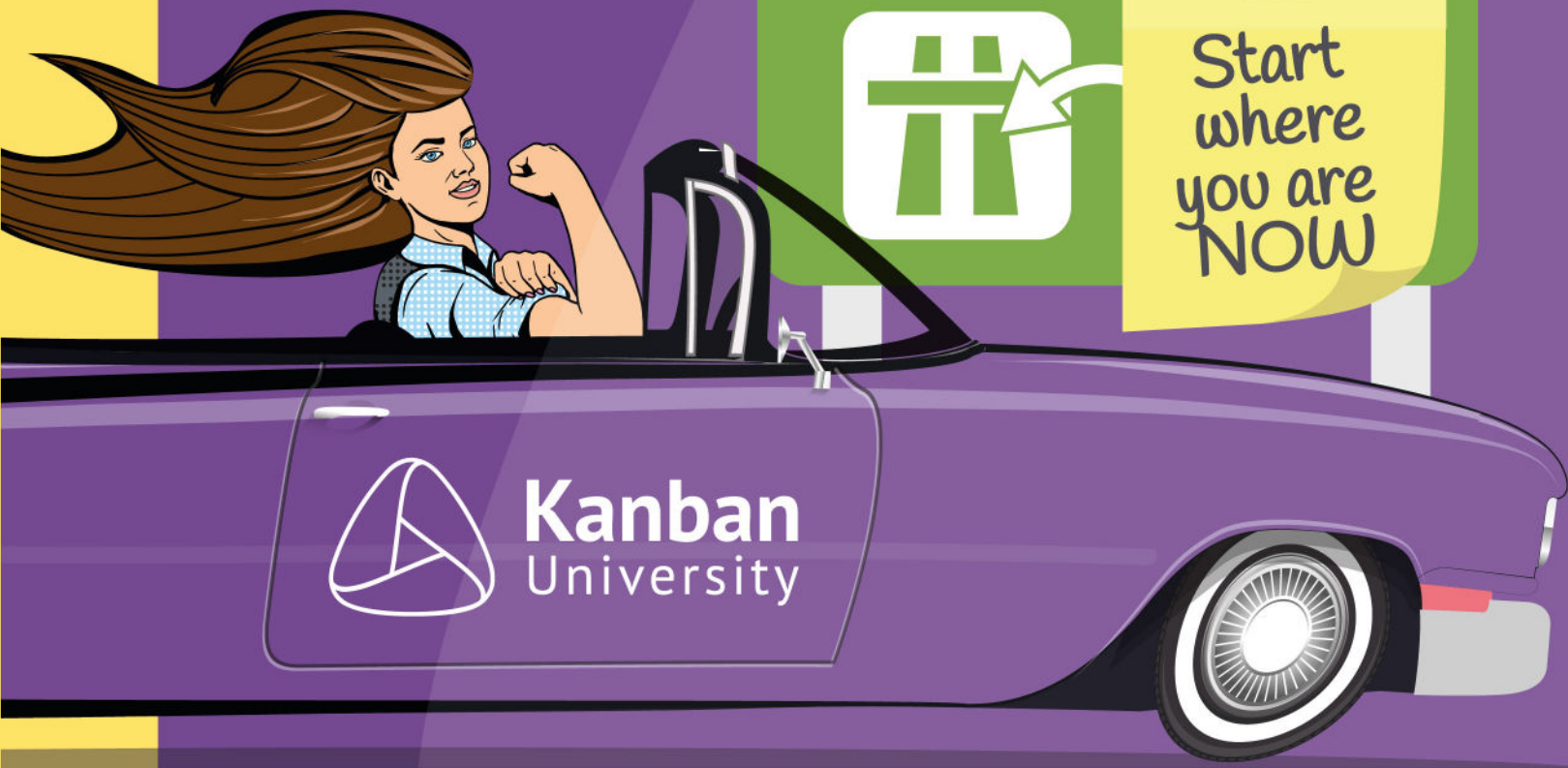


#YesWeKanban

THE OFFICIAL GUIDE TO THE KANBAN METHOD

V.2
DEC'22



Index

The Kanban Method	3
What is Kanban?	3
Method, Methodology, or Framework?	4
Origins	4
Areas of Application	4
Principles and Practices of the Kanban Method	5
Kanban Principles	5
Change Management Principles	5
Service Delivery Principles	5
Kanban General Practices	6
Visualize	6
Limit Work in Progress (WIP)	6
Manage Flow	6
Make Policies Explicit	6
Implement Feedback Loops	7
Improve Collaboratively, Evolve Experimentally	7
“Kan-Bahn” – An Introductory Metaphor	8
Utilization vs. Throughput	9
Types of Work	9
Classes of Service	9
Managing the Flow of Work	9
Visualize	9
Limit Parallel Work	9
Pull	10
Flow of Work	10
Blockers	10
Explicit Policies	10
Feedback Loops	10
Improve the System	10
Options, Commitment Point, Lead Time	10
Specific Practices	11
STATIK	11
Kanban Boards	12
WIP Limits and Pull	12
Core Kanban Metrics	13
Kanban Cadences	14



The Kanban Method

This guide is targeted at people new to Kanban and interested in learning about the basics of the method. That is why we included an introductory metaphor (KanBahn) to help people connect to the concept. Our hope is that this guide provides an easy entry into the vast Kanban body of knowledge.

For alumni of Kanban University classes for instance who want to review certain aspects, we recommend the “Essential Kanban Condensed” e-book as a reference.

What is Kanban?

Maybe the simplest way to put it is: With Kanban, you can manage work. It is a method to manage all types of professional services, also referred to as knowledge work.

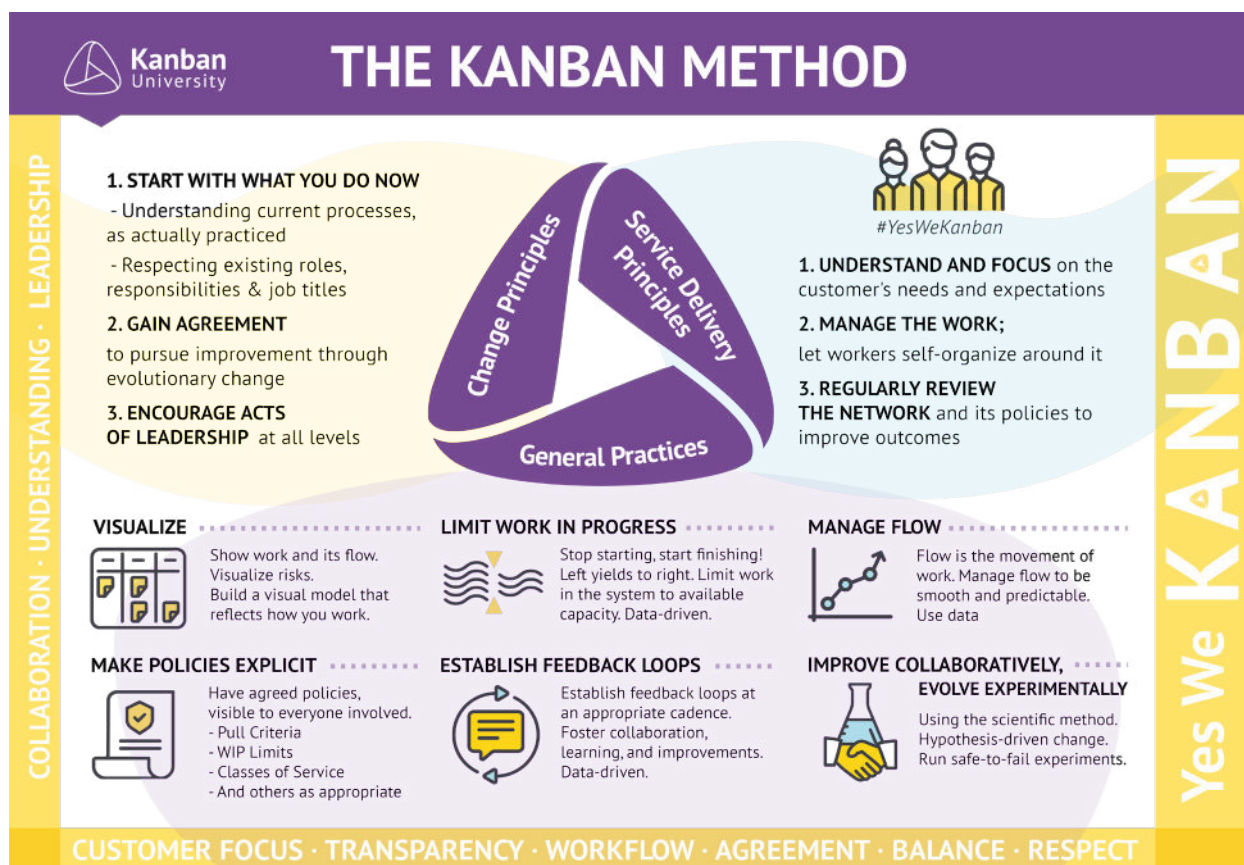
Using the Kanban method means applying a holistic way of thinking about your services with a focus on improving them from your customers’ perspective.

