

Kanban in der Praxis

Vom Teamfokus zur Wertschöpfung

Bearbeitet von
Klaus Leopold

1. Auflage 2016. Buch. 237 S.

ISBN 978 3 446 44343 3

Format (B x L): 17,9 x 24,7 cm

Gewicht: 534 g

Wirtschaft > Betriebswirtschaft: Theorie & Allgemeines > Wirtschaftsinformatik, SAP,
IT-Management

schnell und portofrei erhältlich bei


DIE FACHBUCHHANDLUNG

Die Online-Fachbuchhandlung beack-shop.de ist spezialisiert auf Fachbücher, insbesondere Recht, Steuern und Wirtschaft. Im Sortiment finden Sie alle Medien (Bücher, Zeitschriften, CDs, eBooks, etc.) aller Verlage. Ergänzt wird das Programm durch Services wie Neuerscheinungsdienst oder Zusammenstellungen von Büchern zu Sonderpreisen. Der Shop führt mehr als 8 Millionen Produkte.



Leseprobe

Klaus Leopold

Kanban in der Praxis

Vom Teamfokus zur Wertschöpfung

ISBN (Buch): 978-3-446-44343-3

ISBN (E-Book): 978-3-446-44654-0

Weitere Informationen oder Bestellungen unter

<http://www.hanser-fachbuch.de/978-3-446-44343-3>

sowie im Buchhandel.

Inhalt

Warum dieses Buch?	VIII
Danke!	XI
Der Autor	XIII
1 Warum machen wir Kanban?	1
1.1 Schiffe im Arbeitsfluss	6
1.1.1 Wenn wir schneller arbeiten, wird trotzdem nicht mehr Arbeit fertig	9
1.1.2 Wir haben genug Zeit für Aufgaben, für die wir nie Zeit haben	11
1.1.3 Wir werden verlässlicher, wenn wir uns Grenzen setzen	12
1.1.4 Wenn alles Priorität hat, hat nichts Priorität	13
1.1.5 Je später wir anfangen, desto besser für den Kunden	14
1.1.6 Lokale Optimierung führt zu globaler Suboptimierung	15
1.2 Kanban: Viva la evolución!	17
1.3 Wertgenerierung: In Services denken	24
1.4 Ebenen der Gestaltung: Kanban Flight Levels	26
1.4.1 Flight Level 1: Team mit unreguliertem Input und Task-Fokus	27
1.4.2 Flight Level 2: Team mit koordiniertem Input, Etablierung eines Arbeitsflusses	28
1.4.3 Flight Level 3: Optimierung des Wertstroms	30
1.4.4 Flight Level 4: Optimierung von Strategie und Portfolio	32
2 Kanban-Systeme betreiben und verbessern	37
2.1 Visualisierung, WIP-Limits und Arbeitsfluss	37
2.1.1 Mit WIP-Limits arbeiten	40
2.1.2 Über Wert und Fluss	47
2.1.3 Mit mehreren Arbeitstypen umgehen	50
2.1.4 Veränderung von Arbeitstypen im Zeitverlauf	51
2.1.5 Nicht geplante Arbeit	54
2.1.6 Definition of Done	55

2.2	Umgang mit Blockaden	58
2.2.1	Blocker Clustering	58
2.2.2	Umgang mit Rückfluss und Defekten	65
2.2.3	Priorisierung von Lösungen	68
2.3	Kundenvalidierung	75
2.4	Wissenstransfer	79
2.4.1	Kapazitätsengpass (Capacity Constrained Resource)	79
2.4.2	Spezialistenengpass (Non Instant Availability)	81
2.4.3	Spezialisten vs. Generalisten	82
2.5	Koordination	87
2.5.1	Von der Idee zum koordinierten Input – Nachschubmeeting	88
2.5.2	Von der Input Queue ins Kanban-System – Regular Standup Meeting	89
2.5.3	Besser werden – Retrospektive	91
2.5.3.1	Check-in	95
2.5.3.2	Fakten sammeln	97
2.5.3.3	Den Problemkern definieren	98
2.5.3.4	Das Problem verstehen	99
2.5.3.5	Lösungen finden und Maßnahmen beschließen	100
2.5.3.6	Check-out	101
2.5.3.7	Verbesserungsarbeit sichtbar machen	102
3	Kanban im Großen	105
3.1	Praxisbeispiel: Eine Verkaufsplattform mit mehr als 200 Projektmitarbeitern ..	108
3.2	Kanban vergrößern	114
3.2.1	Aggregation von Services	116
3.2.2	Verbinden von Services	117
3.2.3	Geteilte Services	119
3.3	Kanban im ganz Großen bei Bosch Automotive Electronics	121
4	Forecasting	127
4.1	Voraussetzungen für das Forecasting	129
4.2	Forecast für eine Arbeitseinheit	134
4.3	Forecast für mehrere Arbeiten ohne historische Daten	139
4.3.1	Bestimmung von Minimum und Maximum	139
4.3.2	Monte-Carlo-Simulation	143
4.3.3	Kontinuierliches Durchsatz-Forecasting	147
4.3.4	Interview mit Troy Magennis	155
4.4	Kann man dem Forecast trauen?	156
4.4.1	Zusammenhang zwischen Work in Progress, Durchlaufzeit und Durchsatz	157
4.4.2	Messen der Stabilität eines Systems	159
4.4.3	Stabilitätsmuster interpretieren	161
4.4.4	Interview mit Dan Vacanti	164

5	Von der Priorisierung zur Risikobewertung	169
5.1	Die Nachfrage durch Verzögerungskosten managen	174
5.2	Verzögerungskosten quantifizieren	179
5.2.1	Schritt 1: den Wert bestimmen	179
5.2.2	Schritt 2: die Verzögerungskosten bestimmen	182
5.2.3	Schritt 3: Sequencing	187
5.2.4	Bestimmung der Verzögerungskosten in der Praxis	192
5.3	Weitere Nachfüllfaktoren: Risikobetrachtungen	197
5.3.1	Risikoarten	199
5.3.2	Risiken quantifizieren	200
5.3.2.1	Freie Risikobewertung	201
5.3.2.2	Modellbasierte Bewertung	204
5.4	Interview mit Markus Andrezak	208
6	Vom Ticket-Handling zur strategischen Weiterentwicklung – Kanban bei der STUTE Logistics (AG & Co.) KG	213
	Index	223

Warum dieses Buch?

„Und hier waren sie wieder, meine drei Probleme: kein Job, kein Geld und keine Ahnung, wie es weitergehen soll“, sagte Otto Waalkes in „Otto – Der Film“. In diesem Buch will ich Lösungen für die wichtigsten Probleme anbieten, die bei der Arbeit mit Kanban auftreten können. Für Kanban-Praktiker sollen es Anregungen sein, wie es nach einem erfolgreichen Start weitergehen kann.

Warum gibt es eigentlich dieses Buch, es gibt ja schon „Kanban in der IT“? Gute Frage! In „Kanban in der IT“ ging es Sigi Kaltenecker und mir um eine Kurzeinführung in das Thema. Der Schwerpunkt lag auf der Frage, wie man Kanban überhaupt in einem Unternehmen etablieren kann. Dabei ging es um Grundzüge des Veränderungsmanagements und wir haben sehr praxisnahe beschrieben, wie man ein Kanban-System aufsetzen und erfolgreich starten kann. Aber der Start ist eben nur der Anfang einer langen Reise auf dem Weg der kontinuierlichen Verbesserung.

Der Titel „Kanban in der Praxis“ beschreibt eigentlich recht gut, was Sie von diesem Buch erwarten können. Ich glaube, dass ich mich als richtigen Praktiker bezeichnen kann. Ich habe einige Zeit an der Universität gearbeitet, so richtig glücklich wurde ich dabei jedoch nicht. Die Forschungsarbeit gefiel mir sehr gut, aber ich sah nie die Anwendung der Resultate im „echten“ Leben, also im Praxiseinsatz in Unternehmen. Zu oft dachte ich mir: „Das ist eine tolle Lösung, aber es gibt kein Problem dafür.“ Seit ich als Kanban-Trainer und -Berater arbeite, hat sich das zum Glück vollkommen geändert. Alleine im Jahr 2015 war ich am Bau und Start von mehr als 100 Kanban-Systemen beteiligt. Seit Beginn meiner Kanban-Karriere im Jahr 2008 haben sich die Erfahrungen aus mehreren Hundert Kanban-Systemen angesammelt. Meistens besuche ich Unternehmen einige Zeit nach dem Start erneut, analysiere die eingeführten Kanban-Systeme und halte dann Verbesserungsworkshops. Einige Problemuster zeigen sich dabei immer wieder. Ich dachte mir: Wenn ich die Probleme gut verstehe, kann ich noch bessere Lösungen anbieten. Daher begann ich damit, alle ausgearbeiteten Lösungen zu protokollieren und so entstand eine umfangreiche Lösungsbibliothek für alle möglichen Probleme. In diesem Buch habe ich mir zur Aufgabe gemacht, die wesentlichen Erfahrungen zu strukturieren und der Allgemeinheit zugänglich zu machen.

Welche Probleme sind das nun, die nach dem Start von Kanban-Systemen immer wieder auftreten? Sie lassen sich in diesen Fragestellungen zusammenfassen:

- Machen wir überhaupt „richtiges“ Kanban?
- Wie können wir unser Kanban verbessern?
- Wie können wir Kanban vergrößern (ausweiten)?
- Was hilft uns dabei, in unserer Arbeit vorhersagbarer zu werden?
- Woran sollen wir als Nächstes arbeiten?

In **Kapitel 1** geht es noch einmal um die Fundamente von Kanban. Fortgeschrittene Praktiker werden vielleicht denken: „Das haben wir doch alles in der Pfeife geraucht, kennen wir schon.“ Gut möglich, aber lassen Sie sich trotzdem noch einmal auf die Basics ein. Oft kommt der Verbesserungsmotor deshalb nicht in Schwung, weil es am grundlegenden Verständnis der Wirkmechanismen von Kanban mangelt. Warum machen wir das überhaupt? Was ist ein Arbeitsfluss? Wieso sollten wir überhaupt die Arbeit im System limitieren? Wo kann Kanban im Unternehmen eingesetzt werden? Diese Fragen beantworte ich in Kapitel 1.

In **Kapitel 2** „Kanban-Systeme betreiben und verbessern“ tauchen wir in die Praktiken von Kanban ein. Auch hier sehe ich in der Praxis immer wieder, dass gut gemeint das Gegenteil von gut gemacht ist. Kennen Sie das: Dauernd beklagt sich jemand, dass die Projekte zu lange dauern. Kanban wird eingeführt, WIP-Limits werden etabliert und eigentlich sollten die Projekte jetzt wesentlich schneller fertig werden. Wenn Projekte trotzdem noch lange dauern, liegt es mit an Sicherheit grenzender Wahrscheinlichkeit daran, dass die falsche Arbeit limitiert wird. Wie man WIP-Limits richtig setzt und viele weitere Tipps und Tricks rund um Visualisierung, WIP-Limits und Arbeitsfluss finden Sie daher in Kapitel 2. Aus meiner Praxiserfahrung zeige ich in diesem Kapitel, wie man aus Blockaden lernen kann, wie man Bugs richtig behandelt und wie man mit Engpässen umgehen kann. Abgeschlossen wird das Kapitel mit den Kanban-Meetings: Welche sind überhaupt sinnvoll und wie gestaltet man sie am effektivsten? Kapitel 2 ist also ganz klar ein Ideenpool für Verbesserungen am technischen Kanban-System und es hilft Ihnen dabei, die Frage zu beantworten, ob Sie Kanban „richtig“ machen.

Meistens wird Kanban nicht mit einem Big Bang eingeführt. Es arbeiten also nicht von einem Tag auf den anderen plötzlich Hunderte oder Tausende Menschen im Unternehmen mit Kanban. Sehr oft beginnt es mit einigen Pilot-Teams, die erste Erfahrungen sammeln. Wenn man sich nicht völlig ungeschickt anstellt, werden bald die ersten positiven Veränderungen sichtbar. Das Motto lautet nun: „Sehr gut, Kanban funktioniert. Wir wollen mehr davon.“ Das Trainer- und Beraterauge tränt vor Freude – Dutzende, vielleicht sogar Hunderte Teams müssen auf Kanban umgestellt werden. Es können also sehr viele Tage fakturiert werden. Schön für ein Beratungsunternehmen, schlecht für das Kanban-Unternehmen in spe. Die Leistung dieses Unternehmens wird im besten Fall gleich bleiben, im Normalfall wird sie sich verschlechtern. In **Kapitel 3** beschreibe ich, warum das so ist, und ich stelle anhand eines Fallbeispiels vor, wie Kanban im Großen besser gemacht werden kann.

Kapitel 4 widmet sich voll und ganz dem Thema Forecasting. Kunden und Auftraggeber wollen verständlicherweise wissen, wann sie mit der Fertigstellung einer Arbeit rechnen können. Es gilt nun also, die Zukunft vorherzusehen, und das ist nicht unbedingt die einfachste Aufgabe. Aus heutiger Sicht gibt es nämlich nicht nur eine Zukunft, sondern viele verschiedene Zukünfte – es kann und wird Unvorhergesehenes passieren. Glücklicherweise sind manche Zukünfte wahrscheinlicher als andere und diesen Vorteil machen wir uns in

Kapitel 4 zunutze, um einen Forecast darüber zu erstellen, wann Arbeit fertig wird. Das Schöne an diesem Ansatz ist, dass keine Zeit aufs Schätzen verschwendet werden muss. Mit wenigen Daten lassen sich bereits Forecasts erstellen, die wesentlich präziser sind als die meisten Schätzungen.

Im Grunde ist es eine gute Situation, wenn viele Arbeiten im Ideenpool, Backlog oder Optionenbereich eines Kanban-Boards liegen. Das bedeutet, dass genügend Nachfrage nach den Services und Produkten existiert. Hoffentlich gibt es auch immer mehr Nachfrage als Kapazität vorhanden ist, um die Nachfrage zu decken. Was passiert nämlich, wenn es wesentlich mehr Mitarbeiter als Arbeit gibt? Stimmt, es werden Mitarbeiter abgebaut. Wir müssen also die Tatsache begrüßen, dass es mehr Nachfrage als Kapazität gibt. Daran knüpft sich aber die Frage: Woran sollen wir zuerst arbeiten und was muss noch warten? Genau diese Frage beantworte ich in **Kapitel 5**. Allerdings will ich gleich vorwegschicken, dass ich nicht beanspruche, die *einzig richtige* Priorisierungsmethode vorzustellen. Es gibt viele verschiedene Ansätze und viele davon haben durchaus ihre Berechtigung. In diesem Buch stelle ich die Verzögerungskosten (Cost of Delay) als meinen favorisierten Ansatz dar. Vor allem am Anfang des Kanban-Lebens empfehle ich, die Verzögerungskosten als Priorisierungsmethode zu verwenden, weil sich der Fokus von den Kosten auf den Wert der Ergebnisse verlagert. Zeit bekommt ein Preisschild und damit können ökonomisch sinnvolle Entscheidungen getroffen werden. Wenn die Ökonomie der Verzögerung in Fleisch und Blut übergegangen ist, können Sie den Werkzeugkasten der Priorisierung um Risikobewertungen erweitern. Unterm Strich ist Priorisierung nämlich nichts anderes, als eine Risikoentscheidung zu treffen.

In **Kapitel 6** berichte ich von der ereignisreichen Kanban-Reise der STUTE Logistics (AG & Co.) KG. IT-Leiter Holger Rieth hat seine Abteilung innerhalb von eineinhalb Jahren zu Höchstleistungen angespornt. Vor Kanban herrschte der „normale“ IT-Alltag: viele offene Tickets und Anfragen, lange Wartezeiten und ein verbesserungswürdiger Verbesserungsprozess. Heute gelangen die Tickets in einen geregelten Arbeitsfluss und durch ein automatisiertes Forecasting wissen die Kunden genau, wann etwas fertig sein wird. Alle Mitarbeiter können Verbesserungsideen einbringen, für die Verzögerungskosten bestimmt werden, um eine effektive Reihenfolge der Abarbeitung zu bestimmen. Im Verbesserungsprozess der IT-Abteilung von STUTE spiegeln sich sämtliche Themen dieses Buchs wider und daher freue ich mich sehr, dass ich diese Geschichte erzählen darf. Ich konnte mir keinen besseren Abschluss für dieses Buch wünschen als ein reales Unternehmensbeispiel, das alle Inhalte bei ihrer Anwendung in der Praxis zeigt.

