums effective?				
13	Do we inspect our progress toward the Sprint Goal in the Daily Scrums?				

Qualitative Criteria

Row	Criteria	Member 1	Member 2	Member 3	Member ...
14	Do we have Sprint Review?				
15	Do we have key stakeholders in the Sprint Review?				
16	Are stakeholders engaged in the Sprint Review?				
17	Is our main intention in the Sprint Review getting feedback from the stakeholders, not a one-way presentation?				
18	Do we have Sprint Retrospective?				
19	Is our Sprint Retrospective effective?				
20	Do we keep a backlog of improvements?				
21	Do we have at least one real improvement in each Sprint?				
22	Do we regularly update the Sprint Backlog?				
23	Do we have a Definition of Done?				
24	Do we use the Definition of Done in our daily work?				
25	Is our Definition of Done getting improved over time?				
26	Does our Definition of Done really make a difference in the quality of the Increment?				
27	Do team members collaborate?				
28	Do we adhere to the Scrum Value of Commitment?				
29	Do we adhere to the Scrum Value of Focus?				

Qualitative Criteria

Row	Criteria	Member 1	Member 2	Member 3	Member ...
30	Do we adhere to the Scrum Value of Openness?				
31	Do we adhere to the Scrum Value of Respect?				
32	Do we adhere to the Scrum Value of Courage?				
33	Do I feel psychologically safe sharing whatever I have on my mind?				
34	Do I feel managers in our organization are supportive?				
35	Do we effectively resolve impediments?				
36	Do we effectively resolve conflicts?				
37	Do we have fun in the team's dynamic?				

How to use this self-evaluation checklist

- ✓ Your Sprint Retrospective is a great place to fill this checklist.
- ✓ Ask every team member to fill out the checklist individually.
- ✓ Tell them not to think too much about each question. The first thing that comes to their mind is the true answer.
- ✓ Just a few minutes are needed to fill this checklist.
- ✓ After filling in the checklist individually, show the answers of all members beside each other. It reveals the trend.

Sample Checklist

No | So So | Yes

Row	Criteria	Member 1	Member 2	Member 3	Member ...
1	Have we met the Sprint Goal?				
2	Have we created a Done releasable Increment?				
3	Do we have a clear Product Vision?				
4	Do we have a transparent Product Goal?				
5	Do we define a Sprint Goal for each Sprint?				
6	Do our Sprint Goals support achieving the Product Goal?				
7	Have we kept our Product Backlog actionable?				
8	Is our Product Backlog Refinement effective?				
9	Do we have Sprint Planning?				
10	Is our Sprint Planning effective?				

- ✓ Prioritize items with the red color for improvement.
- ✓ Items with yellow color are candidates for discussions.
- ✓ Items with various colors are candidates for discussions.
- ✓ Count the number of green boxes and monitor its trend. If your checklist is getting greener over time, you are on track.
- ✓ You can call this checklist the “Scrum Team Green Map.”

Visualization

How to visualize metrics to make the data available and visible to all.

Tools

There are great tools that you can use to visualize your metrics. See below examples.