With the Kanban Method, you **visualize** invisible knowledge work and how it moves through a workflow. This helps to effectively operate your business, including understanding and managing risks in delivering your services to the customers. With Kanban, you and your business will develop an adaptive **capability** over time to respond better and faster to changes in

your customers’ needs and expectations or within your business environment.

Kanban is widely known for usage within teams, to relieve **overburdening** and to (re-)gain control over the work done by the team. While this usually brings quick benefits, applying the Kanban Method at a greater scale, e.g., for a line of service usually encompassing the work of multiple teams or different parts of organizations, brings even greater opportunities. Used with a service focus in mind, Kanban is an effective organizational development tool.

Kanban University (www.kanban.university) is “Home” of the method and the global community of Kanban trainers, coaches and consultants who continue to evolve the method and develop its related body of knowledge.



Method, Methodology, or Framework?

Kanban is often confused with a methodology or framework. In software engineering a methodology is a process definition approach to software development and project management (a bit of a misnomer, as “methodology” should mean “the study of methods”). Methodologies contain prescriptive, defined workflows and processes, including roles, and responsibilities. This means that they are usually specific to a domain, such as software development.

A process framework on the other hand is an incomplete methodology – a set of scaffolding that is intended to have broader applicability but requires customization for each context in order to fill in the missing gaps.

Kanban is not a methodology nor a process framework. Rather, it is a management method or approach that should be applied to an existing process or way of working. There is never a question of using Kanban versus a given methodology or framework. Rather, it is always adding Kanban using an existing methodology, framework, or way of working. Kanban is intended to help you manage work better and to improve **service delivery** to the point where you consistently meet customer expectations. Kanban is a means to improve what and how you already do things. It is not a replacement for what you already do.

Origins

The Kanban Method as described here is based on [Kanban: Successful Evolutionary Change for Your Technology Business, by David J Anderson, 2010](#). The motivation to create the method was mainly to find a way to manage and improve professional service businesses as well as a way to provide a humane change method.

The roots of the method are found in Lean Manufacturing. However, Kanban is meant to be used to manage knowledge work resulting in intangible and virtual goods and services. When compared to manufacturing, the Kanban Method views inventory as usually intangible or invisible and has much lower direct costs attached, variability in the delivery of work is accepted as inherent, the workflow is usually less strict, and the focus on waste reduction is of lower concern. Improving the value and flow of goods and services delivered are the initial focus when using the Kanban Method.

In many aspects, Kanban is strongly founded on lean: The focus on the flow of work, limiting work in progress to establish pull systems, focus on the optimization of the system as a whole rather than managing an individuals' performance, making decisions based on data, and continually improving in an evolutionary way.

Areas of Application

Kanban is a rather abstract “method without methodology” and has a wide area of possible applications.

It is important to understand that the Kanban Method is applied with its principles and practices on top of an existing flow of work and way of working. The work can be of very different kinds. After being introduced in 2010, there were several examples of Kanban being applied to services in the IT sector. Today, there is an ever increasing amount of examples of Kanban being utilized by marketing agencies, human resources, media and design services, customer support, product development, and education.



Principles and Practices of the Kanban Method

When using Kanban, the scope of application (e.g., single team, multiple teams, departments, divisions, etc.) can influence the way the method's principles and practices are applied.

If you have a look at a basic scope within a team for instance, you might find a relatively simple Kanban board with maybe 5 columns indicating the workflow, a few simple metrics and diagrams, a daily held coordination meeting and regular reviews of the team's work and performance.

Now imagine a whole internal services department within an enterprise, which is managed by a set of Kanban boards relating to each other, placed at different levels of granularity, covering different workflows. The amount of work in progress is limited at different levels.

Both instances are proper usage of the Kanban method. There is no “right or wrong” in Kanban, rather more or less appropriate adoption of practices given the business context and cultural environment.

The following two sections describe the general Kanban principles and practices.

Kanban Principles

Change Management Principles

These Change Management Principles are common to all Kanban implementations:

- Start with what you do now
- Agree to pursue improvement through *evolutionary change*
- Encourage acts of leadership at all levels

Kanban is not a big bang transformation going from a current state to some future state. We know from history that those rarely work. Instead, Kanban uses an evolutionary change approach, building on the way of working already in place, seeking to improve it using many forms of feedback and collaboration. The Kanban Method engenders evolutionary change through insights gained by the people working with the Kanban board and taking acts of leadership to continuously improve their way of working. These acts of leadership may not be what are thought of as traditional leadership. They may be small observations and suggestions for improvement by individuals without organizational leadership roles.

Service Delivery Principles

Kanban encourages you to take a service-oriented approach to understanding your organization and how work flows through it. This service-oriented organizational paradigm is based on the idea that your organization is an organic entity consisting of a network of services, each of them living and breathing, and evolving. Customer requests flow through this network of services. If we are to improve service delivery, improvements should be guided by a set of principles. These principles may not be utilized early on by organizations as they may not have developed or evolved a service-oriented or customer service mindset as part of their culture.

The service-oriented principles are:

- Understand and focus on customer needs and expectations
- Manage the work; let people self-organize around it
- Regularly review the network of services and its policies in order to improve outcomes

YES WE KANBAN



Manage with Competence.

Lead with Confidence.

#YesWeKanban

Kanban General Practices

As mentioned earlier, the breadth and depth of Kanban practices applied varies greatly.

In this section, the six general Kanban practices are described. Later in the guide we go into more details on some of the core specific practices that fall within these 6 general practices. Please refer to the ***Kanban Maturity Model (KMM)*** for more details on specific implementation by maturity level.

Visualize

Show work and its flow.
Visualize risks.
Build a visual model that reflects
how you actually work.



Visualize

A good visualization is the key to effective collaboration and to identify improvement opportunities. Many times, work in the organization is hidden. Visualizing that work and the flow of that work greatly improves transparency. The human sense of vision is very old from an evolutionary point of view. It allows us to absorb and process a great deal of information in a short time. In addition, visualization supports cooperation, as everyone involved literally has the same picture. More details on Visualization will be presented in the section on Kanban Boards.