Übrigens: Auf www.kanbaninderpraxis.com finden Sie unterstützende Materialien, mit denen Sie die Ideen in diesem Buch umsetzen können.

Klaus Leopold

Wien, Herbst 2016

Danke!

Es ist geschafft! Dieses Mal hat es ziemlich genau zwei Jahre von der ersten Zeile bis zur Veröffentlichung dieses Buchs gedauert und es haben viele Menschen dazu beigetragen, dass ich dieses Buch über die Ziellinie bringen konnte.

Allen voran will ich mich bei meiner Lebens- und Geschäftspartnerin Katrin Dietze bedanken, die permanent meine unausgegorenen Visionen und Ideen aushalten muss. Sie ist mein wichtigster Sparringspartner und dank ihr reifen Themen nicht nur in meinem Kopf vor sich hin, sondern werden auch auf den Boden der Realität gebracht. Katrin schafft es auch jedes Mal, meine Kritzeleien in schöne Illustrationen zu übersetzen.

Ich möchte mich sehr herzlich bei meiner Text-Fee Dolores Omann bedanken, die es wirklich versteht, meine Gedanken und Texte in lesbares Deutsch zu übersetzen. Die Arbeit mit ihr ist humorvoll und professionell, eine Kombination, die ich wirklich schätze. Dolores findet auch immer wieder die richtigen Stellen, an denen Kommata ganz gut in den Text reinpassen (bestimmt hat sie auch in diesem Absatz sprachlich eingegriffen).

Ein herzlicher Dank geht an Holger Rieth, der sich bereit erklärt hat, seine Kanban-Geschichte als IT-Leiter bei STUTE Logistics (AG & Co.) KG in diesem Buch zu erzählen. Es scheint auch, dass Holger die Texte richtiggehend aufgesaugt hat, denn seine Anmerkungen als Reviewer gingen weit über das Erwartete hinaus.

Ein besonderes Dankeschön sage ich den Mathematik-Profis Ralf Männel von RME und Horst Eidenberger von der Technischen Universität Wien. Die zwei haben die mathematische Korrektheit des Forecasting-Kapitels erhöht, ohne dass ich daraus ein unlesbares mathematisches Formel-Gewirre machen musste. Alles begann mit der lapidaren Anmerkung von Horst: „Das musst du mir aber jetzt zeigen, dass elf Samples ausreichen.“ Das Resultat waren unzählige Diskussionen und E-Mails mit meinem lieben Freund und Forecasting-Profi Troy Magennis. Zum Schluss schaltete sich auch noch mein „Schwieger-Onkel“ Ralf Männel ein, der schließlich den mathematischen Beweis führte, dass im vorliegenden Kontext 19 Samples korrekt sind und nicht elf. Das war sensationell, liebe Leute! Ich möchte mich auch noch mal extra bei Troy bedanken: Er legt seit Jahren die Latte beim Thema Forecasting ständig höher und er ließ sich außerdem von mir für dieses Buch interviewen.

Forecasting ohne Metriken ist wie eine Suppe ohne Salz. Ich bin wirklich glücklich, Daniel Vacanti unter meinen Freunden zu finden, denn er ist das Salz in der Forecasting-Suppe. Ich schätze seine (mathematische) Genauigkeit sehr, denn durch sein penibles Feedback wurde ich dazu angehalten, trotz meines Anspruchs auf einfach zu konsumierende Darstellungen

nicht die notwendige Genauigkeit aus den Augen zu verlieren. Auch Dan hat mir ein spannendes Interview für dieses Buch gegeben.

Ich bedanke mich herzlich bei meinem langjährigen Freund und Kanban-Wegbegleiter Markus Andrezak. Vor allem seine Sichtweise auf Innovation und Unternehmensstrategie halfen mir, genau abzugrenzen, welche Inhalte Teil dieses Buchs sein sollten und welche ich mir für ein nächstes Buch aufsparen sollte. Im Interview haben wir diese Grenze noch einmal genau herausgearbeitet.

Vielen lieben Dank an Andreas Haugeneder und Hans-Oliver Ruoff von Bosch Automotive Electronics. Dort betreiben wir Kanban definitiv im Großen. Andreas und Oliver sind federführend an der Einführung von Kanban in Deutschland, China, Indien, Ungarn und in den USA beteiligt. Ihre meist sehr ambitionierten Ziele veranlassen mich dazu, meine eigenen Limits zu erweitern, und spornen mich zu Höchstleistungen an. Im Interview vermitteln sie einen Einblick, was in dieser Kanban-Initiative passiert.

Matthew Philip war einer der ersten Anwender meines Blocker-Clustering-Verfahrens außerhalb meines Kundenkreises. Im Interview in Kapitel 2 erzählt er von seinen Erfahrungen – vielen Dank, Matt!

Man sollte meinen, nach so viel Unterstützung sei es genug, aber das ist bei diesem Buch nicht der Fall. Ich bekam immer wieder Feedback von meinen Reviewern Andreas Schliep, Anton Spitzer, Arne Roock, Eric-Jan Kaak, Florian Eisenberg, Oliver Finker, Michael Beyer, Siegfried Kaltenecker, Stefan Rook und Wolfgang Wiedenroth. Danke, danke, danke!

Und schließlich: Vielen Dank an David Anderson, der meine Begeisterung für das Thema Kanban geweckt hat.

Der Autor



Dr. Klaus Leopold ist Informatiker und Kanban-Pionier. Seit 2008 arbeitet er als Lean- und Kanban-Berater und zeigt pro Jahr rund 1000 Workshop- und Trainingsteilnehmern und -teilnehmerinnen, wie sie mit Kanban starten und es weiterentwickeln können. Er berät weltweit agierende Unternehmen bei der Einführung von Lean und Kanban, in den damit verbundenen Change-Prozessen und bei der Optimierung der Wertschöpfung. Klaus Leopold ist Co-Autor des Standardwerks „Kanban in der IT“ und veröffentlicht seine Überlegungen und Erfahrungen u. a. auf seinem Blog www.LEANability.com. Außerdem ist er gefragter Speaker auf international renommierten Lean- und Kanban-Konferenzen. 2014 wurde Klaus Leopold in San Francisco mit dem Brickell Key Award für „Outstanding Achievement and Leadership“ ausgezeichnet.

1

Warum machen wir Kanban?

„Increased team performance can lead to decreased system performance“, twittet ein Teilnehmer von unserem Applying-Kanban-Training in Bangkok. Innerlich atme ich in solchen Momenten auf, denn die Teilnehmerinnen und Teilnehmer haben verstanden: Kanban ist keine Methode, um Teams zu optimieren. Wer das denkt, begrenzt von Anfang an die Möglichkeiten seiner Organisation. Dass Kanban eine Teammethode ist, gehört zu den großen Missverständnissen, die mich hartnäckig bei vielen meiner Aufträge, Vorträge und Trainings verfolgen. Auch wenn es von allen Missverständnissen zweifellos das größte ist, ist es aber nur eines von vielen. Wenn mich ein Unternehmen zu Hilfe holt und ich mir die Kanban-Boards ansehe, wenn ich beobachte, wie die Beteiligten mit vorbildlicher Vehemenz auf falschen Annahmen aufbauen, bedeutet das für mich, mit den Akteuren einen Schritt zurückzugehen, die Prinzipien und Praktiken noch einmal genauer zu beleuchten und noch deutlicher zu machen, worum es bei Kanban eigentlich geht.

So wie jede mehr oder weniger neue Methode ist auch Kanban in das Fahrwasser des Trends geraten und das birgt Risiken. In wirtschaftlich turbulenten Zeiten – ob selbst verursacht oder im Sog von Wirtschaftskrisen und Umbrüchen – suchen Unternehmen nach der nächsten Wundermethode, die alle Probleme wie von unsichtbarer Hand löst. Dankbar springen sie auf den nächsten Trend auf und der heißt momentan Agilität. Unternehmen hoffen, damit das Glück zu finden, dem sie seit einiger Zeit vergeblich hinterherlaufen. Allerdings stellen sie irgendwann frustriert fest, dass es mit dem braven Befolgen einiger Praktiken nicht getan ist. Hinter jeder Methode steckt eine Reihe von Erkenntnissen und Einsichten, von Prinzipien und Werten, die man verstanden haben muss, um die Methode erfolgreich anwenden zu können.

Genau das erlebe ich mit Kanban: Alle Hoffnungen ruhen auf einer Methode. Mehrmals pro Monat sehe ich motivierten Menschen in verschiedenen Unternehmen dabei zu, wie sie innerhalb von zwei Tagen die tollsten Kanban-Systeme entwerfen. Die Kunstwerke aus bunten Haftnotizen, akkurat gezogenen Spalten und Swimlanes hängen an der Wand, davor stehen stolze, zufriedene Mannschaften. Und dann fällt dieser eine Satz: „Ich hoffe wirklich, dass uns Kanban schneller macht.“ Oder: „Hoffentlich nützt uns Kanban irgendwie.“ Während diese Sätze langsam in meinem Hinterkopf versickern, versuche ich mir den Alltag in diesen Unternehmen vorzustellen: Jeden Tag sitzen Menschen in einem Halbkreis vor dem Kanban-Board. Ein Mitarbeiter schreibt auf ein rotes Post-it: „Blockade: Test-Infrastruktur nicht fertig.“ Er steht auf, geht zum Board und klebt es auf einen der Zettel in der Spalte „TEST“. Er dreht sich um. Die Kolleginnen und Kollegen werfen ihm fragende Blicke zu. Einige Sekunden Stille. Schließlich fragt jemand: „Und das hilft?“



@drunkcod (Torbjörn Gyllebring): „Sometimes ppl don't know what their problem is. So they get a tool to solve it. They now have n+1 problems.“

Überraschenderweise hilft das nicht! Nur weil man etwas mit kreativem Eifer und bunten Zetteln visualisiert hat, verschwinden die Schwächen eines Systems nicht. Oder sind Sie schon einmal schlanker geworden, weil Sie neben einem Heimtrainer gesessen sind? Ob man sich einen Heimtrainer kauft oder eine neue Methode, um effizienter zu arbeiten: Beides entfaltet seine Wirksamkeit nicht durch den Kauf, sondern durch das Handeln. Menschen sind die treibende Kraft hinter jeder Veränderung. „Es“ – ob Heimtrainer oder Kanban – ist nicht mehr als ein Werkzeug. Genau so möchte ich Kanban vermitteln: Es hilft uns zu sehen, wo die Schwachpunkte eines Systems liegen, und offenbart damit, wie wir für den Kunden besser Wert generieren können. Manchmal schreit uns das Kanban-Board das Problem (und sehr oft auch die Lösung) förmlich entgegen, manchmal gibt es die Knackpunkte etwas langsamer preis. Das kann so ein Board aber nur, wenn es von mitdenkenden Menschen bedient wird, die Probleme erkennen wollen und die verstanden haben, dass Kanban nur durch die Energie etwas verändern kann, die sie selbst investieren. Wenn man das verstanden hat, kann man über Verbesserungen nachdenken. Dann kann man gezieltere, weil bewusstere Entscheidungen treffen. Das funktioniert unter zwei Voraussetzungen:

1. Die Beteiligten entwerfen ihr eigenes Arbeitssystem statt kritiklos Best Practices zu übernehmen, die eine bestimmte Methode vorschreiben. „Best Practice“ ist ohnehin ein zweifelhafter Begriff, denn eigentlich bedeutet er das Ende jeder Erkenntnismöglichkeit. Auch wenn man noch so viele Kanban-Trainings besucht, um zu verstehen, wie man bestimmte Dinge verbessern kann: Tun muss man es selbst.

2. Die Beteiligten haben ihr Problem verstanden, bevor sie sich die Lösung überlegt haben. Allerdings ist gerade das einer der großen Schwachpunkte des heutigen Managements: Es werden wahllos Lösungen installiert, ohne das zugrunde liegende Problem zu verstehen. Der Fokus auf Lösungen macht ein Unternehmen aber nicht anpassungsfähiger an die Unwägbarkeiten und Überraschungen einer komplexen Umwelt. Wenn Kanban in einigen Unternehmen also wider Erwarten nicht hält, was es nie versprochen hat, liegt die Ursache dafür vermutlich ganz woanders: Es unterstützt nicht die Schablonenwünsche eines faulen Managements. Wer eine To-do-Liste oder ein paar einfach umzusetzende Patentrezepte sucht, sollte es mit Kanban erst gar nicht probieren und auf der Liste der trendy Methoden einfach abhaken.



Toyota Production System: Was ist richtig?

Viele Ideen zu Kanban in der Wissensarbeit stammen aus dem Toyota Production System (vgl. Ohno, Bodek 1988). Toyota zeigt jedem, den es interessiert, sehr bereitwillig seine Arbeitsweise. Allerdings ist diese Arbeitsweise immer nur eine Momentaufnahme, die derzeit passende Lösung zu den gegebenen Voraussetzungen, Problemen und Herausforderungen. Viele der Besucher verstehen das nicht.

Mike Rother erzählt in seinem Buch „Toyota Kata“ die Geschichte eines Freundes, der in Japan mehrmals eine Fabrik von Toyota besucht hatte (vgl. Rother 2010). Bei seinem ersten Besuch beobachtete er, dass die Mitarbeiter an den

Fertigungsstraßen die Arbeitsstücke aus Durchlaufregalen nahmen, die mit den Bestandteilen der unterschiedlichen Fahrzeugmodelle befüllt waren. Je nachdem, welches Modell gerade an der Reihe war, konnte der Mitarbeiter also den passenden Bestandteil aus dem Regal nehmen. Dank der Durchlaufregale konnten mehrere Modelle auf einer Fertigungsstraße bearbeitet werden – andere Automobilhersteller haben dieses System oft kopiert.

Nach einigen Jahren besuchte Mike Rothers Freund dieselbe Fabrik erneut, aber: Alle Durchlaufregale waren verschwunden. Die Bestandteile steckten nun in einem Bausatz, der gemeinsam mit dem jeweiligen Modell die Fertigungsstraße entlangfuhr. Kam das Fahrzeug in einer Fertigungszelle an, musste der Mitarbeiter nur mehr zum immer an der gleichen Stelle angebrachten Bausatz greifen.

Mike Rothers Freund war aufgebracht. Welcher Ansatz denn nun der richtige sei, fragte er seinen Gastgeber von Toyota. Das Durchlaufregal oder der Bausatz? Der Gastgeber verstand die Frage nicht. Er antwortete: „Als Sie vor einigen Jahren zum ersten Mal hier waren, haben wir auf dieser Straße vier verschiedene Modelle gefertigt. Heute montieren wir acht verschiedene Modelle auf derselben Linie – es war nicht mehr möglich, die unterschiedlichen Bestandteile in einem einzigen Regal zu halten. Abgesehen davon wollen wir uns einem One-piece-Flow annähern. Immer wenn Sie uns besuchen, sehen Sie einfach die Lösung, die wir für eine spezielle Situation zu einem bestimmten Zeitpunkt entwickelt haben.“

Welche Probleme versuchen wir zu lösen?

Wenn ich eine Beratungsanfrage bekomme, versuche ich die Unterhaltung auf den richtigen Weg zu bringen: Warum will das Unternehmen eigentlich Veränderungen herbeiführen? Welche Probleme versucht es mit dieser Veränderung zu lösen? Solange diese Fragen nicht geklärt sind, läuft jede Methodik – ob sie nun Kanban oder anders heißt – völlig ins Leere. Ohne Antworten darauf kann ein Unternehmen nicht die Parameter definieren, an denen es festmacht, ob sich die Situation verbessert hat. Sehen wir uns dazu Bild 1.1 an: Welches der beiden Systeme ist Ihrer Meinung nach besser?