Microsoft Power BI

Google Looker Studio

Metabase

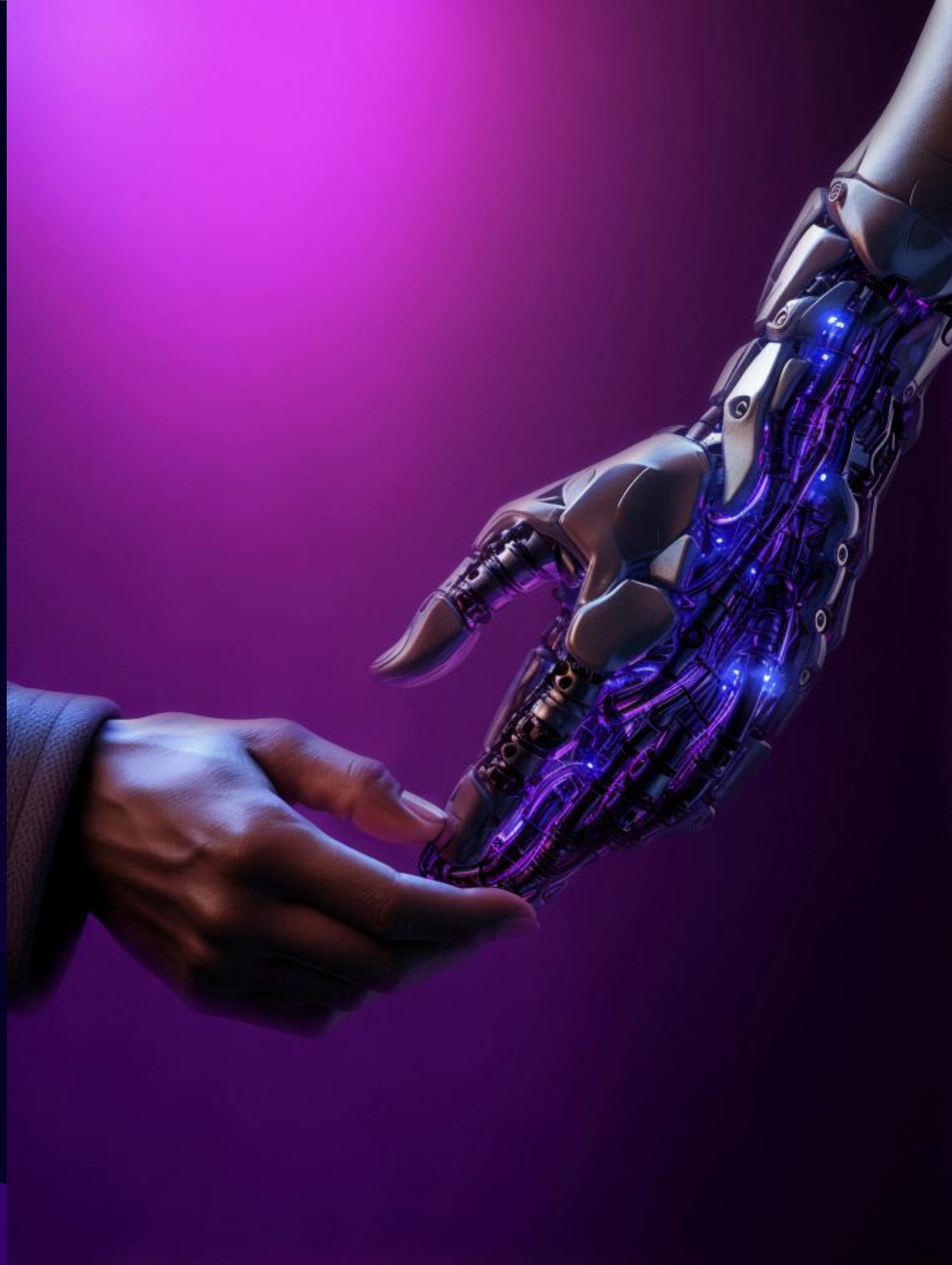
Evaluation Process

Setting up the required infrastructure and process for the Scrum Team performance evaluation needs a relatively high amount of time and energy.

1. Sit with your team and stakeholders to define quantitative criteria.
2. Decide the interval for the evaluation of each metric.
3. Define from where you want to collect the data for each metric. Fortunately, task management tools (like Jira, Azure DevOps) have almost all the required data.
4. Define which tool you want to use to visualize the metrics.
5. Connect your visualization tool with your data sources.
6. For qualitative criteria, fill out the self-evaluation checklist.
7. Sprint Retrospectives are great places to review the latest state of the metrics.
8. Interpret metrics together.
9. Metrics data are just a means to show your current status. They don't tell you how to improve them or what your ideal state should be for each metric.
10. You can use AI deeply to interpret the metric data and create insights.