Limit Work in Progress (WIP)

WIP (Work in Progress) states the number of work items in progress at a certain time. Through Kanban we have discovered that effective systems focus more on flow of work and less on worker utilization. When resources are fully utilized there is no slack in the system and the result is very poor flow, just as in rush hour on the freeway. In knowledge work we also have the issue of context switching that can drastically reduce the effectiveness of workers.

In Kanban, we limit the WIP to balance utilization and still ensure the flow of work. Later we will describe WIP limits and how they are used in a “pull system”.

Manage Flow

The goal of managing the flow of work is to complete work as smoothly and predictably as possible, while maintaining a sustainable pace. As mentioned before, limiting WIP is one of the key ways that helps us ensure smooth and predictable flow. The monitoring or measuring of the ***workflow*** results in important information that is very useful for managing expectations with customers, for forecasting, and for improvements. This will be discussed in the section on Core Kanban Metrics.

Make Policies Explicit

Every day, countless decisions are made about the organization of work, either by individuals or between groups of people.

Imagine a new employee starting in your area. Ideally, she will quickly understand how work is organized through explicit policies. These include:

- Policies such as ***replenishing*** the board (when, how much, by whom)
- Definition of when a work ***activity*** is completed, and the work item can move on (“pull criteria”)
- WIP Limits
- Policies for handling work items of different classes of service
- Meeting times and content
- Other principles and collaboration agreements

All of these policies should be agreed to jointly between all parties involved including customers, stakeholders, and employees responsible for work on the board. The policies should be placed in a clearly noticeable area, preferably right next to the board. At team level, a team agreement is a good way to introduce policies. Like all other building blocks of the system, it is necessary to check and adapt these regularly.

Please note that policies are not like work instructions freeing people from the burden to make meaningful decisions. Rather, policies should enable self-organization within the group of people running a Kanban system.

Policies should be:

- sparse
- simple
- well-defined
- visible
- always applied
- readily changeable by those providing the service

Implement Feedback Loops

Feedback loops are required for a coordinated delivery and for improving the delivery of your service. A functioning set of feedback loops appropriate for the given context strengthens the learning capabilities of the organization and its evolution by means of managed experiments.

Some commonly used means for feedback loops in Kanban systems are the board, metrics, and a set of regular meetings and reviews which are referred to as cadences.

Improve Collaboratively, Evolve Experimentally

Back to the Change Management Principles, in the Kanban Method we “Start with what you do now” and “Agree to pursue improvement through evolutionary change.” Kanban is a method for continuous change, and we make those changes collaboratively using designed **experiments** based on models and the **scientific method**. This is where feedback and metrics are so important to guide us in the evolutionary path. We design safe-to-fail experiments so that if our hypothesis is correct and the experiment gives good results, we keep the change, but if the results are not positive, we can easily roll back to the prior state.



“Kan-Bahn” – An Introductory Metaphor

The basic concepts of Kanban will be introduced here by means of a metaphor. Before we start, please consider George E.P. Box’s famous [quote](#) “All models are approximations. Essentially, all models are wrong, but some are useful. However, the approximate nature of the model must always be borne in mind”. An international group of Kanban coaches and trainers created this metaphor in 2016 at a Kanban Leadership Retreat in Barcelona.

It is based on an “Autobahn”, a German form of highway, hence the name. Our board (or system) is represented by a highway. Traffic (the work) flows – divided into packages – in the form of different vehicles through our system, a defined section of the route.



Utilization vs. Throughput

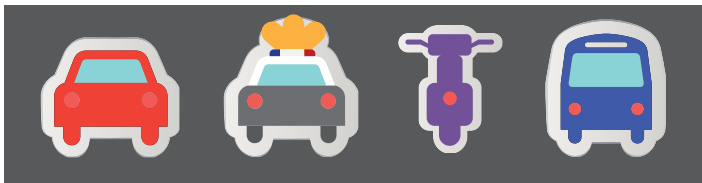
When traffic jams are on the motorway, the roads (resources or **capacity**) of our system are fully utilized (utilization), but very little is moving: Very few vehicles (**work items**) per unit of time pass through the system (**throughput**), all spend a very long time (**lead time**) in this section of the route. Consequently, we are late (**delays occur**), and we miss our appointments (**promises to deliver** might be broken).

Are you really looking forward to high utilization on the road when driving? Unfortunately, this form of optimization is still a widespread management paradigm.

With Kanban, we optimize differently. As many vehicles (**work items**) as possible should be able to pass through our system in a smooth manner, as quickly and predictably as possible. Operating well below full capacity (slack) is desired here and conducive to the flow.

Types of Work

Different types of vehicles from motorcycles to cars, minibuses, trucks and buses pass through the section of the route. The Kanban equivalent are different types of work (**work item types**). These have different characteristics – they vary in their purpose, size, speed, and passenger or cargo capacity.



Classes of Service

Different types of vehicles such as police cars, fire trucks or ambulance cars may pass through the system preferentially. This is an example for treating defined items in a differentiated manner. In Kanban, this concept is called “**Class of Service**”.

The example described above could be mapped to a service class typically called “expedite”. For this purpose, there are agreed rules and criteria for vehicles known to all drivers that are allowed to use this service class: vehicles must be clearly recognizable (e.g., by blue light and certain paintwork) and may pass through the system even if the **WIP limit** is fully exhausted (motorway congested), while others have to form a rescue lane. This will lead to the “**expedite**” vehicles being able to pass quicker, while the travel for the other cars will take longer.

Another example for the use of Classes of Service are restricted traffic lanes that are exclusively reserved to e.g., only buses and taxis, electric cars or vehicles with two or more occupants (“carpool lanes”, also known as High Occupancy Vehicle or “HOV” lanes in the United States).

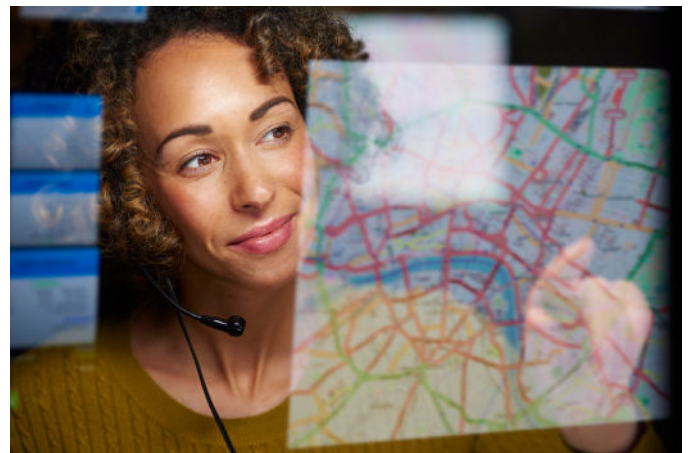
Managing the Flow of Work

Depending on the location and time, the volume of traffic varies, i.e. the total number of vehicles (**work items**) and the distribution of vehicle types (**work types**). In metropolitan regions there will usually be more personal or private traffic with extremely high volumes during peak times (e.g., rush hour). Conversely, on major transit routes between metropolitan regions, there will be less extreme peaks in volume consumed mostly by shipping trucks.

Our system is being designed to cope with variability of traffic volume. In doing so, we might control the inflow of vehicles (work items), the available capacities (e.g., number of lanes and their quality of expansion), and the speed limit.

Visualize

Imagine working in a traffic control center. Due to the complexity of the system and the variability in the behavior of each vehicle and unpredictable events, each day will be different.



In the picture above, a control board (**kanban board**) is used by a traffic dispatcher to see at a glance which sections of the route are busy, where there are construction sites, and where there have been accidents or breakdowns causing congestion (**bottlenecks**). This presentation allows decisions to be made more quickly and collaboratively.