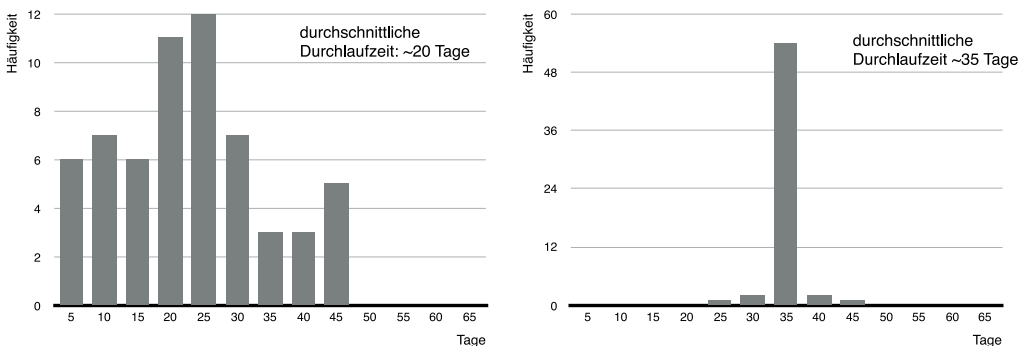
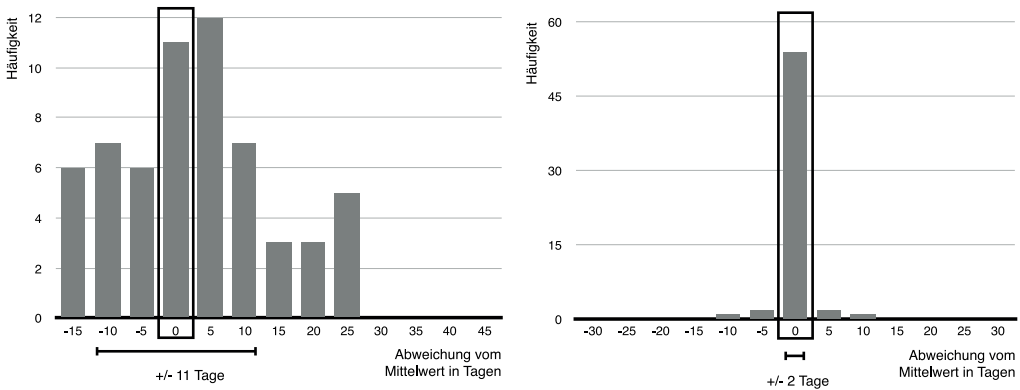


Bild 1.1 Welches ist das bessere System?

In beiden Grafiken sehen wir auf der x-Achse die Durchlaufzeiten, auf der y-Achse ist aufgetragen, wie oft welche Durchlaufzeiten auftreten. Berechnen wir den Mittelwert, so liefert

System 1 etwa alle 20 Tage und ist damit zweifellos schneller als System 2, das bis zur Lieferung durchschnittlich 35 Tage braucht. Schneller ist immer gut, oder? Was ist aber, wenn Schnelligkeit zwar wichtig ist, aber etwas anderes noch wichtiger ist? Sehen wir uns die zwei Systeme noch einmal aus einem anderen Blickwinkel an (Bild 1.2).



Das System ist schneller.

Das System ist zuverlässiger.

Bild 1.2 Geschwindigkeit vs. Vorhersagbarkeit – was ist besser?

Auf der x-Achse ist nun die Standardabweichung vom Mittelwert zu sehen. Jetzt ergibt sich ein anderes Bild: System 1 liefert zwar meistens alle 20 Tage, aber diese Aussage ist mit einer Unsicherheit von ± 11 Tagen behaftet. Sagen die Betreiber von System 2 hingegen, dass sie ihr Service nach 35 Tagen liefern, kann der Kunde davon ausgehen, dass er die Leistung mit einer Abweichung von ± 2 Tagen bekommt. System 2 ist berechenbarer. Also: Welches System ist besser?

Wahrscheinlich merken Sie schon, dass keines der beiden Systeme besser ist, solange wir nicht wissen, in welchem Kontext und in welcher Branche wir uns bewegen. In der Automobilindustrie laufen zu einem Stichtag die Produktionsstraßen für ein neues Modell an. Bis dahin muss die Entwicklung aller Einzelteile – und wir sprechen hier von bis zu 10 000 Einzelteilen – abgeschlossen sein, das Ergebnis der Entwicklung wird jetzt nur noch reproduziert. Jede weitere Verzögerung würde Unsummen kosten. Vorhersagbarkeit ist in diesem Fall jenes Kriterium, nach dem die „Fitness“ der Entwicklungsabteilung für den Zweck „neues Modell“ beurteilt werden könnte. Geht es hingegen um ein Software-Startup, ist vielleicht die Liefergeschwindigkeit wichtiger als alles andere, weil die Kunden regelmäßig etwas Neues sehen wollen, auch wenn die Qualität möglicherweise noch nicht so ausgereift ist.

Die eigentliche Frage lautet somit: **Was ist dem Kunden wichtig?** Jedes Unternehmen und jede Organisationseinheit in einem Unternehmen erfüllt einen Zweck. Für manche Unternehmen besteht er darin, möglichst schnell möglichst viel Geld zu verdienen. Andere Unternehmen wollen aber keine Eintagsfliegen sein, wollen länger am Markt bestehen und sehen es als wichtigsten Zweck, dem Kunden – also denjenigen Menschen, die schlussendlich die Gehälter bezahlen – guten Service zu bieten. Die internen Prozesse sollten darauf ausgerichtet sein, den Zweck bestmöglich zu erfüllen, damit der Kunde bekommt, was er braucht und will. Dazu muss eine Organisation erkennen, was ein Kunde an einem Service schätzt:

Nach welchen *Fitnesskriterien* (vgl. Anderson 2013) beurteilt er, ob er es für gut oder weniger gut befindet? Solche Fitnesskriterien sind sehr oft (aber nicht ausschließlich):

- Lieferzeit,
- Vorhersagbarkeit/Termintreue,
- Qualität,
- Sicherheit,
- Konformität mit regulatorischen Anforderungen.

Ein gutes Fitnesskriterium erkennt man daran, dass es aus der Sicht des Kunden definiert und messbar ist. In der Realität sind solche Kriterien – zum Beispiel Effizienz, Agilität oder Qualität – aber Begrifflichkeiten, die von jedem Menschen anders verstanden und mit verschiedenen Inhalten gefüllt werden. Wenn ich Kanban in einem Unternehmen einführen soll, höre ich zum Beispiel oft, dass man sich davon eine steigende Effizienz erwartet. Gut, aber wann ist das Unternehmen effizienter? Effizienz an sich ist ein solcher Puddingbegriff, der genauer definiert werden muss, denn die Vorstellung von Effizienz unterscheidet sich nicht nur von Unternehmen zu Unternehmen, sondern bereits von Mensch zu Mensch. Daher frage ich in solchen Fällen: „Was wollen wir mit Kanban verbessern und wie können wir messen, ob wir das erreicht haben?“ Durch diese beharrlich wiederholte Frage zerbricht die leere Begriffshülle der Effizienz recht schnell und weicht handfest quantifizierbaren Messgrößen.

In der intensiven Auseinandersetzung mit dem Zweck und den dafür gültigen Fitnesskriterien wird meistens auch deutlich, dass man nicht alles haben kann und Schwerpunkte setzen muss. Eine schnelle Lieferung und hohe Ressourcenauslastung sind in der Innensicht eines Unternehmens selten gleichzeitig zu erreichen, auch wenn es in Relation zu anderen Unternehmen beides erreichen kann. Auch eine immer kürzere Lieferzeit und Vorhersagbarkeit sind entgegengesetzte Pole – der Flugverkehr ist dafür das beste Beispiel: Piloten fliegen in der Regel langsamer als es technisch möglich wäre. Schneller zu fliegen bringt keinen Vorteil, weil sich Airlines an die Flugpläne halten und sich am Verkehr orientieren müssen, der am Zielflughafen aktuell vorherrscht. Schnell zu fliegen ist also nicht das Kriterium für die Fluggesellschaft, um einen Flug erfolgreich abzuschließen. Die Piloten müssen mit der *richtigen* Geschwindigkeit fliegen, um am Zielflughafen auch wirklich landen zu können. In der Automobilindustrie würde es sich ein Zulieferer auf Dauer mit den Kunden verscherzen, wenn er seine Bauteile plötzlich eine Woche früher liefern und so die Lager unnötig vollräumen würde. Nur weil man etwas kann, heißt das noch lange nicht, dass es in einem bestimmten Kontext auch sinnvoll ist.

Guter Service aus der Sicht des Kunden ist für ein Unternehmen daher immer mit der bewussten Entscheidung verbunden, Schwerpunkte zu setzen, aber in der Entwicklung nicht stehen zu bleiben. Das setzt voraus, dass man den Kunden erstens kennt und zweitens versteht. Ihren Kunden verstehen zu wollen, ist aber sehr oft das, worin Unternehmen am wenigsten Energie investieren. Ich merke das meistens daran, dass der Kunde in keinem einzigen Gespräch vorkommt, und wenn er doch erwähnt wird, dann meistens in der Rolle des Störfaktors. Im klassischen Change Management bekommt der Kunde immerhin eine Minute Ruhm durch die prominente Nennung auf der ersten PowerPoint-Folie – bis zum Organigramm ist er aber wieder vergessen. Den Kunden verstehen zu wollen, ist der Aus-

gangspunkt jeder Verbesserung und damit auch der Kern der Frage, warum man eigentlich mit Kanban arbeiten will.

Will sich ein Unternehmen „Fit für den Zweck“ machen, muss es seine eigenen Antworten auf zwei Fragen finden:

1. Wie finden wir heraus, was unsere Kunden wollen?
2. Wenn wir wissen, was unsere Kunden wollen: Wie werden wir den Anforderungen gerecht? Wie muss unser Arbeitsfluss aussehen, damit wir beim Kunden Wert generieren?

Die Antworten auf Frage 1 sind das Thema von Discovery Kanban (www.discovery-kanban.com), Praktiken wie Lean Startup, Lean UX oder Design Thinking. Eine spannende Materie und dringend zu empfehlen, wenn man seine Kunden wirklich verstehen will. Der Fokus dieses Buchs liegt aber auf der Antwort auf Frage 2. Dazu gehen wir davon aus, dass sich ein Unternehmen bereits eingehende Gedanken über seine Kunden gemacht hat. Wenn man also weiß, worauf der Kunde Wert legt: Wie kann man ein Arbeitssystem bauen, das die Kundenwünsche bestmöglich erfüllt?

Nun haben wir bereits festgestellt, dass Kanban – so wie jede andere Methode – nicht durch seine bloße Existenz Wunder vollbringen kann. Kein Schema oder Framework könnte jemals für jedes Unternehmen dieser Welt passen. Warum sollte Kanban dann ein geeignetes Mittel sein, um Frage 2 zu lösen?

■ 1.1 Schiffe im Arbeitsfluss

Kanban ist nicht immer intuitiv. Am Beginn mancher Kanban-Trainings, noch bevor es sich erwartungsvolle Teilnehmerinnen und Teilnehmer in der aufmerksamen Berieselungshaltung gemütlich machen können, erkläre ich daher mittlerweile ... nichts. Ich fange mit einer Übung an, die in einem Buch zwar nicht ganz einfach darzustellen ist, aber wagen wir den Versuch. Fühlen Sie sich in den kommenden Minuten ganz wie Aristoteles Onassis: Wir bauen jetzt Schiffe!



Ein Flussexperiment: Schiffe falten

Was man dafür braucht:

- Einen Stapel Papier
- Eine Stoppuhr
- Mind. sechs Teilnehmer, max. zwölf Teilnehmer in einem Arbeitssystem:
 - Der erste Teilnehmer in der Reihe gibt die Blätter einzeln ins System.
 - Der Teilnehmer am Ende der Reihe ist für die Zeitmessung zuständig.

Versuchsanordnung:

- Die Teilnehmer sitzen entlang einer Tischreihe.
- Mit Ausnahme des vorletzten Teilnehmers in der Reihe ist jeder Teilnehmer für max. zwei Faltschritte zuständig (abhängig von der Teilnehmerzahl). Jede Aktivität ist also mit nur einem Teilnehmer besetzt.

Als Aktivität bezeichne ich die einzelnen Schritte, die in einem Arbeitssystem gesetzt werden, um am Ende einen Wert zu generieren. Eine Aktivität kann unterschiedliche Aufgaben umfassen (im Falle dieses Experiments das Falten in verschiedenen Ausprägungen).

- Der vorletzte Teilnehmer in der Reihe muss mehrere recht komplizierte Faltschritte ausführen.
- Der letzte Teilnehmer misst die Durchlaufzeit, indem er Zeitstempel für jedes fertiggestellte Schiff nimmt.

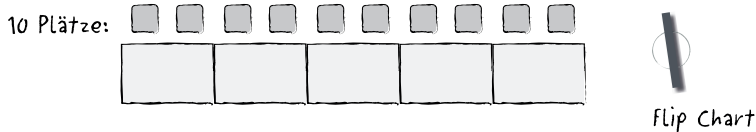
Wir führen das Experiment in zwei Durchgängen durch:

- *1. Durchgang:* Es fließt ohne Limitierung Arbeit in das System. Die Teilnehmer arbeiten nach dem Push-Prinzip und schieben ihre Teilprodukte daher in die nächste Aktivität weiter, sobald sie mit ihrer Aktivität fertig sind.
- *2. Durchgang:* Die Anzahl der Schiffe in jeder Aktivität im System wird mit einem WIP-Limit von 1 begrenzt. Das heißt: Sobald ein Teilnehmer seine Arbeitsschritte abgeschlossen hat, hebt er die Hände und lässt das Teilprodukt vor sich am Tisch liegen. Erst wenn der Kollege in der folgenden Aktivität ein Schiff gezogen hat und der Kollege aus der vorhergehenden Aktivität sein Schiff fertiggestellt hat, kann man das nächste Schiff ziehen (Pull-Prinzip).

Für die zwei Durchgänge gelten folgende Regeln:

- Es wird auf der Grundlage von First In First Out (FIFO) gearbeitet – die Schiffe werden also in der Reihenfolge bearbeitet, in der sie in das System gekommen sind.
- Die „Mitarbeiter“ dürfen einander nicht helfen und der Prozess darf nicht verbessert werden, da in beiden Durchgängen Messdaten erhoben werden, die nach dem Experiment ausgewertet werden – es geht also um Vergleichbarkeit.
- Es wird sehr schnell gearbeitet, da wir den Output des Systems maximieren wollen. Da jedes Schiff um 10 Euro verkauft werden kann, wollen wir in beiden Durchgängen so viel Geld wie möglich verdienen.

Setup für Schiffe falten für eine Gruppe mit 10 Teilnehmern



Material
pro Gruppe: □ 100 Blatt Papier weiss
■ 2 Blatt Papier rot

Falt-Anleitung: <https://youtu.be/swRaeAzfirw>

Bild 1.3 Setup für das Flussexperiment „Schiffe falten“

Push it! Diese Situation kennen alle bestens: Es gibt unendlich viel Arbeit zu erledigen, die Sales-Kollegen ziehen einen Auftrag nach dem anderen an Land. *Ursula* gibt bei meinem heutigen Training die Sales-Managerin: Sie faltet Blatt für Blatt einmal durch die Hälfte und schiebt im ersten Durchgang des Flow-Experiments ohne Pause das Papier ins System. Am Anfang ist alles noch recht entspannt, denn *Harald*, der Nächste in der Reihe, muss nicht mehr machen, als das Blatt ein zweites Mal durch die Hälfte zu falten. Er ist flink und schiebt die gefalteten Blätter etwa eine halbe Minute recht flott zu *Anna* weiter, deren Aktivität schon ein wenig komplexer ist: Sie muss die Ecken zur Mitte falten. Auch *Lisa* führt ihre Aufgaben so schnell wie möglich aus. Doch allmählich beginnen sich in jeder Aktivität kleine Zwischenlager aufzubauen und je größer die Häufchen werden, desto seltener wird das Kichern. Stattdessen macht sich im Raum eine angestrengte Stille breit. Alle falten hastig an ihren Schiffchen, um die Stapel so schnell wie möglich abzubauen. Aber während ein Schiff noch nicht mal richtig fertig ist, landen schon wieder zwei weitere am Stapel. Am härtesten trifft es *Tom*, der am Ende der Reihe sitzt und gleich drei Aktivitäten ausführen muss. „Hoffentlich brichst du dir nicht die Hände“, versuche ich zu scherzen, denn er faltet und faltet und faltet. In den ersten paar Sekunden hält er mit dem Tempo der anderen noch ganz gut mit. Doch plötzlich liegen zwei Schiffe in seiner Warteschleife, drei, fünf ... und schließlich ein unüberschaubarer Haufen. Ein Engpass wie aus dem Bilderbuch. Tja, Pech für Tom – die anderen sind selbst viel zu beschäftigt, niemand hat Zeit, um ihm zu helfen. Da hat es *Jürgen* leichter: Er misst, wie viele Schiffe den Arbeitsprozess pro Minute fertig verlassen. Er startet am Beginn des Experiments die Stoppuhr und jedes Mal, wenn ein Schiff die letzte Aktivität verlässt, notiert er sich die dazugehörige Zeit.