AI & Performance Evaluation

Let’s check how AI can help with the Scrum Team performance evaluation.

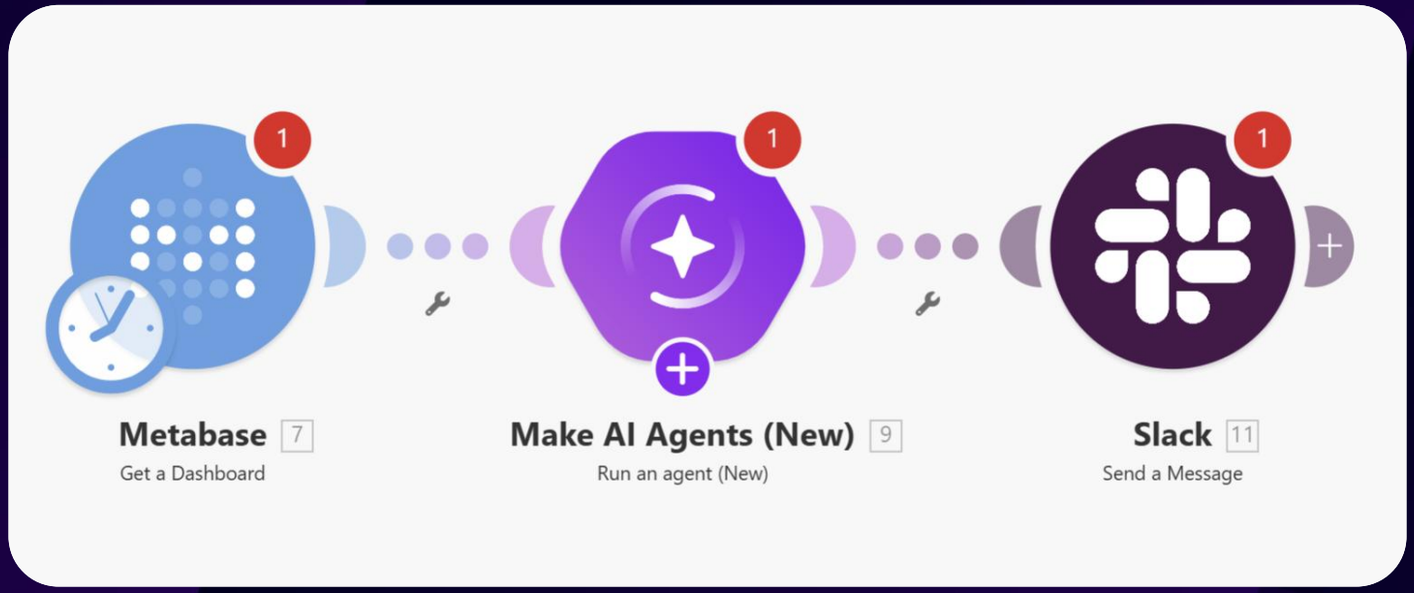


AI & Performance Evaluation

You can create AI Agents to help the team as an assistant to have a great interpretation of metrics, plus give you improvement suggestions. See the AI Agent below.

*** You can use Make.com, n8n, Zapier, etc. to create your AI Agents.

My recommendation is Make.com, which is simple and user-friendly.



This AI Agent connects with Metabase to get the data, interpret metrics, and adds the interpretation into a Slack channel that you already created for this purpose. You can use this AI Agent as your assistant in your Sprint Retrospective.

If you want to learn how to leverage AI in product delivery with Scrum, join one of my PSM-AI Essentials or PSPO-AI Essentials classes.

Professional Scrum Master™ – AI Essentials

Now, it is decision time. If you want to go further and become a true AI-aware Scrum Master, join my upcoming Scrum.org official PSM-AI Essentials class.

Enroll Now

Professional Scrum Product Owner™ – AI Essentials

Now, it is decision time. If you want to go further and become a true AI-aware Product Pro, join my upcoming Scrum.org official PSPO-AI Essentials class.

Enroll Now