Limit Parallel Work

In urban centers, traffic lights are often found at motorway entrances (on-ramps). Ramp metering or signaling, as it is called, controls (see [Ramp Metering: A Proven, Cost-Effective Operational Strategy](#)) the rate at which vehicles can enter the system based on traffic volume and speed in order to avoid overloading.

The Kanban term for this is “Limiting WIP”, where WIP stands for **Work in Progress**.

Pull

When you are driving on the highway, you can see if there is space in front of you. You consider this as a signal to continue, otherwise you must slow down or even stop. In Kanban systems, we call these signals of available capacity **pull signals**. For pull signals to work, you will need to define your WIP limits as an expression of maximum capacity.

Applied to a motorway might look like this: The **system**, our section of the motorway where we are driving our car, would be divided into sectors (e.g., 500m or ~500yds). If there is enough space for your vehicle plus a safety distance on the following sector (i.e., fewer vehicles there than the maximum capacity = WIP limit), something will signal your vehicle (**work item**) to proceed to the following sector, otherwise you will wait at the end of the current sector until capacity becomes available (through other vehicles leaving the sector).

Well, every metaphor has its limits. This signal will populate further up the road, possibly preventing more cars from entering the motorway.

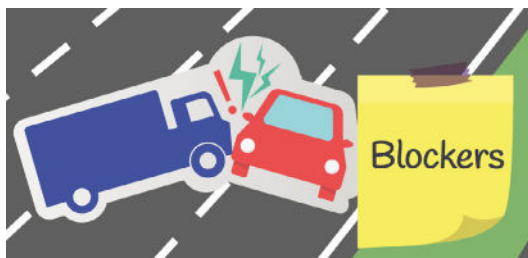
Flow of Work

In the Kanban context, flow refers to the movement of work through a system. Traffic flow is actively controlled on particularly busy sections of the motorway. This requires visualization and measurement data logging and evaluation. This data is collected by sensors for traffic volume and speed, weather conditions etc. In addition to controlling the inflow of vehicles, there are electronic scoreboards that reduce or throttle the speed depending on the traffic situation in order to enable all road users to pass as quickly and evenly as possible.

Over time, a lot can be learned about patterns in the flow by evaluating the historic data gathered. It can be used to further optimize the system, informing authorities where changes would have the biggest effects.

Blockers

Reported accidents or road damage (**blockers**) obstruct the flow and are displayed in the control center and removed as soon as possible. The system is regularly examined for accident hotspots in order to enable future improvements.



Explicit Policies

The signs and signaling systems along the motorway make the rules of traffic (which are known to all road users) visible and are usually followed.

Feedback Loops

On particularly important roads, such as access roads to airports or city centers, there are information boards indicating the estimated time to travel to certain destinations. E.g., “10 Min to Airport”. This data is based on historical data as well as current traffic volume.

Map providers like Google Maps use a combination of real-time data and historic patterns to both navigate you best on your journey (**manage flow**), and to help you plan trips ahead by forecasts.



Improve the System

A motorway system also needs to be continuously developed and improved. Traffic flow measures are optimized, existing routes need to be serviced, potholes repaired, **bottlenecks** and accident hotspots defused. New lanes may be built (**capacity expanded**) in particularly busy sections, which is very costly and time-consuming. All these improvement measures are informed by knowledge of the system, especially supported by visualization and collecting data, and regularly checked for their effectiveness after they are introduced.

Options, Commitment Point, Lead Time

A roundabout at the motorway entrance for instance allows you to enter the motorway. Only when you take the turn to the on-ramp will you exercise this **option**. You commit to traveling on the motorway (discarding other options). If you already see a long traffic jam from a distance, you can also discard the motorway option and, for example, choose a different route or postpone the trip. So how do you build your own Kanban system? Let's learn about some of the **Kanban Method's** specific practices.

Once you decided to enter the motorway, you are “in the system” and the lead time clock starts ticking. Depending on the available capacity, you can now pass through the individual parts of the highway section. Arriving at the end, the **lead time** ends, indicating how long it took you from the entrance to the exit point.

Specific Practices

So how do you build your own Kanban system? Let's learn about some of the Kanban Method's specific practices.

STATIK

A question commonly asked by practitioners is "If every board and Kanban system is unique, how can I design my own system?"

The Systems Thinking Approach To Introducing Kanban (STATIK) is a repeatable and humane way to get started with Kanban. It has been applied numerous times in practice.

The STATIK approach should be applied to each service. It will result in a full Kanban system design. Throughout the process, systems thinking should be applied. The (future) system is always considered as a whole, with the goal of improving the flow of value to **customers**.

The illustration below (Figure 1) summarizes the 6 basic steps in the STATIK approach, which are usually applied in an iterative way. Subsequent steps can uncover new information, and it might make sense to repeat earlier steps.

STATIK workshops tend to iteratively explore the correct system design. STATIK is not intended as a one-pass, sequential process, but rather it is intended to function as a **feedback loop** that informs design and redesign activities.

In practice, this process usually takes between 4 hours and 4 days. It is important to understand that it should be done with at least a representative group of the people involved. While everyone will have a picture of how the work is done in mind, it rarely maps between people. The STATIK approach will reconcile these views into a shared view. As a rule of thumb, it should not be done in isolation e.g., by the Project Manager, Team Lead, or a Coach or Consultant.

1. **Identify sources of dissatisfaction** – What are the people involved in service delivery dissatisfied with? What are customers dissatisfied with? All these sources of dissatisfaction provide motivation for change which is key for a successful Kanban initiative.
2. **Analyze demand** – What do customers request, through which channels? What are types of work and patterns in **demand**? This information is key to develop the full picture of the work that arrives at the system. Remember, manage the work not workers!
3. **Analyze system capabilities** – What are the capabilities of the system with regard to how much of the customer demand is being delivered, of what type, and how fast and predictable? This step typically requires historical data.
4. **Model the workflow** – Which are the activities that each of the identified work item types go through? They might be sequential, parallel, or in no particular order. Later, these will be the basis for defining the columns on the Kanban board.
5. **Identify classes of service** – How do items enter and get treated in the system? See the Classes of Services definition.
6. **Design the Kanban system** – Based on all the insights gained in the prior steps, the Kanban system is then designed. A Kanban system naturally consists of a board and tickets, plus other important elements such as metrics, **cadences**, and policies.

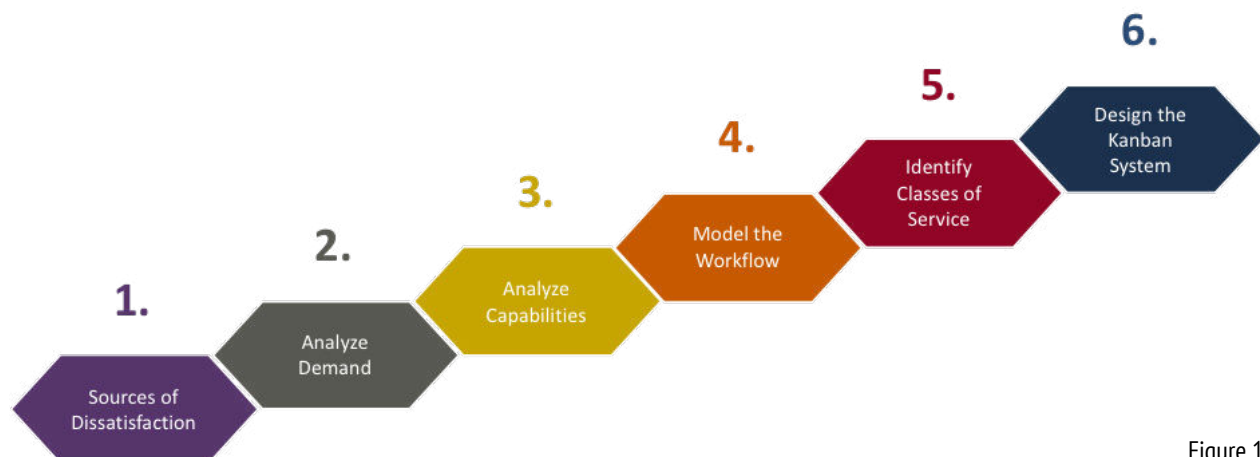


Figure 1

Further details about STATIK are taught in the Kanban University System Design courses.