Nach zwei Minuten schicke ich ein rotes Blatt in die Mini-Reederei und kündige gleichzeitig an, dass dies das letzte Schiff ist. Allmählich schließt einer nach dem anderen die Arbeit ab, nur Tom muss noch einige Zeit weiterfalten. „Vergesst nicht, dass ihr in der Realität weiterarbeiten würdet, da immer wieder neue Kundenaufträge reinkommen“, gebe ich ihnen in der Zwischenzeit eine Denkaufgabe. Endlich schiebt Tom das rote Schiff weiter zu Jürgen. Hier rufe ich Stopp. Jürgen misst die Durchlaufzeit des roten Schiffs:

Ankunftszeit – 2 Minuten = Durchlaufzeit

Ankunftszeit 5:18 – nach einer Durchlaufzeit von 3:18 Minuten ist das rote Schiff also fertig für den Stapellauf. Kurze Verschnaufpause, Tom kann seine Hände ausschütteln, dann geht es in den zweiten Durchgang.

Pull it! Der zweite Durchgang beginnt unter anderen Vorzeichen. In jeder Aktivität darf sich nur jeweils ein Schiff in Arbeit befinden – es gilt das WIP-Limit 1. Ursula hält sich als Sales-Managerin vorbildlich an die Vorgabe, sitzt ruhig vor ihrem Papierstapel, faltet das erste Blatt und hebt anschließend die Hände. Harald zieht das halbierte Blatt, faltet es noch einmal durch die Hälfte und hebt seine Hände, als er damit fertig ist. Das ist das Zeichen für Anna: Sie zieht das geviertelte Blatt zu sich, faltet die beiden Ecken nach unten und hebt die Hände. Ein wenig erinnert es mich an die Welle in einem Fußballstadion, als sich in der Reihe allmählich ein Rhythmus aus „falten – Hände heben – nächstes Schiff holen“ aufbaut. Tom, der im Prozess-Bottleneck am Ende der Reihe sitzt, faltet ebenfalls gleichmäßig vor sich hin und hat weniger Stress, weil sich keine Schiffe vor ihm auftürmen. Aber weil er drei Faltschritte statt nur einem durchführen muss, bestimmt sein Tempo die Geschwindigkeit der ganzen Reihe. Also heißt es ab einem bestimmten Zeitpunkt für alle: warten. Es ist spannend, die Teilnehmer und Teilnehmerinnen dabei zu beobachten, was sie während dieser Wartezeit machen. Einige verfolgen das Geschehen ganz genau, manche denken nach, andere wiederum nutzen die Pause für Qualitätsverbesserungen und streichen noch einmal sorgfältig über die Falzungen. Nach zwei Minuten schicke ich wieder ein rotes Blatt in das System: Es kommt um 2:45 Minuten bei Jürgen an – die Durchlaufzeit beträgt somit 0:45 Sekunden. Stopp.

1.1.1 Wenn wir schneller arbeiten, wird trotzdem nicht mehr Arbeit fertig

Rekapitulieren wir kurz: Im ersten Durchgang nach dem Push-Prinzip arbeiteten alle wie verrückt. Im zweiten Durchgang hatten einige Leute sogar Zeit genug, um einfach in die Luft zu schauen oder ihre unfertigen Schiffchen zu verbessern. Das kann doch nur bedeuten, dass der zweite Durchgang nicht sehr produktiv war. Oder?

Als ich nach dem Pull-Durchgang in die Runde frage, wie es den Teilnehmerinnen und Teilnehmern ergangen ist, bekomme ich folgende Antworten:

- Es war entspannter.
- Ich konnte mich konzentrieren.
- Wir konnten an der Qualität arbeiten.

Das ist schon mal interessant, aber jetzt kommt noch ein spannender Moment: Wir werten gemeinsam Jürgens Messungen aus. Ich zeichne am Flipchart eine Tabelle: Wie viele Schiffe sind in den beiden Durchgängen in jeder Minute bei Jürgen angekommen? Das jeweils erste Ergebnis des ersten und zweiten Durchgangs streichen wir, weil das System zu diesen Zeitpunkten noch nicht zur Gänze befüllt war. Das letzte Ergebnis des ersten Durchgangs streichen wir auch, weil bei der Ankunft des letzten Schiffs die Minute noch nicht voll war. Da wir den durchschnittlichen Durchsatz pro Minute berechnen wollen, würde es das Resultat verfälschen. Wir berechnen also den durchschnittlichen Durchsatz pro Minute (Tabelle 1.1).

Tabelle 1.1 Schiff-Durchsatz pro Minute

Minute	Push und ohne Limit	Pull mit WIP 1
0	3	3
1	5	6
2	6	5
3	6	
4	5	
5	2	
Ø Durchsatz	5,5 Schiffe pro Minute	5,5 Schiffe pro Minute
Durchlaufzeit rotes Schiff	3:18	0:45

Ich mag diesen Moment, wenn ich die Fragezeichen regelrecht hochschießen sehe. Wie kommen solche Ergebnisse zustande?

- In beiden Durchgängen verließen pro Minute rund fünf Schiffe das System.
- Den Pull-Durchgang haben wir früher abgeschlossen, weil das rote Schiff nach 2:45 Minuten fertig gefaltet war – im ersten Durchgang hingegen erst nach 5:18 Minuten. Die Durchlaufzeit hat sich um den Faktor 4 von 3:18 auf 0:45 reduziert.
- Stress im ersten Durchgang, „Slack“ – also freie Zeit – im zweiten Durchgang.

Die erste Erkenntnis kommt meistens sehr schnell, wenn ich einen gezielten Blick an das Ende unserer Mini-Reederei werfe: Richtig, dort gab es einen Engpass im Prozess (wohl-gemerkt: im *Prozess* – Menschen sind meistens nicht die Engpässe).



Ob wir in einem Push- oder in einem WIP-limitierten Pull-System arbeiten: **Der Engpass eines Systems bestimmt den Durchsatz.** Spätestens dort bleibt der Arbeitsfluss stecken. Einfach schneller zu arbeiten, ist also keine Lösung, denn völlig egal, wie schnell wir arbeiten: Am Ende des Tages schaffen wir immer gleich viel.

Für Verantwortliche in Produktionsbetrieben gehört das zum kleinen Einmaleins. Genauso war beim Falten der Schiffe – also in der „Produktion“ – dieser Engpass sehr leicht zu identifizieren. Auch in der Wissensarbeit ist das Konzept der Engpässe bekannt. Aber meistens fällt es zum einen schwer, einen Engpass überhaupt einwandfrei zu identifizieren. Das liegt unter anderem daran, dass Engpässe in der Wissensarbeit meistens nur temporär an einer Stelle auftreten und dann zu einer anderen weiterwandern, weil in der Wissensarbeit nicht immer das Gleiche ausgeführt wird und Aufträge nie völlig identisch sind.

Meistens wird deshalb auch zu viel Energie darin investiert, vermeintliche Engpässe zu beseitigen. Die offensichtlichste Lösung wird dabei meistens außer Acht gelassen: Man kann den Zufluss an Arbeit beschränken und frei werdende Ressourcen dafür nutzen, um den Engpass zu erweitern. Näheres zu Engpässen in Abschnitt 2.4.

1.1.2 Wir haben genug Zeit für Aufgaben, für die wir nie Zeit haben

Betrachten wir es wirtschaftlich: Wenn wir jedes Schiff für zehn Euro verkaufen können, verdienen wir nach beiden Prinzipien – sowohl Push als auch Pull – gleich viel, nämlich $5,5 \times € 10 = € 55$ pro Minute. Denken wir an die übliche Situation in einem Unternehmen: Obwohl die Menschen im System bereits hart an den Leistungsgrenzen arbeiten, wird Arbeit noch und nöcher in das System gepresst. Ironischerweise feiern viele Unternehmen das als hohe Auslastung. Weil aber der Engpass den Durchsatz eines Systems bestimmt, ist es vollkommen sinnlos, den Druck ständig zu erhöhen: Das System wird dadurch weder produktiver noch schneller. Zusätzlich wird in Push-Systemen durch „Zwischenlager“ viel Kapital im System gebunden. In der Wissensarbeit fehlt die Sensibilität dafür, weil diese Arbeit größtenteils unsichtbar ist. Meistens handelt es sich dabei um Dateien auf Festplatten, an denen nicht weitergearbeitet wird.

W. Edward Deming würde jetzt sagen (Deming 2000): „It's not about working harder, it's about working smarter.“ Und genau das passiert im limitierten Pull-System: Auch hier fließt die Arbeit nur so schnell durch das System, wie es der Engpass zulässt. Aber erstens können die Menschen im System mit ihren Kräften haushalten und zweitens bleibt sogar noch Zeit übrig (was ebenso ironischerweise reflexartig als schlechtes Zeichen gedeutet wird). „Also“, frage ich in die Runde, „wie könnte man die Zeit nutzen, in der alle auf das Freiwerden des Engpasses warten?“

1. **Wir könnten dem Kollegen im Engpass helfen.** In der Praxis lautet einer der häufigsten Einwände gegen diese Idee: „Aber wir sind doch keine Spezialisten für die Engpassarbeit.“ Die gute Nachricht lautet: Darauf kommt es in den meisten Fällen überhaupt nicht an! Es geht um Entlastung und die kann auf unterschiedlichste Art und Weise passieren: administrative Aufgaben übernehmen, Vertretung in Meetings – Helfer müssen keine Spezialistentätigkeiten ausführen. Entscheidend ist, dass sich die Kollegen im Engpass ganz auf ihre Aufgaben konzentrieren und sie erledigen können. Was ist die Konsequenz? Auch wenn ein Nichtspezialist im Bottleneck hilft, geht der Durchsatz nach oben (siehe dazu auch Abschnitt 2.4) – vielleicht langsamer als durch den Einsatz eines Spezialisten (der gerade nicht verfügbar ist), aber er geht nach oben.
2. **Ist das überhaupt ein optimaler Prozess?** Wenn in einem Bereich ständig alle auf Hochtouren laufen und die Aufgaben dennoch nur langsam abgeschlossen werden, stellt sich die grundsätzliche Frage nach der Balance. Möglicherweise ist die Art und Weise, wie die Arbeit derzeit verteilt ist und wie gearbeitet wird, nicht ideal. In unserem Experiment ist das Ungleichgewicht deutlich: Während in jedem anderen Prozessschritt nur eine Aufgabe erledigt werden muss, sind im Engpass drei Aufgaben zu erledigen. Die Verteilung der Arbeit würde den Engpass auflösen, der Durchsatz ginge wieder nach oben.

Mehr Durchsatz bedeutet: Im selben Zeitraum kann mehr geliefert werden. Und wie wir im Experiment gesehen haben, kann Slack (freie Zeit) auch für die Arbeit an der Qualität genutzt werden. Einer dürfte sich besonders freuen, wenn mehr in besserer Qualität geliefert werden kann: der Kunde. Das Schöne daran ist, dass ein Unternehmen für Verbesserungen dieser Art, die durch die Zeit für Überlegungen im Prozess entstehen, nichts bezahlen muss. Es muss weder teure Prozessingenieure noch Berater wie mich anheuern, die das Nachdenken übernehmen.



Slack bietet die Gelegenheit, bei gleichem Output die Schwachpunkte des Systems zu erkennen und zu untersuchen sowie an der Qualität zu arbeiten und Möglichkeiten zur Verbesserung zu finden. In einem nicht limitierten System gibt es diese Gelegenheit wegen chronischer Überlastung nicht.

1.1.3 Wir werden verlässlicher, wenn wir uns Grenzen setzen

Die Stabilität und damit die Verlässlichkeit eines Systems sind unmittelbar mit einem der unbeliebtesten Themen der Geschäftswelt verbunden: mit dem Schätzen. In überladenen klassischen Push-Systemen mit – juhu – voll ausgelasteten Mitarbeitern, ist immer jemand mit der undankbaren Aufgabe betraut, einen Projektplan erstellen zu müssen. Also geht er zu den einzelnen Personen, Teams oder Abteilungen hin und fragt alle reihum: „Ich hätte da was. Wann seid ihr fertig?“ Jeder würfelt eine Zahl und schlägt einen Puffer von 100 Prozent auf. Der geübte Projektmanager addiert die Teilschätzungen und multipliziert sie vorsichtshalber mal drei, um endlich zum Termin zu gelangen ... an dem das Produkt oder Projekt sicher nicht fertig sein wird. Und weil das Ziel so oft verfehlt wird, hält der Projektmanager seine Kollegen in einem System, in dem ohnehin schon nichts mehr fertig wird, regelmäßig mit der Frage „Wie weit seid ihr?“ von der Arbeit ab. Diese Frage äußert sich meistens in Kontrollmaßnahmen wie Statusreports, die zusätzliche Arbeit verursachen und die Mitarbeiter von ihrer eigentlichen Arbeit abhalten. Dadurch verschärft sich das Problem weiter. Außerdem werden die Mitarbeiter in eine Schätz-Schulung gesteckt, damit sie endlich lernen, wie man es richtig macht. Das Problem ist nur, dass die Mitarbeiter nicht das Problem sind. **Die Ursache von Fehlschätzungen ist, dass sich die Anzahl der Arbeiten im System ständig ändert. Eine verlässliche Vorhersagemethode, egal welcher Art, funktioniert nur auf Basis eines stabilen Systems.**

Niemand kann in einem unlimitierten System, in das ständig neue Aufgaben gelangen, die sofort bearbeitet werden sollen, eine valide Schätzung abgeben, wann welche Arbeit fertig wird. Dazu kommt, dass Menschen auf das Schätzen in Aufwänden konditioniert sind. Sie stellen sich grundsätzlich die Frage „Wie lange brauche ich dafür?“ statt „Wann ist es fertig?“, was auf die Durchlaufzeit abzielen würde. Ob das Arbeitssystem eines Unternehmens verstopft ist, erkenne ich übrigens schon an den ersten Gesprächen, in denen es lediglich um das Finden eines Termins geht: „Da haben wir keine Zeit ... da geht es nicht ... da müssten wir mal schauen ...“ Je länger die Terminfindung dauert, desto höher ist die Wahrscheinlichkeit, dass es sich um ein chronisch überlastetes Unternehmen handelt.

Denken wir noch einmal an das Flussexperiment: Was wäre passiert, wenn ich das rote Schiff nicht nach zwei Minuten, sondern erst nach 15 Minuten in Auftrag gegeben hätte?

- Im Push-System hätte sich in dieser Zeit bereits ein massiver Rückstau gebildet und wir könnten keine Aussage darüber treffen, wann das rote Schiff vom Stapel laufen wird. Wir würden lediglich wissen, dass es wesentlich länger dauern wird – aber wie lange genau? Mit anderen Worten: Dieses System ist *instabil*. Der Kunde bekommt keine verlässliche Aussage darüber, ab wann er mit seinem Schiff in See stechen kann.
- Im Pull-System ist die Antwort einfach: Wir wissen, wie viel Arbeit im System ist. Das System ist *stabil*. Daher können wir zum Beispiel durch Schätzungen oder Forecastings

sagen, wann die Arbeit fertig wird. Wenden wir beim Falten der Schiffe das Pull-System an, verändert sich die Durchlaufzeit nicht. Wegen dieser Stabilität kann die Reederei selbst nach 15 Minuten ihrem Kunden eine verlässliche Auskunft darüber geben, wann sein Schiff fertig wird. In der Wissensarbeit funktioniert das nicht ganz so einfach, da nicht immer das Gleiche gebaut oder entwickelt wird, sondern der Umfang der Arbeit variiert. Trotzdem ist ein limitiertes System auch in der Wissensarbeit die Voraussetzung, um sagen zu können, bis wann eine Arbeit spätestens fertiggestellt werden kann (siehe Kapitel 4).