Kanban Boards

Kanban boards are the most common means of visualizing a Kanban system. Common to all boards is pulling work from left to right through the board: on the left, new work items enter the board. When they exit on the right, value is delivered to customers.

In a Kanban system, there is at least one clear commitment and delivery point as well as a representation of the permitted amount of work (Work in progress, WIP).

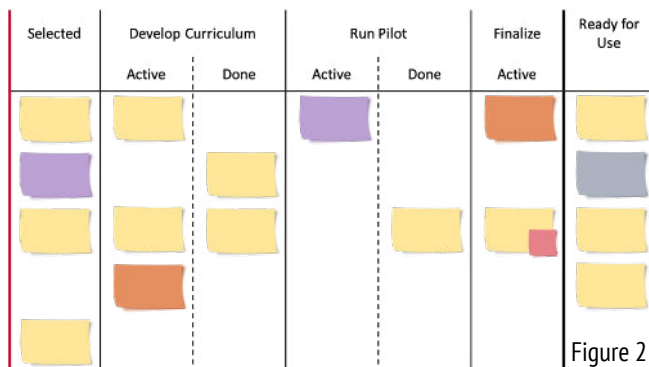
Work items can be of different types and sizes, from tasks to requirements, types of artifacts, (groups of) product features and topics to projects or product packages on higher level boards. Examples are campaigns in agencies, user stories in software development teams, job positions in HR, or products for a product development group.

Work items are typically displayed on individual (paper) notes, which are usually called **cards** or tickets.

The series of activities these work items go through are referred to as workflow. Kanban is based on the “Start where you are now” approach, so, the actual workflow (not a wishful future image) is being modeled on the Kanban board.

The individual steps in the workflow and buffers are shown in columns. Lanes are often used for different work types, projects, etc. to distribute capacity.

Imagine the work of an in-house training service provider in a larger company. Ideas or requirements for new courses are collected first. After a selection and refinement process, new courses are developed, piloted, then finalized and are then ready for use. The image below (Figure 2) shows a possible, simplified board layout:



The workflow is modeled on the board. Different colored notes can be used, for example, to represent different types of courses (e.g., online vs. classroom training), or different groups of clients.

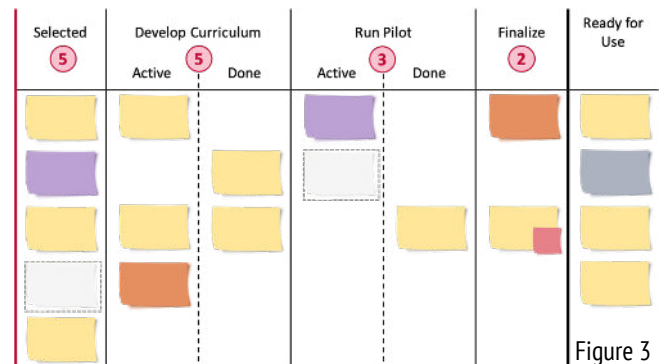
The flow of work and its risks should be realistically shown in their true current state rather than a wishful image of the future at all times. Your Kanban board should reflect your specific workflow, which is usually more than columns

labeled as To Do, Doing, Done. The possibilities vary greatly. Each Kanban system and Kanban board are unique.

WIP Limits and Pull

The so-called WIP limit, i.e., the maximum number of work items allowed at a time, can be defined per work state(s), per person, per lane, per type of work, for a whole Kanban system, etc.

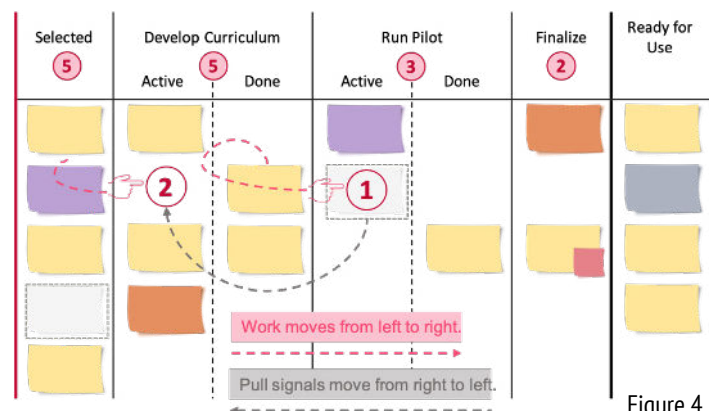
WIP limits are typically represented by a number in a circle, above the respective columns:



In Figure 3, a maximum of three courses may be piloted at the same time. Furthermore, the design of the system is such that both Active and Done columns are limited by a total WIP limit. Currently, there is a purple item in the Active column, a beige item in the Done column and there is capacity for another course, indicated by the grey dashed note (slot).

Limiting the work that is allowed to enter the system is an important key to reducing *delay* and context switching which may result in poor timeliness, quality, and potentially waste. The aim is to create a balance between demand and *capability* over time.

Limiting the work that is allowed to enter the system also creates a continuous flow of work in which drawing or “pulling” work only happens if there is *capacity*. A virtual pull signal is generated when the WIP limit is not fully utilized. While work on the board moves to the right, pull signals move to the left, further upstream (Figure 4).



A pull system is an important distinguishing point from traditional project management, where work items are scheduled based on deterministic planning (push). In pull systems, completed work is regarded as more valuable than starting new work. This is often a cultural change. “Stop starting, start finishing” is a good mantra to remember for starters!

WIP limits are one specific example of a policy in Kanban. For more information, please refer to the **“Make Policies Explicit”** section under **Kanban General Practices** of this guide. They need to be agreed to by everyone actively involved. WIP limits serve as an enabling constraint, which gives focus and develops behaviors such as collaboration and finishing started items with high quality. WIP limits are key to establishing a pull system.

Core Kanban Metrics

There are a number of basic **metrics** in Kanban:

- Lead time is the time it takes for a single work item to pass through the system from the start (**commitment point**) to completion
- **Delivery rate** is the number of completed work items per unit of time, such as features per week, training classes per months, or new hires per month
- WIP (work in progress) is the amount of work items in the system (or a defined part of it) at a certain point in time

These core metrics are used in various graphical representations to understand system behavior and identify opportunities for improvement.

Figure 5 represents a **run chart**. The lead times of completed work items are plotted sequentially on a timeline. This is useful to observe lead time trends:

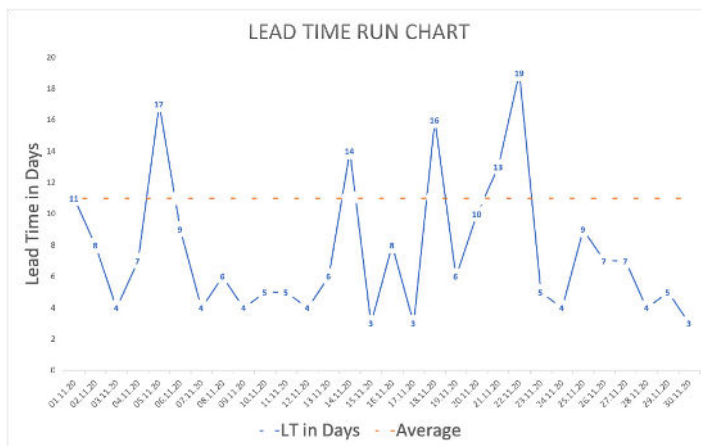


Figure 5

Figure 6 shows the distribution of lead times:

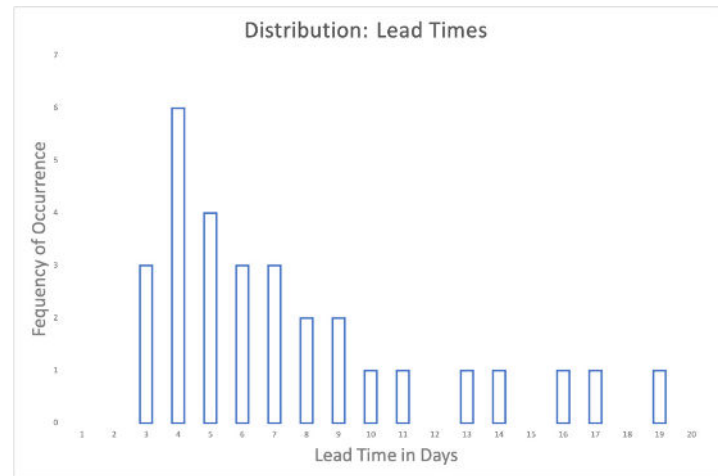


Figure 6

This chart depicts the range of observed lead times (min and max) and their frequency of occurrence (how often). The purpose of managing flow would be to optimize this distribution: by narrowing down the range as much as possible (predictability) and shifting it to the left (timeliness).

Figure 7 represents a **cumulative flow diagram** (CFD). The CFD contains useful information regarding the flow of work across multiple activities. The colored areas in the diagram represent the number of work items within a particular activity in the workflow and how these work items move across all activities, from top to bottom, over time until done.

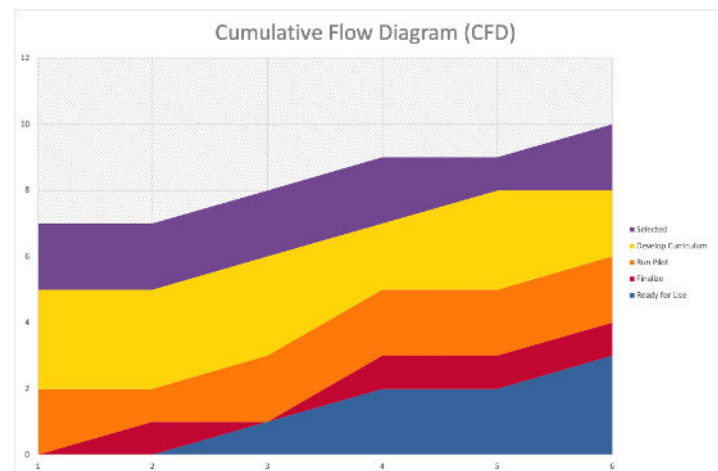


Figure 7

Kanban Cadences

While in early Kanban implementations feedback loops might be almost completely lacking, with growing maturity, feedback loops evolve, which in turn advances maturity. We encourage you to build up your cadences gradually.

Please note that like all elements of a Kanban implementation, the cadences can and should be set up to fit within the given organizational context. In practical terms, this means:

- Identify existing meetings and **reviews** that already serve a similar purpose and continually evolve them
- Keep the existing names or use the standard cadence naming or come up with something else. It is the purpose that matters
- Pick the frequency and duration based on your context. In many cases, having more frequent but shorter meetings over time increases agility

As a side effect of many Kanban initiatives, we observe better focused, structured and tightly managed regular meetings with fewer attendees.

Example: Figure 8 Cadences at team level

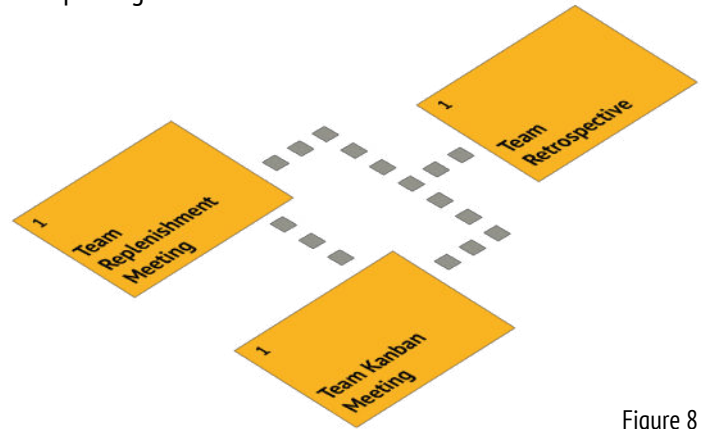


Figure 8

Example: Figure 9 Service-oriented Cadence

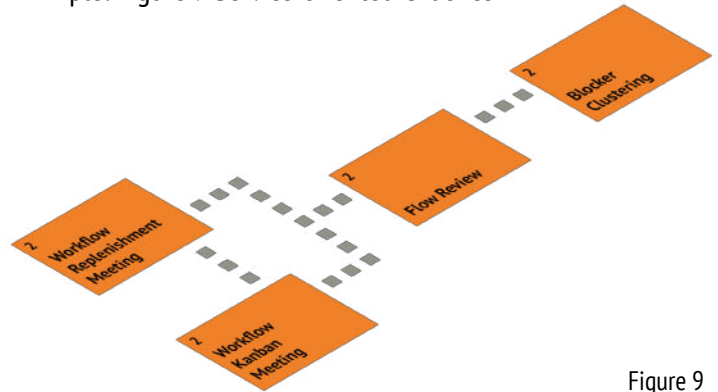


Figure 9

Cadence	Example frequency	Purpose
Team Kanban Meeting	Daily	To observe and track the status and flow of work (not the workers). How can we deliver the work items in the system quickly? Has capacity become available? What should we pull next?
Team Retrospective	Biweekly or Monthly	Reflect on how the team manages their work and how they can improve.
Internal Team Replenishment Meeting	Weekly or as needed	Select items from the pool of work to do next.



Kanban University

About Kanban University

Kanban University works to assure the highest quality coaching and certified training in Kanban for knowledge work and service work worldwide. Our Accredited Kanban Trainers, Accredited Kanban Consultants, and Kanban Coaching Professionals follow the Kanban Method for evolutionary organizational change.

Kanban University offers accreditation for Kanban trainers, a professional designation for Kanban coaches, and certification for Kanban practitioners.