Sowohl für den Erbringer einer Leistung als auch für den Kunden ist das wohl einer der größten Vorteile von **WIP-Limits**: Verlässlichkeit. Termintreue ist nur möglich, wenn die Zahl der begonnenen Arbeiten begrenzt und das System dadurch stabilisiert wird.

1.1.4 Wenn alles Priorität hat, hat nichts Priorität

Nehmen wir an, in unserem instabilen Push-System gibt der Kunde das rote Schiff nach 15 Minuten in Auftrag. Dadurch schnellst die Durchlaufzeit in die Höhe, nach einer Stunde ist das Schiff gerade mal zu 80 Prozent fertig. Und nun passiert, was so oder ähnlich in der Realität jeden Tag in Unternehmen passiert:

Kunde: „Ich könnte das Schiff in zwei Minuten falten und Sie erzählen mir, dass Sie nach einer Stunde noch immer nicht fertig sind? Geben Sie gefälligst Gas!“

Manager: „Sie wissen, dass bei uns der Kunde im Mittelpunkt steht.“

Und das bedeutet so viel wie: Das rote Schiff hat ab sofort höchste Priorität und wird an allen anderen in Bau befindlichen Schiffen vorbei durch das System gereicht. Pech für die Abnehmer der weißen Schiffe, die müssen nun länger warten. Gott sei Dank, das rote Schiff wird vor allen anderen fertiggestellt. Der zappelige Kunde ist halbwegs zufrieden.

Nun hat ein Unternehmen meistens mehr als einen Kunden. In unlimitierten Push-Systemen steigt damit die Wahrscheinlichkeit, dass mehrere unzufriedene, wartende Kunden ihre Aufträge nach vorne schieben wollen und irgendwann schwimmen unzählige rote Schiffe durchs System. Was passiert?

- Der Vorteil der Priorität schlägt ab einem gewissen Punkt in das Gegenteil um. Sollen mehr hoch priorisierte Schiffe gebaut werden, werden die normal eingestuft Schiffe noch später fertiggestellt, als es wegen der Staus in den einzelnen Prozessschritten ohnehin schon der Fall ist. Aber nicht nur das: Auch alle roten Schiffe werden allmählich ausgebremst.
- Gibt es irgendwann nur noch hoch priore Schiffe im System, heben sich die Priorisierungen auf. Wenn alles priorisiert ist, ist nichts priorisiert. Alle roten Schiffe haben nun die gleiche Durchlaufzeit, wie die weißen Schiffe sie ursprünglich hatten. Und die weißen Schiffe werden vermutlich nie mehr fertig – zumindest solange nicht, bis jemand ihre Priorität ganz nach oben stuft. Das gesamte System wird wieder langsamer.

Man möchte nicht für möglich halten, welche grotesken Auswüchse das unkontrollierte Priorisieren in der Realität annehmen kann. In vielen Unternehmen erlebe ich, dass es nicht nur Priorität 1 gibt, sondern auch die Prioritäten 1+ und 1++. Ist ein Unternehmen in diesem Stadium angelangt, werden als nächste Priorisierungsstufe Task Forces gegründet, die wiederum priorisiert werden – eine faszinierende Endlosschleife.



Sollen für die Kunden **Prioritäten** gewährleistet werden, müssen sowohl das System als auch die Prioritäten limitiert werden.

1.1.5 Je später wir anfangen, desto besser für den Kunden

Bleiben wir noch etwas bei unserem besonderen Kunden, der das rote Schiff nach 15 Minuten in Auftrag gegeben hat. Weil die Reederei so lange nicht geliefert hat und sich die Welt für ihn in der Zwischenzeit weitergedreht hat, will der Kunde ab Aktivität Nr. 4 etwas Neues. Er will jetzt ein grünes Schiff. Erinnern wir uns noch einmal an das Experiment: Im ersten Durchgang kam das rote Schiff nach zwei Minuten ins System und hatte eine Durchlaufzeit von 3:18 Minuten. Wäre das Schiff nach 15 Minuten in das System gelangt, hätte die Durchlaufzeit bereits eine knappe Stunde betragen. Das heißt, je länger man in einem instabilen System arbeitet, desto größer werden die Durchlaufzeiten. Im zweiten Durchgang haben wir hingegen gesehen, dass die Durchlaufzeit konstant geblieben ist, weil zuerst die bereits im System befindlichen Schiffe fertiggebaut wurden, bevor etwas Neues begonnen wurde.

In einem Push-System würde nun Folgendes passieren: Das rote Schiff wird aus dem Prozess genommen. Stattdessen wird ein grünes Schiff wieder ganz am Anfang des Prozesses in Auftrag gegeben und damit in ein System eingespeist, das ohnehin schon überlastet ist. Abgesehen davon wurde gerade die gesamte bisher am roten Schiff geleistete Arbeit vernichtet. Das schmerzt in den Bilanzen.

Was wäre in einem WIP-limitierten Pull-System passiert? Nichts. Die Arbeiten am roten Schiff hätten zum Zeitpunkt der Meinungsänderung noch gar nicht begonnen und somit wäre auch kein finanzieller Verlust entstanden.

Eines der grundlegenden Prinzipien eines limitierten Systems ist, dass die Arbeit so lange wie möglich außerhalb des Systems gehalten wird (Late Commitment). Das hat einen einfachen Grund: Wird eine Arbeit begonnen, soll sie auch so schnell wie möglich fertiggestellt werden. Geplante Arbeiten werden daher nicht im System aufgestaut (das würde die Durchlaufzeiten erhöhen), sondern *davor* – im Fall von Kanban im sogenannten Optionenpool. Für das Arbeitssystem hat das gravierende Effekte:

- Überlegt es sich der Kunde noch einmal anders, sinkt durch die späte Zusage die Wahrscheinlichkeit, dass bereits geleistete Arbeit vernichtet wird.
- Je niedriger das WIP-Limit in einem System angesetzt ist, desto kürzer sind die Durchlaufzeiten. Durch kurze Durchlaufzeiten sinkt die Wahrscheinlichkeit, dass der Kunde während der Bearbeitung etwas an seinem Auftrag ändern muss oder ändern will.



Mit einem limitierten System streben wir danach, die Durchlaufzeiten zu verkürzen und gleichzeitig die Zusage für eine Arbeit so spät wie möglich zu machen (Late Commitment). Dadurch sinkt das Risiko, dass Auftragsänderungen zu einem späteren Zeitpunkt die bereits geleistete Arbeit vernichten.

1.1.6 Lokale Optimierung führt zu globaler Suboptimierung

Stellen wir uns vor, in den einzelnen Aktivitäten unserer unlimitierten Push-Papierreederei säßen nicht einzelne Personen, sondern Teams. Während alles andere beim Alten bleibt, beschließt nun das Management, eine Teamoptimierung durchzuführen. Die zwei ersten Teams werden dafür ausgewählt, sie sollen mit der neuen Teammethode schneller liefern als bisher. Zunächst geht die Rechnung auf: Jede Woche übertreffen sich die Teams aufs Neue in ihren Leistungen, ihr Durchsatz steigt an, die Durchlaufzeiten verkürzen sich in den ersten beiden Aktivitäten des Prozesses. Fröhlich pumpt das Management noch mehr Aufträge ins System. Eigentlich ein Grund, zufrieden zu sein. Aber warum steigt plötzlich die Durchlaufzeit des Gesamtsystems?

Geblendet durch die lokal begrenzte Leistungssteigerung wird meistens genau das Falsche getan: Wird noch mehr Arbeit ins System geschaufelt, erhöht sich damit der Work in Progress (WIP) im gesamten System. Ein höherer Work in Progress bedeutet: längere Durchlaufzeiten. Besteht außerdem irgendwo im System ein Engpass (und diesen Engpass gibt es immer), bestimmt wiederum der Engpass den Gesamtdurchsatz und die Gesamtdurchlaufzeit eines Arbeitssystems. Die lokale Verbesserung einzelner Teams hat zur Folge, dass das gesamte Arbeitssystem langsamer wird.

Der Glaube an die positive Wirkung lokaler Optimierungen ist in den meisten Unternehmenskulturen fest verankert. Gerade in der agilen Gedankenwelt steht das Team im Zentrum der Optimierung, weil der Wertstrom durch die Besetzung mit Vertretern aller nötigen Disziplinen in einem einzigen Team abgebildet werden soll. Das verführt zu Trugschlüssen, weil in den seltensten Fällen ein einziges Team die gesamte Wertschöpfungskette abdecken kann. Besonders deutlich wird das anhand des hochaktuellen Themas der agilen Skalierung: Viele Unternehmen, die agil werden wollen, sind der Meinung, man müsse lediglich alle Teams in agile Methoden einschulen – und schon hat man ein agiles Unternehmen. Mitnichten. Ich bin ein Fan des Organisationstheoretikers Russell L. Ackoff, von dem die zeitlos gültige Aussage stammt:

„And therefore, the performance of the whole is never the sum of the parts taken separately, but it's the product of their interactions.“ (Ackoff, Gharajedaghi 1984)

Wenn nicht gerade die neueste Arbeitsverbesserungswunderwaffe getestet wird (z.B. Scrum oder Kanban), manifestiert sich der falsche Glaube an die lokale Optimierung meistens in Form von Zielvorgaben für Einzelpersonen oder Teams, in Boni und anderen Gratifikationen. Mit diesen indirekten Optimierungsversuchen wird in Performancesystemen aber lediglich der Druck erhöht, der im besten Fall keine Auswirkung auf die Gesamtleistung hat und im Normalfall mit einer Leistungsverschlechterung des Gesamtsystems einhergeht. Wie groß wird die Motivation sein, einem Kollegen im Engpass (in einem anderen Team oder einer anderen Abteilung) zu helfen, wenn das Team ausschließlich am Erreichen

seiner eigenen, lokalen Ziele gemessen wird? Völlig unbeachtet bleibt dabei die Tatsache, dass ein Team *nur ein* Element der Wertgenerierung auf dem Weg zum Kunden ist. Lokale Optimierungen sind für mich daher ein klarer Fall von Fehlsichtigkeit: Es fehlt die Sicht auf das System in seiner Gesamtheit. Betrachten wir lokale Optimierungen aus der Sicht eines Kunden: Für ihn ist es völlig irrelevant, ob Team 1 und 2 nun schneller arbeiten. Bezogen auf das Beispiel des Schiffefaltens generiert diese Form der Leistungssteigerung für den Kunden keinerlei Wert, weil er a) ohnehin auf ein ganzes Schiff wartet und nicht auf teiloptimierte Komponenten und b) auf dieses ganze Schiff nun länger warten muss, weil Team 1 und 2 mit ihrer Performance das gesamte System ausbremsen. Lokale Optimierung führt zu globaler Suboptimierung.

Vielleicht verstehen Sie nun, warum sich in mir alles sträubt, wenn Kanban als Teammethode bezeichnet wird. Setzen Sie Kanban ausschließlich für einzelne Teams oder Abteilungen ein, ist das in vielen Fällen ein Schuss ins Knie. Kanban entfaltet sein Potenzial am besten, wenn man sich in Veränderungsinitiativen von der Teamebene löst und mit den Optimierungsbetrachtungen auf der höchst möglichen Ebene beginnt (mehr dazu beim Thema Kanban Flight Levels in Abschnitt 1.4). Visualisierung ist daher mit gutem Grund eine zentrale Praktik von Kanban. Vor allem zu Beginn einer Veränderungsinitiative hilft die Visualisierung, so viel wie möglich von einem System zu sehen und damit zwei wesentliche Fragen zu beantworten:

1. Wie generieren wir Wert für unseren Kunden?
2. Wo liegen die Probleme in unserem Gesamtsystem?

Hätte sich das Management unserer fiktiven Reederei zunächst einen Überblick verschafft, wäre deutlich geworden: Eigentlich ist der Engpass jene Stelle im Prozess, an der es mit Optimierungen ansetzen sollte. Im Schiffsexperiment drängt sich der Eindruck auf, eine Person sei das Bottleneck. In der Realität wird genauso gehandelt: Wenn Menschen in einem System arbeiten, dann müssten wir doch bei ihnen mit den Optimierungen ansetzen, oder? Daher zielen Schulungen und Trainings hauptsächlich darauf ab, Einzelpersonen schneller zu machen – nicht das System.

Natürlich hängt es von den Fähigkeiten ab, wie gut ein Mensch seine Aufgaben erfüllen kann. Nun habe ich dieses Experiment aber bereits Hunderte Male auf verschiedenen Kontinenten mit den unterschiedlichsten Menschen durchgeführt. Egal, wo und mit wem ich das Experiment mache: Von wenigen Ausreißern nach oben und unten abgesehen, pendelt der Durchsatz immer zwischen fünf bis sechs Schiffen pro Minute – viel mehr und viel weniger ist einfach nicht drin. Faszinierend, oder? Es liegt nicht an den Faltfertigkeiten der Probanden, dass die Schiffe durch ihre Aktivitäten langsamer fließen als durch alle anderen. Es liegt daran, dass die Probanden in diesen Aktivitäten mehr Faltungen vornehmen müssen als alle anderen – das Arbeitssystem ist also das Problem. Nicht nur im Schiffsexperiment, sondern natürlich auch bei meiner Arbeit mit Unternehmen sehe ich, dass die Fähigkeiten von Menschen einen recht geringen Anteil am Gesamtoutput einer Organisation haben (außer, diese Fähigkeit äußert sich in einer besonderen Spezialisierung dieser Menschen). Beim genaueren Hinsehen zeigen sich die weit größeren Schwächen immer in den Arbeitsprozessen, nicht in den Menschen.

Das Ziel von Kanban ist daher nie die lokale Optimierung einzelner Bestandteile eines Systems, sondern die Verbesserung des Wertstroms vom Zeitpunkt des Commitment (Zusage einer Arbeit) bis zur Lieferung an den Kunden.



Lokale Optimierung führt in den meisten Fällen zu einem schlechteren Ergebnis des Gesamtsystems. Um die Leistung eines Systems zu verbessern, muss der Fokus auf den Interaktionen zwischen den einzelnen Systemeinheiten liegen.

Am Ende des Schiffsexperiments, bevor ich tiefer in die Prinzipien und Praktiken von Kanban einsteige, frage ich die Teilnehmerinnen und Teilnehmer immer, was sie aus dieser Übung mitnehmen und welche Schlüsse sie für sich ziehen. Also lasse ich sie auch an dieser Stelle zu Wort kommen. Hier eine kleine Sammlung von Antworten.



- Man sollte die (gelebten) Prozesse kennen.
- Der Durchsatz wird vom Engpass bestimmt.
- Was ist dem Kunden wirklich wichtig?
- WIP-Limits einführen und beobachten: Man muss sich an das am besten passende WIP herantasten.
- Priorisierungen sollten durchgeführt werden, *bevor* die Arbeit in das System gelangt und nicht dann, wenn sie sich bereits im System befindet.
- Aktives Management fördern und fordern
- Ohne WIP-Limits haben wir keine Zeit für Optimierungen.
- Mit WIP-Limit haben wir weniger Stress bei gleichem Output/Durchsatz.
- Mit WIP-Limit kann man auch spät noch auf neue Anforderungen reagieren.
- WIP-Limits führen zu vorhersehbareren und schnelleren Durchlaufzeiten.
- Auch Prioritäten müssen limitiert werden.
- Der Fokus sollte auf dem Gesamtsystem liegen, nicht auf Teilen davon.

■ 1.2 Kanban: Viva la evolución!