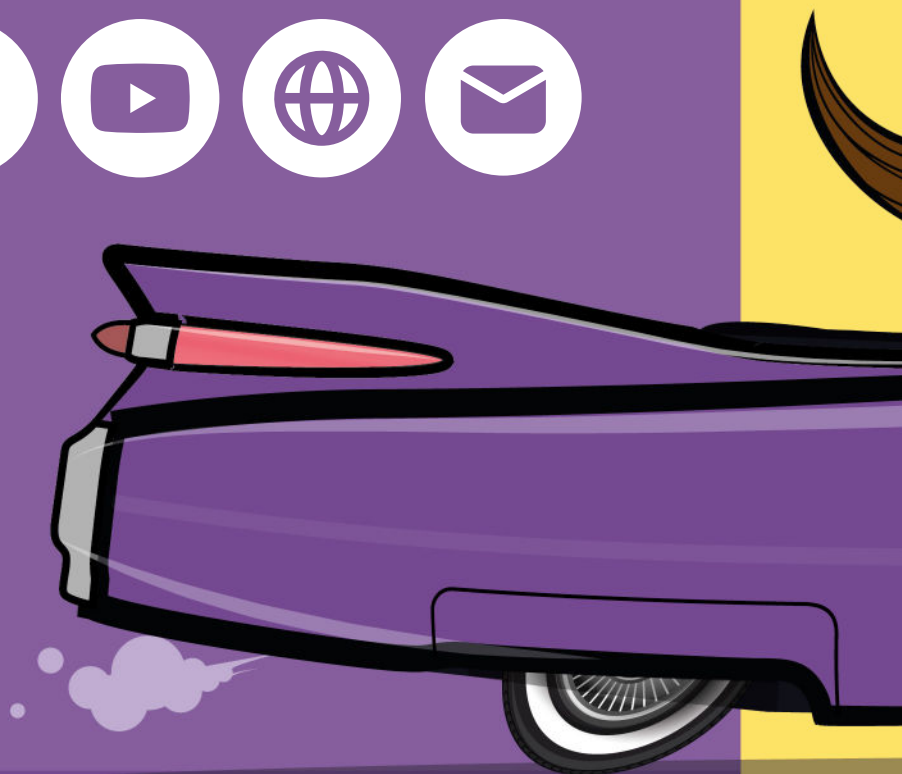
Acknowledgements

We would like to express a special thank you to Susanne and Andreas Bartel of Flow.Hamburg for putting together The Official Guide to The Kanban Method with the collaboration of the Kanban University team.

Another big thank you to the following participants in the creation of the Kan-Bahn during one of the Kanban Leadership Retreats in Barcelona: David Lowe, Jose Casal, Martin Hoppen, Susanne Bartel, Andy Carmichael, Teodora Bozheva, Ruben Olsen, and Ward Schwillens. We are grateful to have you part of the Kanban University community.



#YesWeKanban



The Kanban Method Glossary

Activity

Actions taken in the service delivery workflow that progress a work item to the next step of the **knowledge discovery**. One or more activities can be visualized as a column in a **kanban board**.

Blocker

Something that is impeding the flow of a work item. The blockage could be partial or complete.

Bottleneck

A constrained activity that limits the flow and possible delivery rate of the entire workflow.

Cadences

A type of meeting or review providing feedback from one or more services for the purpose of reviewing, coordinating, or improving the work being delivered.

Capability

A measurement of system performance. Measurements may include lead time and throughput. For more information on measurements, please see **metrics**.

Capacity

Represents the amount of work a kanban system can hold while work is still flowing efficiently. Also used for the allocation to a particular **activity**, work item type, or class of service.

Card

A visual representation of a **work item**. A card may also be referred to as a ticket.

Change Management Principles

The **Kanban Method** approaches **evolutionary change** using the following three principles.

- Start with what you do now
- Gain agreement to pursue improvement through evolutionary change
- Encourage acts of leadership at all levels

Class of Service

A specific level of service applied to the treatment of a **work item** established through a defined set of policies. A Class of Service can apply to one or more **work item types**. A class of service will often set a **service level expectation**. The choice of class of service may reflect a relative value, risk, or **cost of delay**. The four common archetypes of class of service that are widely recognized are:

- **Expedite**
- **Fixed date**
- **Intangible**
- **Standard**

Commitment Point

The point at which the decision is made to start activities to deliver a **work item**. Before this point, work done supports the decision whether to deliver the work item (see **Upstream**).

Control Chart

A chart, usually a **run chart**, showing control ranges outside of which a process may be considered “out of control” in some specific sense.

Cost of Delay

The rate at which the expected value of a product, initiative, or **work item** decays as its delivery is delayed. Cost of delay involves both urgency and impact. Cost of delay may be used to inform time-related decisions, including the ordering of work items during **replenishment** or the assignment of a **class of service**.

Cumulative Flow Diagram

A chart showing the cumulative number of arrivals and departures of **work items** from each step in the workflow over a time period. The diagram is a quick visual indicator of the stability and characteristics of the workflow in terms of amount of work in progress in the different stages. In general, parallel lines between the steps indicate that the inflow is consistent with the outflow which is good, and deviation is something to be investigated.

Customer

The person, or more likely group of people making the service request and accepting its delivery. Customers can be internal or external to an organization: When one service makes a request to another service within the same organization, the requesting service can be considered an internal customer of the requested service.

Customer Lead Time

Customer lead time is the time between receiving the customer's request and delivering on it.

Delay

In knowledge work there are a number of potential sources of delay. A **work item** could be waiting for **capacity** or it could be **blocked** due to a defect or a **dependency**. There are also delays due to switching between different work items in progress, e.g., when a work item has to give way to another work item with higher urgency.

Delivery Rate

See **throughput**.

Demand

The work requested by a service's customers. Kanban systems seek to bring demand and **capability** into balance.

Evolutionary Change

In the Kanban Method, managed evolutionary change is the intentional design of small incremental changes, tested via **experiments** using the **scientific method**. The current way of working serves as the starting point, with the goal of being fitter for purpose in a continuously changing environment. Also known as **Evolve Experimentally**.

Evolve Experimentally

One of the 6 general practices in the Kanban Method. See evolutionary change.

Expedite

An archetype of a class of service that is applied to work items where the impact of delay is both high and immediate.

Experiment

An intentional design to determine if a hypothesis is valid or not. Observations and evidence from the experiment will be used to determine the validity of the hypothesis and what further experiments should be planned. See: **Scientific Method**

Explicit Policies

An explicit description of various agreements that shape how the service delivery works. Policies can reflect how work items move from one activity to another or can include how a team, individuals, and services interact with each other.

Failure Demand

Demand generated because of previous poor-quality deliverables or demand that never should have been received. Examples can include poor requirements or defects during delivery requiring rework.

Feedback Loop

One of the **6 general practices** described as Establish Feedback Loops in the **Kanban Method**. A feedback loop starts with feedback information about the existing **system** and then requires action based on whether the result is in line with the desired goals or must be altered in order to improve or align with those goals. Feedback information without action is not a feedback loop.

Fixed Date

An archetype of a class of service that is applied to work items where the impact of delay occurs at a specific date.

Hypothesis

An idea, based on observations, that lends itself to an **experiment** using the **scientific method** to determine if the idea is valid.

Intangible

An archetype of a class of service that is applied to work items where the impact of delay is unknown.

Kanban Board

A kanban board is a visual display of **cards** that represent the work items in a kanban system. Typically, boards are arranged in vertical columns that represent **activities**. Some boards use horizontal swimlanes to further enhance the visualization of policies, types of work, classes or service, or other attribute that is important for the management of the work. Additional dimensions may be represented by color or other card attributes. Cards move rightward from column to column as the work items they represent progress through the system. **Work in progress limits** and other **policies** may be represented visually.

Kanban General Practices

The six general practices of the Kanban Method are:

- **Visualize**
- **Limit WIP**
- **Manage Flow**
- **Make Policies Explicit**
- **Use Feedback Loops**
- **Evolve Experimentally**

Kanban Maturity Model

A model describing common patterns of evolutionary change within organizations. The model can serve as a map to organizational agility, resilience & reinvention. See: [KMM.plus](https://kanban.university/partners/)

Kanban Meeting

A meeting to help the delivery teams coordinate and manage the flow of work. The meeting is held at the cadence appropriate for the team, usually somewhere between daily and weekly. A common approach is to “walk the board” meaning to look at the work on the Kanban board, going from right to left, and ask the question “what is keeping this item from flowing?” and “who can help?”

Kanban Method

A method for the definition, management and improvement of services that deliver knowledge work.

Kanban System

A model of a kanban **service** delivery **workflow**. A kanban system is designed using **STATIK**. A kanban system contains the board, work items represented as cards, policies, metrics, and cadences.

Knowledge Work

Developing goods and services through activities that add to **knowledge discovery**. Examples of Knowledge Work are Marketing, Software Development, all kinds of Product Development. Knowledge work is performed by knowledge workers.

Lead Time

The difference in time (e.g., minutes, hours, days, etc.) between one point in the workflow to another point in the workflow. See **Customer Lead Time** and **System Lead Time**.

Lead Time Distribution

A chart showing the frequency of the observed lead time of work items. Lead times should be consistent, ie., all **system lead times** or all **customer lead times**. Different work item types or different classes of service may have different lead time distributions. The distribution of the lead time can be an indicator of the predictability of the **system**.

Limit WIP (Work in Progress)

One of the **6 general practices** in the **Kanban Method**. We want to limit work in progress to enable a **pull system** to improve predictability and flow. When work in progress is not limited, the system can frequently become **overburdened** resulting in poor **throughput**, predictability, and quality. Work in progress is commonly limited by **work in progress limit** policies.

Make Policies Explicit

One of the 6 general practices in the Kanban Method. The practice of defining agreements on how the service delivery will be handled. Examples of policies can include agreements on replenishment criteria, pull criteria, class of service, service level expectations, work in progress limits, and completion criteria. Policies are commonly visualized on or near the kanban board. They can also be stored and shared through various digital tools. For a complete list of tools, please visit <https://kanban.university/partners/>

Manage Flow

One of the 6 general practices in the Kanban Method. The point of a workflow is to result in delivery. We focus on managing the flow of the work in order to get smooth and predictable and potentially faster and efficient delivery. To this extent we manage the work, not the workers. Workers are encouraged to self-organize to improve the delivery of the work.

Meeting

A type of feedback loop designed to focus on managing the work. Examples include:

- Team replenishment meeting
- Team kanban meeting
- Team retrospective

Metrics

Metrics are a means of feedback to indicate how a **system** is performing. The metrics most used for kanban systems are:

- **Lead time** in different versions
- **Delivery rate** or **throughput**
- Levels of **WIP** in various parts of the system
- **Blockers**
- **Failure demand**

Other metrics such as quality or rework may be quite valuable.

Teams should not have too many metrics to start, but at the same time should be cognizant that “you get what you measure”, so collection of metrics should be designed to counter or minimize any gaming of the system while supporting the management of flow.

Common visualizations of metrics used in understanding a system are:

- Lead time distribution
- Lead time or delivery rate run chart
- Cumulative flow diagram

Option

Prior to the commitment point we have many ideas, requirements, or needs that have come from customers and which may have value. Some people may refer to this as a backlog, but in the Kanban Method we prefer to call the things that may have value as options. They are managed in the upstream part of the Kanban system. We wish to vet these options and constrain them against available capacity, also considering their urgency. This frequently results in discarding many of the options prior to service delivery which helps ensure that the service delivery is focused on high value requests.

Overburden

A situation when demand exceeds capacity and demand is not limited by available capacity. Overburdening is commonly observed when **work in progress** is not **limited** resulting in work being **pushed** into the system.

Pull Signals

A work item in an upstream can be pulled downstream only when capacity is available. The available capacity is the signal. In a kanban system the signal is that the work in progress is below the current work in progress limit.

Pull System

A system for delivering work only when both demand exists and delivery capacity is available. A kanban system is an example of a pull system which uses pull signals to indicate available capacity. **WIP limits** are the means to implement pull systems in Kanban.

Push System

A system or activity where work is placed into the system without regard to whether capacity is immediately available. The opposite of **pull system**.

Replenishment

The act of reviewing the requested work items that fulfill the “ready for delivery” criteria and selecting the ones to be pulled into the system based on available capacity.

Review

A type of cadence designed to reflect on the performance or risk of one or more services with a goal to improve service delivery. Reviews are largely performed based on data and observations. Examples include:

- Service Delivery Review
- Risk Review
- Operations Review

Run Chart

A chart that shows an observed metric in a time sequence. Commonly used to visualize the past Lead Times or Delivery Rates. One of the main benefits of the Run Chart is to see any trending of the data. A common question to consider is, “Are lead times going up or down, or within the expected variation?”

Scientific Method

A feedback loop designed to iterate towards a better understanding of a given situation. The steps of the scientific method are:

- Observe
- Hypothesis
- Experiment
- Collect data
- Analyze results
- Accept or Reject hypothesis
- Repeat

Service

A service starts with a customer request, fulfilled through **service delivery**, which results in acceptance by the customer. From a service perspective, a request can scale from a task to the development of a product, project, or initiative.

Service Delivery

The performance of a series of activities, also known as a workflow, to fulfill customer requests. For example, the activities performed in a service support the delivery of a work item. A service can be managed by one or more kanban systems.

Service Delivery Principles

The core principles of the Kanban Method for Service Delivery. The principles are:

- Understand and focus on customer needs and expectations
- Manage the work; let people self-organize around it
- Regularly review the network of services and its policies in order to improve outcomes

Service Level Expectation

What can be expected in future service delivery based on prior performance. In knowledge work we have inherent uncertainty and prefer to set expectations based on historical probabilities. If we know that historically we have a stable system that delivers 85% of work items within 10 days, we can set a corresponding Service Level Expectation (SLE) of 10 days with 85% confidence.

Standard

An archetype of a class of service that is applied to work items where the impact of delay is typical of current conditions.

STATIK

An acronym for the Systems Thinking Approach to Introducing kanban, a recommended approach to introducing kanban in a new context. It is a “how-to” for designing service-oriented kanban systems. Common activities when using STATIK include:

- Identify sources of dissatisfaction
- Analyze demand
- Analyze capability
- Model the workflow
- Identify classes of service
- Design the kanban system

System Lead Time

The elapsed time it takes for a work item to move from the commitment point to the first column on its kanban board that has no work in progress limit.

Team Retrospective

A type of **review** focused on exploring potential improvements based on learnings from the past. At higher levels of organizational maturity it is frequently found that this review does not need to exist on its own but becomes part of other reviews and **meetings**.

Throughput

The number of **work items** exiting from the system or a given part of it; measured in work items delivered per time period. Throughput is often referred to as delivery rate.

Visualize

One of the **6 general practices** in the **Kanban Method**. We want to make work and its flow visible so that it is not hidden. We also want to make information about the work, its risks or the process visible through such elements as charts.

Workflow

A series of activities, often performed within a kanban service, that results in products or services being delivered. Usually, the workflow starts with a request and ends with delivery. The workflow or a selected part of it is represented on a **kanban board** by a set of sequential columns showing the knowledge discovery activities that the **work item** passes through. For activities that occur in parallel or in no particular order, workflow-related information can be added to the kanban **card**.

Work in Progress (WIP)

The work items which have entered the system or activity under consideration and have not exited.

Work in Progress Limit (WIP Limit)

A policy that constrains the amount of work allowed in a given part of the kanban system. Can be applied for instance on columns, swimlanes, per person, or for the whole system. Limits prevent starting new work when downstream capacity is already utilized and unable to complete the work. This is also known as limiting work in progress.

Work Item

A deliverable or a component thereof resulting from the **demand** placed at the system that will be worked on by the service.

Work Item Type

A grouping of work items that behave similarly and follow the same workflow. These different types depend on the **demand** of the service and will vary in their form and size, specific to each kanban system. Examples of work item types are information requests, campaigns, incidents, software bugs, product features, whole products, or projects.