Das Schiffsexperiment zielt darauf ab, die Unterschiede zwischen zwei Systemen aufzuzeigen: In dem einen System kommt nie ein gleichmäßiger und kontinuierlicher Arbeitsfluss zustande, weil der Rhythmus durch Druck von außen bestimmt wird. Das andere System erzielt gleiche und sogar bessere Ergebnisse, weil es den *Zufluss* an Aufträgen *reguliert* und genau dadurch einen kontinuierlichen Arbeitsfluss etabliert. Es ist ein effektives System und paradoxerweise gibt es in einem effektiven – flussbasierten – System vieles, das absolut nicht intuitiv klingt. Einiges davon haben wir im Schiffsexperiment – das eigentlich ein Flussexperiment ist – bereits erkannt:

- Wenn wir schneller arbeiten, wird trotzdem nicht mehr Arbeit fertig.
- Wir haben genug Zeit für Aufgaben, für die wir nie Zeit haben.
- Wir werden verlässlicher, wenn wir uns Grenzen setzen.

- Wenn alles Priorität hat, hat nichts Priorität.
- Je später wir anfangen, desto besser für den Kunden.
- Lokale Optimierung führt zu globaler Suboptimierung.

An dieser Stelle möchte ich noch einmal auf den Einwand zurückkommen, den einige Teilnehmerinnen und Teilnehmer in den Diskussionen rund um das Flussexperiment meistens einbringen: Wenn wir in unserem Fall über die Anwendung von Kanban sprechen, dann reden wir nicht von industrieller Fertigung, sondern von über weite Strecken unsichtbarer Wissensarbeit. Wie mittlerweile allgemein bekannt ist, wurde der Begriff „Kanban“ durch das von Taiichi Ohno entwickelte pull-basierte Toyota Production System berühmt und bedeutet so viel wie „Signalkarte“. Damit wird in der Automobilfertigung angezeigt, dass eine Fertigungsstufe ihre Arbeit abgeschlossen hat und Materialnachschub entgegennehmen kann. Tatsächlich sind Kanban-Systeme aber keine Erfindung des 20. Jahrhunderts: Im kaiserlichen Garten in Tokio wird die Besuchermenge mit kleinen Plättchen geregelt, die am Eingang ausgeteilt und am Ausgang wieder eingesammelt werden (vgl. Anderson 2010). So sind immer genau so viele Besucher unterwegs, wie das System kaiserlicher Garten verträgt, um den Menschen Erholung und Freude zu bieten.

Nun haben wir es in der Wissensarbeit mit wenig Greifbarem zu tun. Wissen liegt nicht in Halbfertigteilen herum und wir können die Auslastung eines Wissensarbeitssystems auch nicht daran messen, wie viele Mitarbeiter sich gerade in einem Raum tummeln. Die Praktiken von Toyota können also nicht direkt in die Wissensarbeit transferiert werden, es gibt jedoch systemunabhängige Prinzipien *hinter* den Praktiken, die in jedem Arbeitssystem anwendbar sind. Jedes System hat nur eine gewisse Kapazität und so ist das Limitieren der begonnenen Arbeiten in einem System beim Nachdenken über eine neue Software genauso sinnvoll wie bei handwerklichen Tätigkeiten oder in der industriellen Fertigung. Auch Engpässe gibt es in jedem System dieser Welt. Ob im kaiserlichen Garten von Tokio, in den Fertigungshallen von Toyota oder bei Kanban in der Wissensarbeit: Der springende Punkt in allen diesen Systemen ist

- a) der Blick auf die Generierung eines Werts für den Kunden und
- b) der daraus resultierende Wille, das eigene System beweglich zu halten und stets Ausschau nach Verbesserungen zu halten.

Kanban in der Wissensarbeit bietet daher kein starres Schema an, denn Schemata zielen immer darauf ab, etwas Bestimmtes zu liefern. Damit könnte auf Dauer gesehen niemand wirklich erfolgreich werden, denn keine zwei Unternehmen bzw. Unternehmenssysteme sind in allen ihren Details miteinander vergleichbar. Wenn wir uns daran erinnern, dass jedes Arbeitssystem im Grunde die Wünsche seines Kunden erfüllen will bzw. sollte, dann sind rezepthafte Vorgehensweisen ohnehin fehl am Platz. Kanban ist kein vordefinierter und standardisierter Prozess. Kanban ist in der Wissensarbeit immer ein maßgeschneideres System, weil es offen lässt, wie es gestaltet wird, und keinen Zielzustand vorschreibt. Es setzt auf den evolutionären Gedanken: Beginne mit dem, was du momentan tust, und verbessere es anhand dessen, was dir die Anwendung der Prinzipien und Praktiken von Kanban offenbart. So formt sich im Laufe der Zeit der für das einzelne Unternehmen jeweils passende Prozess heraus.

Das ist das völlige Gegenteil von dem, was üblicherweise in Unternehmen passiert: In künstlich initiierten Veränderungsprozessen meinen Personen, die oft gar nicht in den

Arbeitsalltag involviert sind, für die Kolleginnen und Kollegen eine optimale Arbeitsweise neu erfinden zu müssen, um das Unternehmen fit für den Markt zu machen. Sobald die neue Arbeitsweise installiert ist, wird das Ziel als erreicht betrachtet.

Prinzipien und Praktiken

Kanban gibt sich in dieser Hinsicht bescheidener: Wäre in einem Unternehmen tatsächlich alles schlecht und miserabel, wäre es vom Markt bereits aussortiert worden. Daher gesteht Kanban dem, was derzeit vorhanden ist, seine Berechtigung zu, analysiert die aktuelle Situation und baut darauf auf. Dazu gehört auch, nicht zu wissen, wie Mitarbeiter arbeiten sollen. Kanban hilft aber, schrittweise zu *entdecken*, wie Arbeitsweisen verbessert werden können. Beschreitet ein Unternehmen den Weg mit Kanban, ist das mit ständiger Reflexion verbunden: Was hat die letzte Verbesserung bewirkt? War es überhaupt eine Verbesserung? Wenn ja, machen wir mehr davon – wenn nein, versuchen wir etwas anderes. Best Practices, wie man sie in vielen anderen Methoden rezepthaft vorfindet, gibt es aus meiner Sicht in Kanban daher – wenn überhaupt – nur im individuellen Kontext eines Unternehmens. Den Gedanken der evolutionären Verbesserung konkretisieren wir in Kanban anhand von drei Prinzipien. Diese drei Prinzipien sind der Nährboden, auf dem die sechs Praktiken, die den Fluss in einem System etablieren, nachhaltig gedeihen können.



Die Prinzipien von Kanban

1. Starte mit dem, was du jetzt machst.
2. Verfolge evolutionäre Verbesserung.
3. Fördere Leadership auf allen Ebenen der Organisation.

In Gesprächen merke ich häufig, dass sich der Begriff „evolutionär“ für viele nach einer Taktik der kleinen Schritte anhört. Das ist damit aber nicht zwangsläufig gemeint. „Evolutionär“ trifft keine Aussage darüber, wie klein oder groß Verbesserungsschritte sein sollten. Evolutionär bedeutet nicht mehr, als den Ist-Zustand zu analysieren, zu verstehen, was im System gerade passiert, und dementsprechend den Grad der Verbesserung je nach Bedarf zu justieren. Bei einer evolutionären Verbesserung kann es sich also durchaus um eine massive Veränderung handeln: Möglicherweise wird schon nach dem ersten Entwurf eines Kanban-Systems deutlich, dass es einen völlig unnötigen Prozessschritt gibt und dieser gleich gekappt werden kann. Aus der Sicht von Kanban ist es eine evolutionäre Verbesserung, weil sie die Erkenntnis aus der Betrachtung des Systems in seinem aktuellen Zustand ist und diese Erkenntnis von mehreren Personen hierarchie- und funktionsübergreifend geteilt wird. Trotzdem ist diese evolutionäre Verbesserung sehr tiefgreifend. Auch die Einführung von Visualisierung und WIP-Limits ist meistens kein kleiner, trivialer Schritt, sondern eine bedeutsame Veränderung, die mit den Stakeholdern eines Systems geklärt werden muss. Stakeholder eines Systems sind alle, die an einem Prozess unmittelbar beteiligt oder von dessen Auswirkungen spürbar betroffen sind.

Solche evolutionären und dennoch massiven Verbesserungen sind aber von dem zu unterscheiden, was nichts anderes als ein Rückfall in traditionelle Verhaltensweisen unter neuen Vorzeichen ist: Kanban wird als *die* neue Methode zur Lösung aller Probleme erkannt und genauso am Reißbrett für eine Big-Bang-Implementierung entworfen, wie man es bei ande-

ren zum Scheitern verurteilten Change-Initiativen macht. Das Management arbeitet fixfertige Kanban-Systeme aus, hängt sie ihren Mitarbeiterinnen und Mitarbeiter vor die Nase, bezeichnet alles bisher Dagewesene als Blödsinn und befiehlt Verbesserung. Das kann durchaus einer gut gemeinten Intention entspringen, aber: Nein, so ist evolutionäre Verbesserung nicht gemeint! Auch gut gemeinte Neuerungen riechen für die Mitarbeiter vor allem in sehr starren und hierarchisch geprägten Strukturen und Kulturen nach Überwachung und Kontrolle. Zu spüren bekommen es die Initiatoren als passiven Widerstand. Die Einführung von Kanban selbst braucht die Übereinkunft mit all jenen, die davon betroffen sein werden, und ist Gegenstand einer sorgfältig vorbereiteten Veränderungsinitiative (die Details dazu finden Sie in meinem ersten Buch, das ich gemeinsam mit Siegfried Kaltenecker geschrieben habe: „Kanban in der IT. Eine Kultur der kontinuierlichen Verbesserung schaffen“). Wesentliches Element einer gelingenden Kanban-Initiative ist das *gemeinsame* Design des Kanban-Systems durch alle Beteiligten bzw. deren Delegierte. Und auch diese Systeme sind nur der aktuelle Stand des Irrtums – Kanban-Systeme sind keine Skizzen für die Ewigkeit.

Hat man sich dazu entschieden, evolutionäre Verbesserung zu verfolgen, ergibt sich Prinzip 3 von Kanban fast zwangsläufig: Fördere Leadership auf allen Ebenen der Organisation. Ich habe zwei Perspektiven auf dieses Prinzip:

1. Wie bereits angemerkt, ist Kanban dann erfolgreich, wenn es nicht diktiert, sondern mit den davon Betroffenen erarbeitet wird. Die Menschen, die mit Kanban arbeiten sollen, müssen also zunächst abgeholt werden. In „Kanban in der IT“ ging es uns um genau diesen Punkt: Wie holt man die Beteiligten ins Boot, um eine möglichst breite und stabile Unterstützung als Basis zu schaffen, auf der eine Kanban-Initiative aufbauen kann?
2. Evolutionäre Verbesserung funktioniert nur mit den Menschen, die jeden Tag direkt im Geschehen sind. Sie wissen am besten, was funktioniert und was nicht funktioniert. Nicht in allen Unternehmen ist es aber üblich, dass Mitarbeiter frei sagen können, was ihnen auffällt. Mit der Einführung von Kanban müssen die Beteiligten ermutigt werden, Verbesserungsvorschläge zu machen, und das funktioniert am besten mit ehrlichem Interesse an ihrer Meinung. Sonst landet eine Kanban-Initiative schnell wieder im alten Paradigma, dass eine kleine Gruppe einer großen Gruppe sagt, was zu tun ist.

Kanban ist also keinesfalls ein Top-down-Ansatz, bei dem das Management am besten weiß, in welche Richtung es mit den Verbesserungen gehen soll und die dazugehörigen Maßnahmen ins Unternehmen drückt. Kanban ist aber auch kein Bottom-up-Ansatz. Zum einen wäre das nichts anderes als Top-down in umgekehrter Richtung und zum anderen scheitert dieser Ansatz oft daran, dass übergeordnete Hierarchiestufen es zwar dulden, aber vom Nutzen nicht überzeugt sind. Es bleibt daher bei einer lokalen Optimierung. Damit Kanban sein Potenzial entfalten kann, sollten alle Hierarchieebenen einbezogen werden, denn wir erinnern uns: Worum es bei der Arbeit mit Kanban eigentlich geht, ist der Kunde. Das Ziel ist ein möglichst durchgängiges, flussbasiertes und wertgenerierendes System, das zeigt, was auf dem Weg bis zum fertigen Projekt, Produkt oder Service im Sinne des Kunden verbessert werden kann. Bei der Etablierung dieses flussbasierten Systems, das den Weg der Wertgenerierung bis zum Kunden abbildet, helfen uns in Kanban sechs Praktiken.



Die sechs Praktiken von Kanban

1. Visualisiere
2. Limitiere den Work in Progress
3. Manage Flow
4. Mache Prozessregeln explizit
5. Implementiere Feedback-Loops
6. Führe Verbesserungen basierend auf Methoden und Modellen durch

Auch wenn es manchmal beinahe gebetsmühlenartig klingt, zeigt mir die Praxis, dass ich es immer wieder betonen muss: Kanban ist nicht mehr als die Umsetzung dieser drei Prinzipien und sechs Praktiken. Es wird nicht vorgeschrieben, *wie* sie umzusetzen sind – ja, es wird nicht einmal vorgeschrieben, dass sie überhaupt umgesetzt werden müssen. Das Schöne, für viele vielleicht aber auch Verfluchte an Kanban ist, dass es wie Burger King funktioniert: You can have it your way! Es legt ein System unter das Vergrößerungsglas und macht deutlich, was im Sinne der Unternehmensfitness zielführend und nicht zielführend ist, viel mehr nicht. Kanban per se macht die Situation in einem Unternehmen weder besser noch schlechter, sondern löst höchstens Reaktionen auf etwas plötzlich Sichtbares aus. Wer sich von Kanban unumstößliche und gebrauchsfertige Wahrheiten erhofft, ist damit schlecht beraten. Selbst denken ist in Kanban ausdrücklich erlaubt! Sehen wir uns also die Praktiken genauer an, die in der Wissensarbeit ein System in Fluss bringen.

Visualisiere

Wenn sich die unfertigen Schiffe in der letzten Aktivität des Flussexperiments stauen und sich der bemitleidenswerte Faltende beinahe die Hände bricht, wird sehr deutlich, wo im System der Engpass liegt. Jede Fertigung physischer Produkte offenbart die Stellen, an denen es Probleme gibt durch überfüllte Lager, Qualitätsprobleme, Blinken, Rauchen, Explodieren – jedenfalls in den meisten Fällen sehr deutlich wahrnehmbar. In der Wissensarbeit versteckt sich das Objekt der Verbesserung in den Köpfen der Mitarbeiter. Meistens zeigt sich das gleiche Bild: Menschen sitzen vor Monitoren und tippen. Unsichtbar bleibt, woran die Menschen arbeiten, warum sie daran arbeiten, wie lange sie bereits daran arbeiten, wie es ihnen dabei geht und ob alles in Ordnung ist. Wir können nicht *sehen*, wo wir in unserer Arbeitsweise mit Verbesserungen ansetzen sollten und somit können wir auch nicht sehen, ob unser Arbeitssystem für den Kunden noch den Zweck erfüllt, den er sich erwartet. Durch die Sichtbarkeit von Arbeit werden Zusammenhänge deutlich und es lassen sich bessere Entscheidungen treffen.

Limitiere den Work in Progress

Das Flussexperiment zeigt sehr plastisch, wie sich ein Arbeitssystem richtiggehend entspannt, wenn die Anzahl der begonnenen Arbeiten beschränkt wird. Im Experiment setzen wir die einfachste Methode ein, um zu einem WIP-Limit zu gelangen: Pro Aktivität wird immer maximal eine Arbeitseinheit zugelassen. Das ist aber nicht der einzige Weg und in manchen Fällen ist ein WIP-Limit von 1 auch nicht sinnvoll.

Was ist die Konsequenz, wenn wir in einem WIP-limitierten System arbeiten? Im Falle des kaiserlichen Gartens in Tokio müssen die Besucher vor den Eingangstoren warten, wenn alle Eintritts-Kanban verteilt sind – das System hat sein WIP-Limit erreicht. Und so ist es auch in einem WIP-limitierten Arbeitssystem: Irgendwann ist der Punkt erreicht, an dem keine neue Arbeit mehr angefangen werden kann, bevor nicht eine andere Arbeit abgeschlossen ist, die sich bereits im System befindet. Deswegen lautet der Stehsatz der Kanban-Welt auch: Stop starting, start finishing! Die simple ökonomische Überlegung dazu lautet: Eine zu 100 Prozent abgeschlossene Arbeit, die beim Kunden Wert generiert, bringt mehr Geld ein als zehn zu je zehn Prozent abgeschlossene Arbeiten.

Manage Flow

WIP-Limits sind der Schlüssel, um in einem System einen Arbeitsfluss herstellen und steuern zu können. Das Managen des Arbeitsflusses setzt aber bereits früher an: Es beginnt mit dem Sichtbarmachen der Aktivitäten, die ein Produkt (bzw. dessen Teile – Features, User Storys etc.) oder eine Dienstleistung auf dem Weg der Wertgenerierung durchläuft. Ziel eines Arbeitsflusses ist es, dass Arbeitseinheiten möglichst zügig und ohne Reibungsverluste diese Sequenz von Aktivitäten durchlaufen. Daher gehört nicht nur das Setzen von WIP-Limits zu den Instrumenten des Flow-Managements, sondern auch das Erkennen und Interpretieren der Zeichen, die uns ein Arbeitssystem durch die Visualisierung gibt.

Dass der Arbeitsfluss optimiert werden soll, ist keineswegs selbstverständlich. Meistens ist die Ressourcenauslastung Gegenstand der Optimierung, im Grunde also die Menschen, die in diesem System arbeiten. Arbeiten wir unter der Prämisse, beim Kunden Wert generieren zu wollen, ist hohe Auslastung aber in den seltensten Fällen eine sinnvolle Strategie. Wenn alle Mitarbeiter *immer* arbeiten, werden Dinge langsamer fertig. Fahren Sie gerne auf zu 100 Prozent ausgelasteten Autobahnen? Oder läuft Ihr Computer am besten, wenn er zu 100 Prozent ausgelastet ist? Dem Kunden ist es wichtiger, seine Bestellung möglichst schnell geliefert zu bekommen – zu welchem Prozentsatz die Mitarbeiter eines Unternehmens dabei ausgelastet sind, ist ihm prinzipiell egal. Mit ständiger Auslastung am Anschlag nimmt sich ein Unternehmen die Chance, die tatsächlichen Schwächen seines Arbeitssystems zu beheben und nachhaltige Verbesserungen umzusetzen. Genau diese Verbesserungen schlagen sich aber in der Wertgenerierung für den Kunden nieder.

Je nachdem, was Sie optimieren, senden Sie immer auch ein Signal an die Menschen, mit denen Sie arbeiten. Menschen den Raum zu geben, sich auf eine Arbeit fokussieren zu können, setzt ein Signal der Wertschätzung und des Vertrauens. Auslastungsbefehle und Mikromanagement hingegen transportieren immer die unterschwellige Botschaft: Die Menschen arbeiten nicht genug für das Geld, das sie bekommen. Natürlich wird auch in Kanban die Auslastung berücksichtigt, aber die Wertgenerierung steht an erster Stelle. Der Fokus liegt auf der Optimierung des Arbeitsflusses und nicht auf der Optimierung der Menschen, daher lautet die Devise von Kanban: Manage die Arbeit, nicht die Arbeitenden!

Mache Prozessregeln explizit

Kanban an und für sich ist kein Prozess, sondern wird auf eine bestehende Arbeitsweise angewendet. Genau diese Arbeitsweise ist es, die sichtbar gemacht und aufgeschrieben werden soll, damit alle Beteiligten vom selben Wissensstand ausgehen und gemeinsam das Arbeitssystem von diesem Standpunkt aus weiterentwickeln. Achtung: Bestehende Arbeits-

weisen unterliegen bestimmten Regeln. Meistens gibt es offizielle Vorschriften, der tatsächlich gelebte Prozess kann aber anders aussehen. Gegenstand von Kanban ist immer der tatsächlich gelebte Prozess, nicht das Wunschszenario.

Sind die Regeln des aktuellen Arbeitsprozesses aufgeschrieben und sichtbar gemacht, kommt das sonst so vorschrittsfreie Kanban nun doch mit zwei Vorschriften daher:

1. Wir halten uns penibel an diese Regeln.
2. Ist eine Regel nicht mehr sinnvoll, ändern wir sie.

Nur wenn man sich an Regeln und bestehende Arbeitsweisen hält, kann man feststellen, wo es Probleme in der aktuellen Arbeitsweise gibt. Umgekehrt kann man ein Arbeitssystem nicht mehr verbessern, wenn man die Regeln und Prozesse nicht mehr ändert.

Das Explizitmachen von Regeln wirkt sich übrigens auch auf die Zusammenarbeit aus. „Du hast aber damals gesagt ...“, „Er hat aber das getan ...“ – mit einem Regelwerk vor Augen hören Menschen meistens damit auf, sich gegenseitig die Schuld in die Schuhe zu schieben. Stattdessen rückt das Regelwerk in den Mittelpunkt der Diskussionen und die Beteiligten fokussieren sich darauf, eine Lösung für das identifizierte Problem in der Arbeitsweise zu finden.

Implementiere Feedbackloops

Eine Verbesserung setzt immer eine Veränderung voraus, aber nicht jede Veränderung ist eine Verbesserung. Um herauszufinden, wie sich eine Veränderung auswirkt, ist schnelles Feedback vom Arbeitssystem notwendig. Im Betrieb eines Kanban-Systems liefern die ersten drei Praktiken bereits sehr viele Informationen. Aber es gibt noch mehr Möglichkeiten. Die einfachste und nach wie vor wirksamste ist das Gespräch, daher schlage ich drei bekannte Formate für den regelmäßigen Austausch zwischen den Beteiligten vor: regelmäßige Standups, das Nachschubmeeting und die Retrospektive.

Messungen sind eine weitere, sehr objektive Form des Feedbacks. Soll das Arbeitssystem zum Beispiel schneller werden, bietet sich das Messen der Durchlaufzeiten an. Ändert man nun etwas unter der Annahme, dass es das System beschleunigt, liefert die Durchlaufzeit den Beweis oder falsifiziert die Annahme. Der springende Punkt daran ist: Messungen liefern dieses Feedback schnell. Wir sprechen in Kanban also nicht von „Lessons learned“, die im klassischen Projektmanagement am Ende eines Projekts gesammelt werden, sondern von unmittelbarer Information, die *sofort* Veränderungen oder Anpassungen nach sich zieht. Nur mit Feedback in regelmäßigen, kurzen Abständen kann man im Vorhinein etwas besser machen, statt es im Nachhinein reparieren zu müssen.

Verbesserungen basierend auf Methoden und Modellen durchführen

Kanban ist kein Dogma. Es gibt keine ultimative Kanban-Bibel, deren Gebote zu beachten sind, um nicht in der Kanban-Hölle braten zu müssen. Kanban by the book wäre ein Widerspruch in sich (wobei this book natürlich eine Ausnahme ist ...) und daher darf sich jeder, der Kanban einsetzt, aus dem reichen Fundus von Methoden und Modellen bedienen, die es mittlerweile zur Optimierung von Prozessen gibt. Nachdenken ist ausdrücklich erlaubt! Wobei die ganze Wahrheit aber lautet: Kanban zwingt uns zum Nachdenken. Bunte Post-its an der Wand haben lediglich dekorativen Wert, wenn man auf ein Wunder hofft.

■ 1.3 Wertgenerierung: In Services denken

Gleich zu Beginn habe ich mich darüber beschwert, dass Kanban häufig als Methode zur Optimierung von Teams betrachtet wird. Das zielt zum einen darauf ab, Menschen optimieren zu wollen, zum anderen wird damit die Struktur eines Unternehmens (eben in Form von Teams) in den Fokus der Verbesserung gerückt. Es werden also einzelne Bestandteile optimiert und wie das Flussexperiment auf einfache Art und Weise zeigt, ist das nicht förderlich für das gesamte System. Wird Kanban als strukturelle Verbesserungsinitiative verstanden, kreist eine Organisation letztendlich nur um sich selbst. Das ist aber nicht das Ziel – die Kundenzufriedenheit ist das Ziel. Nicht die Organisationsstruktur ist Gegenstand der Optimierung, sondern die Wertgenerierung.

Was wollen wir mit Kanban erreichen? Wir wollen erreichen, dass der Kunde zufrieden ist, weil wir für ihn einen Wert generiert haben. Wenn wir ein Unternehmenssystem durch die Kanban-Brille betrachten, versuchen wir daher zu identifizieren, in welchen Schritten wir vom Wunsch eines Kunden zum zufriedenen Kunden gelangen. Wir suchen also die Wertschöpfungskette in unserem System, die die Verbindung zwischen Kundenwunsch und Kundenzufriedenheit knüpft. Der Kunde sieht diese Wertschöpfungskette von außen als *eine* gesamte Leistung, als *einen* Service. Erfolgreich ist ein Unternehmen dann, wenn dieser Service gut funktioniert. Zoomen wir näher in diese gesamte Leistung hinein, sind die Glieder der Wertschöpfungskette nichts anderes als Teilservices, die dem großen Ganzen zum Erfolg verhelfen. Allerdings entsprechen diese Services nicht unbedingt der organisatorischen Struktur (mit einem Service ist also nicht ein Team oder eine Abteilung gemeint), denn wenn ein Arbeitsfluss etabliert und sinnvolle *Serviceeinheiten* definiert werden sollen, können diese auch mehrere organisatorische Einheiten umfassen. Services bilden also unterschiedliche Spannweiten aus, je nachdem, für wen der Service erbracht wird. Zum Beispiel kann die Softwareentwicklung interne Services für den Verkauf und die Analyse erbringen. In Projekten, in denen an externe Kunden geliefert wird, ist es möglicherweise sinnvoll, die Einzelservices Analyse, Softwareentwicklung und Test zu einer Serviceeinheit zusammenzufassen, um die Wertgenerierung zu optimieren. Die größte Spannweite würde sich bis zur Implementierung beim externen Kunden ergeben.

Idealerweise steht am Beginn einer Kanban-Initiative daher die Überlegung, bei welchen Kunden *wie* Wert generiert wird, die *gelebte* Realität wird am Kanban-Board explizit gemacht. Auf diesem Board wird der Kundenwunsch als Work Item bei seinem Weg durch das System verfolgt, sodass immer klar ist, wo sich die Arbeit gerade befindet. Dabei steht nicht die Frage im Vordergrund, ob alle Kapazitäten voll ausgelastet sind. Im Vordergrund steht der effiziente Fluss der Servicelieferung durch das System. Die Struktur einer Organisation kommt nur insoweit ins Spiel, als dass in einem zweiten Schritt überlegt wird, wer in welchen Stufen der Wertgenerierung benötigt wird. Das können einzelne Personen, Teams oder andere Organisationseinheiten sein. Deren Interaktionen sind dann Gegenstand der Verbesserungen, die zu einem *Arbeitsfluss* führen sollen. Das heißt: Das Unternehmen organisiert sich um die Arbeit herum, statt die Arbeit zu organisieren.



Bestellen wir ein paar Bücher!

Stellen Sie sich vor, Sie bestellen beim Online-Buchhändler Ihres Vertrauens ein Buch. Sie wollen einfach nur ein Buch. Aus Sicht des Händlers stellt sich die Frage, wie dieser Wunsch am besten erfüllt wird:

1. Zunächst werden die bestellten Bücher organisiert. Sie werden entweder aus dem Lager geholt oder beim Verlag geordert.
2. Danach wandern die Bücher in die Verpackung.
3. Im nächsten Schritt wird eine Rechnung erstellt.
4. Anschließend ist das Buch bereit für die Auslieferung und kann dem Lieferservice übergeben werden.
5. Hat der Kunde die Rechnung bezahlt, ist der Prozess für den Verkäufer abgeschlossen.

Eine sehr vereinfachte Abbildung dieser Schritte als Workflow könnte so aussehen:

Bücher organisieren → Bücher verpacken → Rechnung schreiben → Paket liefern → auf die Zahlung warten → Vorgang abschließen

So könnten durchaus auch die Spalten (Aktivitäten) auf einem ersten Kanban-Board aussehen. Was haben wir gemacht? Wir haben die Organisationsstruktur außer Acht gelassen und uns auf den Fluss des Buchs durch das Arbeitssystem hin zum Kunden konzentriert – auf das, was beim Kunden einen Wert schafft. Denn Sie als Kunde wollen einfach nur Ihr Buch und wir haben sichtbar und explizit gemacht, wie dieser Wunsch erfüllt wird!

David J. Anderson schreibt in seinem Blogbeitrag „The Kanban Lens“ sehr treffend (Anderson 2013): „Looking at a current organization through a service-oriented lens and seeing services where currently people only see functions and specializations is liberating and empowering.“ Das Denken in Serviceeinheiten statt in Funktionen und Spezialisierungen hat etwas Befreiendes, es kann sich hin zum Kunden öffnen. Würde man den Weg der Wertgenerierung anhand der organisatorischen Strukturen visualisieren, würde man ziemlich bald an die Grenzen des Überschaubaren stoßen und die Visualisierung wäre mehr Ablenkung als Hilfe. Wenn man Services nicht an organisatorischen Einheiten oder sogar an einzelnen Personen festmacht, sondern an Aktivitäten, nimmt man Komplexität aus der Frage, wie ein Unternehmen Wert generiert und erleichtert die Darstellung. Abgesehen von diesem rein praktischen Effekt entspannt die serviceorientierte, strukturfreie Sichtweise die Zusammenarbeit und führt Diskussionen weg von gegenseitigen Schuldzuweisungen auf eine sachlichere Ebene.

Die serviceorientierte Sichtweise macht aber auch deutlich: Kanban ist keine Methode zur Optimierung von Teams, denn in der Servicesicht kommt der Terminus „Team“ gar nicht vor. Eine Serviceeinheit entspricht nicht zwangsläufig einem Team. Natürlich muss man bei einer Kanban-Initiative irgendwo anfangen und dieser Anfang kann auch auf der Teamebene liegen, wenn es sinnvoll oder nicht anders möglich ist. Wenn wir von der „Spannweite“ eines Service sprechen, wird aber deutlich, dass Kanban an mehreren Ebenen ansetzen kann – dazu habe ich das Erklärungsmodell der Kanban Flight Levels entwickelt.

■ 1.4 Ebenen der Gestaltung: Kanban Flight Levels

Wie Sie nun bereits wissen, werde ich als Berater bei meinen ersten telefonischen Kontakten mit potenziellen Klienten in den meisten Fällen mit dem Satz konfrontiert: „Wir würden gerne diese neue agile Methode Kanban in unserem Unternehmen einführen und damit die Performance unseres Softwareentwicklungsteams verbessern.“ In diesem Stadium sprechen die Anfragenden und ich von zwei völlig verschiedenen Dingen. Nach meinem Verständnis

- ist Kanban keine agile Methode – Unternehmen können ihre Agilität mit Hilfe von Kanban verbessern.
- ist Kanban keine Softwareentwicklungsmethode – auch wenn Kanban oft in der Softwareentwicklung eingesetzt wird.
- fokussiert Kanban nicht auf Teams – auch wenn Teams Kanban natürlich für ihre Zwecke nutzen können.

In den ersten persönlichen Gesprächen versuche ich die Perspektiven daher zu justieren. Ich bringe das Denken in Services ins Spiel, um zunächst klar zu machen: „Versuchen Sie das, was Ihr Unternehmen gerade macht, als ein Set von Services zu sehen, das verbessert werden kann.“ So wird dann deutlicher, dass man mit Hilfe von Kanban entdecken kann,

- wie Wert generiert wird,
- wie und warum Arbeit in das System gelangt und
- wie die Arbeit durch die Serviceeinheiten fließt.

Der Fokus von Kanban ist immer die gesamte Wertschöpfung. Daran knüpft sich meistens auch schon die Frage, an welchen Stellen im Unternehmen Kanban implementiert werden kann. Die Prinzipien und Praktiken von Kanban haben einen entscheidenden Vorteil: Sie werden skalierungsfrei geliefert, das heißt, es ist nebensächlich, auf welcher Ebene einer Organisation man sie einsetzt – sie entfalten immer die gleiche Wirkung. Da Kanban den Anspruch hat, die gesamte Wertschöpfung eines Unternehmens auf dem Weg zum zufriedenen Kunden optimieren zu wollen, lautet der Rat daher: **Versuchen Sie, mit Kanban so breit wie möglich zu starten!**

Die Kanban Flight Levels sind in diesen Vorgesprächen einerseits ein Kommunikationsinstrument, mit dem ich zeigen kann, auf welchen Ebenen Kanban eingesetzt werden kann und welche Wirkung es auf diesen Ebenen entfaltet. Gleichzeitig hilft es bei der Reflexion darüber, wo die derzeitigen Probleme (für die man sich mit Kanban eine Lösung erhofft) tatsächlich liegen und auf welchen Ebenen der Start mit Kanban sinnvoll, vor allem aber möglich wäre. Wenn Sie Teamleiter sind und derzeit noch keinen Weg sehen, auf einer höheren Ebene zu starten, weil Ihnen zum Beispiel die Verbündeten fehlen – vollkommen in Ordnung. Nur bedenken Sie bei Ihrer Initiative, dass Kanban die gesamte Wertschöpfungskette eines Unternehmens im Sinn hat und eben keine dezidierte Teammethode ist. Virale Verbreitung erwünscht!

Wichtig ist: Die Kanban Flight Levels sind keine Skalierungsschablone. Auch wenn es Flight Level 1, 2, 3 und 4 gibt, bedeutet das nicht, dass es sich um aufeinander aufbauende

Ebenen handelt! Die Flight Levels bilden keinen Entwicklungspfad ab, den eine Organisation zwangsläufig gehen muss, und sie sind auch kein Überprüfungsinstrument, um „gutes“ und „schlechtes“ Kanban zu unterscheiden. In den meisten Unternehmen existieren sogar mehrere Flight Levels gleichzeitig, entweder durch eine bewusste Vorgehensweise oder weil es sich im Laufe der Zeit so entwickelt hat.



Die **Kanban Flight Levels** sind ein Kommunikationsinstrument, um die Einsatzmöglichkeiten von Kanban deutlich zu machen und herauszufinden, wo für eine Organisation der sinnvolle und/oder mögliche Ausgangspunkt liegt, um eine evolutionäre Verbesserung zu starten.

1.4.1 Flight Level 1: Team mit unreguliertem Input und Task-Fokus

Die erste Flugebene gehört den Teams. Häufig sind die involvierten Personen hoch spezialisierte Experten, vor allem in Hightech-Umgebungen begegne ich solchen Teams immer wieder. Diese Spezialisten arbeiten ausschließlich an Teilbereichen eines riesigen Gesamtsystems, etwa an den Einspritzsystemen eines Autos oder an der grafischen Aufbereitung eines Wetterradars für Flugzeuge. Eine andere typische Erscheinungsform von Einheiten auf Flight Level 1 sind crossfunktionale Teams: zum Beispiel Designer, Entwickler und Tester, die gemeinsam an einem kleineren Produkt oder Subsystem arbeiten.

Charakteristisch für Flight Level 1 ist, dass der Input an dieses Team nicht koordiniert wird (Bild 1.4). Es gibt also keine Nachschubmeetings (ich werde sie später als „Spice Girls Meeting“ bezeichnen – mehr dazu im nächsten Abschnitt), in denen Stakeholder ihre Aufträge priorisieren, sondern die Arbeit wird meistens über den Zaun geworfen. Das hat auch einen guten Grund: Wegen der Spezialisierung gibt es unzählige Stakeholder, die das Wissen bzw. die Dienstleistungen dieser Teams oder besser gesagt „Spezialeinheiten“ brauchen. Mir sind Fälle begegnet, in denen ein neunköpfiges Team mehr als 100 Projekte bedienen musste. 100 Projektverantwortliche in einem Nachschubmeeting zu koordinieren und sie entscheiden zu lassen, welche Aufgaben das Team als Nächstes bearbeiten soll, ist reine Science-Fiction. Praktisch bedeutet das aber auch, dass solche Teams mit einer Menge gleich priorisierter Aufgaben konfrontiert sind, und es gewinnt, wer am lautesten schreit. Kanban leidet auf Flight Level 1 daher oft an Unmengen von Expressarbeiten (Arbeiten mit hoher Priorität), verursacht durch ständige Repriorisierung.

In diesen Spezialistenteams kommt auch nie ein Arbeitsfluss zustande, es werden einfach Aufgaben oder Arbeitspakete abgearbeitet. Das Kanban-Board ist im Wesentlichen ein Task-Board mit den Spalten To Do, Doing, Done und manchmal findet man noch eine Warten-Spalte. Flight Level 1 ist auch die Ebene der Einzelkämpfer: Eine Person erledigt einen Arbeitsauftrag vollständig alleine, ohne Hilfe von anderen Teammitgliedern und ohne Hand-offs.

Zweifellos verbessert Kanban auf Flight Level 1 die Effizienz des Teams. Der Start ist einfach und zumindest im Team selbst macht Kanban mit Techniken wie der Visualisierung und den WIP-Limits deutlich, was verbessert werden muss. Nach außen stößt das Team aber an seine Grenzen, denn der Schwachpunkt ist die fehlende Inputkoordination. Lautes

Schreien, um als Erster bedient zu werden, nützt vielleicht der subjektiven Befindlichkeit, es nützt aber selten einem Unternehmen in seiner Gesamtheit. Solche Teams arbeiten mitunter sehr effizient an den – in der Gesamtbetrachtung – falschen Dingen. Auf der Strecke bleiben dadurch die Effizienz und noch viel mehr die Effektivität der gesamten Wertschöpfungskette, die sich auf diese Weise nicht verbessern kann.

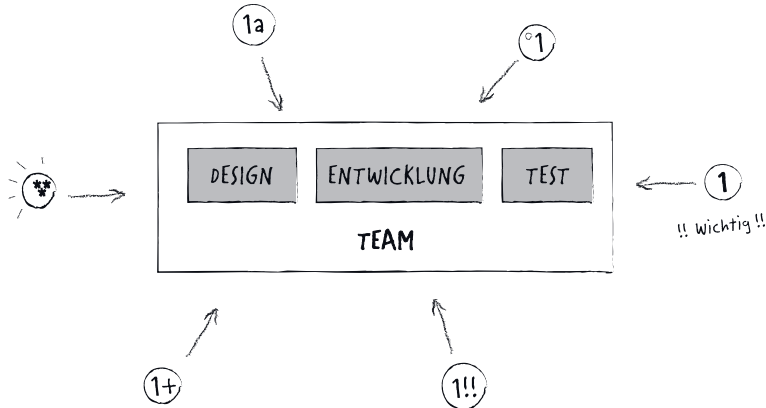


Bild 1.4 Flight Level 1: Team mit unreguliertem Input

1.4.2 Flight Level 2: Team mit koordiniertem Input, Etablierung eines Arbeitsflusses

Auf Flight Level 2 betrachten wir – so wie auf Flight Level 1 – ein Team, aber es gibt zwei entscheidende Unterschiede: Der Input wird koordiniert und es ist ein Arbeitsfluss etabliert.

Betrachten wir zunächst die Inputkoordination: Die Arbeiten werden zum Beispiel durch die Entscheidungen zuständiger Personen in einem Nachschubmeeting ausgewählt und in eine *limitierte* Input Queue gebracht – der Zufluss von Arbeit in das System wird also beschränkt. Dieses Nachschubmeeting wurde von fünf Managementvordenkerinnen der 1990er-Jahre inspiriert: von den Spice Girls. Lange bevor Kanban in der Wissensarbeit Einzug hielt, stellten sie in ihrem Grundlagenwerk „Wanna be“ bereits fest: Tell me what you want, what you really really want!¹ Auf Kanban übersetzt bedeutet diese Metapher: Man kann nicht unendlich viel Arbeit ins System kippen, um zu schauen, was passiert. Der Repriorisierungswahnsinn beginnt dann nämlich, wenn die Arbeit bereits *im* System ist. Man forciert daher eine bewusste Entscheidung darüber, was als Nächstes ins Arbeitssystem gelangen und abgeschlossen werden soll. Sinnvollerweise ist das keine Entscheidung nach dem Motto „Ich will, dass ...“ (what you want), sondern eine Entscheidung nach ökonomischen Kriterien.

¹ <https://www.youtube.com/watch?v=gJLiF15wjQ>

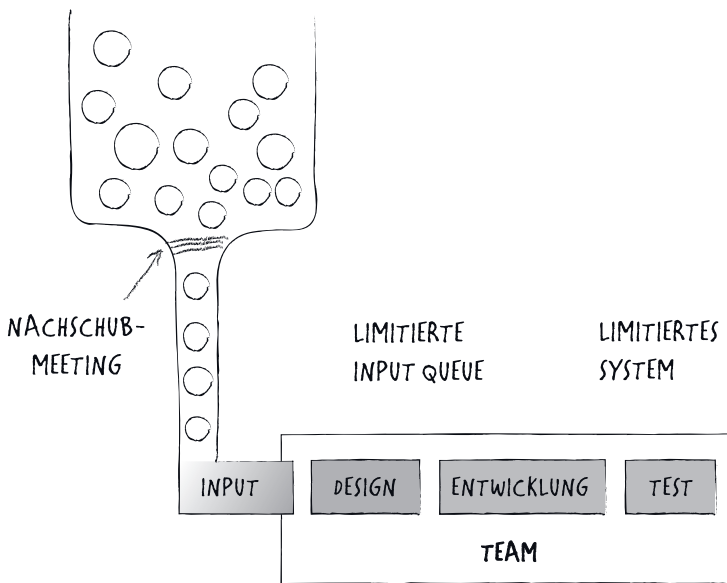


Bild 1.5 Flight Level 2: Koordinierter Input durch die Spice Girls

Zu den Spice Girls gehören Vertreter von Gruppen, die am Output des Systems interessiert sind, sowie Vertreter des Teams selbst. Der große Vorteil dieser Spice Girls Meetings ist der Lerneffekt über die Grenzen einer Organisationseinheit hinaus. Die Spice Girls verstehen auf diese Weise besser, dass sie es nicht mit einer Black Box zu tun haben, in die man nach Belieben Arbeit hineinkippen kann und am anderen Ende kommt das Gewünschte heraus. Es gibt einen Bedarf auf der einen Seite und zeitliche sowie kapazitätsmäßige Möglichkeiten auf der anderen Seite. Koordinierter Input bedeutet, aus einem Pool von Ideen und Wünschen bewusst auszuwählen und im Dialog in eine sinnvolle Reihenfolge zu bringen. Während Kanban-Initiativen auf Flight Level 1 oft Guerilla-Charakter haben, wird auf Flight Level 2 zum ersten Mal deutlich, dass bei der Einführung von Kanban ein begleitendes Change Management nötig ist, denn die Spice Girls müssen sich an diese Form der Zusammenarbeit in der Regel erst gewöhnen.

Im Gegensatz zu Flight Level 1 findet man auf Flight Level 2 keine Einzelkämpfer, die Aufgaben im Alleingang abarbeiten. Hier arbeiten Personen in einem Team zusammen, die es als ihre *gemeinsame* Mission verstehen, einen hoffentlich recht großen Teil der Gesamtwertschöpfungskette abzudecken. Das Team etabliert und managt einen Arbeitsfluss. Das bedeutet, dass die Menge der Arbeiten *im* Arbeitssystem eines Teams auf Flight Level 2 limitiert ist. Durch das koordinierte Vorgehen im Spice-Girls-Meeting wird auf Flight Level 2 seltener repriorisiert, weil das Team das Pull-Prinzip anwenden kann und durch die Koordination im Vorfeld mit höherer Wahrscheinlichkeit die *richtigen* Dinge (zumindest vom Standpunkt der Spice Girls aus betrachtet) in einer sinnvollen Reihenfolge effizienter bearbeitet werden. Sowohl auf Flight Level 1 als auch auf Flight Level 2 sind aber nach wie vor Teams das Ziel von Verbesserungen. Damit wird nur ein Ausschnitt des gesamten Wertstroms eines Unternehmens betrachtet – es wird also lokal optimiert. Unberücksichtigt bleibt die Tatsache, dass die einzelnen Teams einer Organisation in vielfältigen Verbindungen und Abhängigkeiten zueinander stehen.

Lokale Optimierung führt zu globaler Suboptimierung

Das Flussexperiment hat sehr deutlich gezeigt, was passiert, wenn nur ein Glied der Wertschöpfungskette optimiert wird: Die Verbesserung reicht nur bis an die Grenzen des Teams und kann in anderen Teams sogar mehr Schwierigkeiten verursachen. Der größte Nachteil daran ist, dass trotz guter Intentionen eines Teams die Wünsche der Kunden nicht unbedingt besser erfüllt werden. Ich erkläre das gerne am Beispiel einer Tastatur (Bild 1.6).

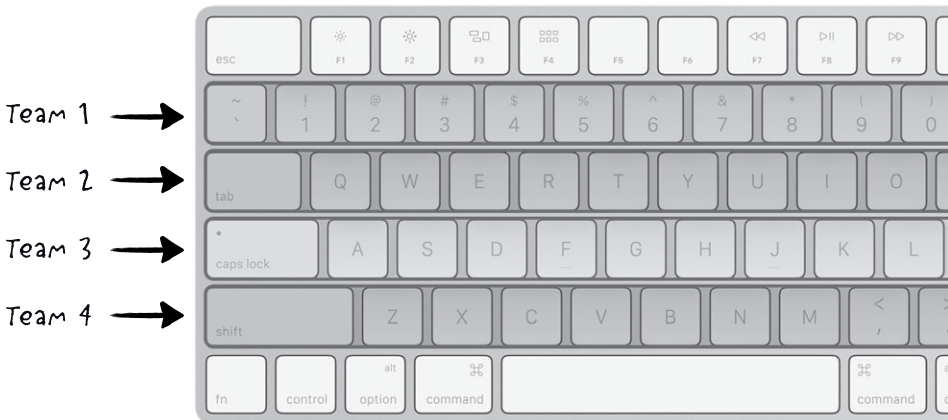


Bild 1.6 Wird ein Brief durch ein schnell angeschlagenes A besser?

Nehmen wir an, ein Kunde wünscht sich einen Brief. Nehmen wir dazu weiter an, dass jedes Team unseres Unternehmens für eine Reihe der Tastatur zuständig ist. Jedes Team ist Meister seiner Reihe, aber Raum für Verbesserungen gibt es immer. Nun kann sich zum Beispiel Team 3 so lange optimieren, bis es einen neuen Weltrekord im Anschlagen des A aufstellt. Wunderbar! Allerdings wird der Brief dadurch kein bisschen schneller fertig. Beim Schreiben eines Briefs geht es auch nicht darum, dass man einen bestimmten Buchstaben besonders schnell anschlagen kann. Wichtiger ist es, den richtigen Buchstaben zur richtigen Zeit anzuschlagen, um eine wirkliche Leistungssteigerung zu erreichen. Daher liegt der Fokus von Kanban auf der gesamten Organisation. Sie ist es, die sich in evolutionären Schritten auf ihre Kunden ausrichten soll: Was muss getan werden, um für den Kunden einen optimalen Wert zu generieren? Ob eine Organisation agil ist, hat nichts damit zu tun, ob sie aus vielen agilen Teams besteht. **Die Interaktionen zwischen den Teams müssen agil sein.**

1.4.3 Flight Level 3: Optimierung des Wertstroms

Übersetzen wir es auf eine alltäglichere Situation in der Geschäftswelt. Bild 1.7 zeigt einen Wertstrom, den ich im Projektgeschäft immer wieder in Variationen sehe. Design, Entwicklung und Test sind nur ein Teil des gesamten Wertstroms. In den meisten Fällen hat es für den Kunden keinen Wert, wenn er nur getesteten Code geliefert bekommt. Der Wert entsteht erst, wenn der Code in ein größeres (Live-)System integriert wird, das den Betrieb aufnimmt. Andererseits spezifizieren Kunden ihre Wünsche selten so kristallklar, dass sie sofort umgesetzt werden können. Ganz im Gegenteil: In der Regel ist vor der eigentlichen

Entwicklungsarbeit noch viel anderes zu tun. Es müssen Leads generiert und Projektteams aufgesetzt werden, das Team muss verstehen, was der Kunde will, Verträge müssen geschlossen werden usw. Sobald man den Blick auf eine andere Sparte wirft, kann der Wertstrom ganz anders aussehen, etwa in der Produktwartung oder im Produktmanagement. Nichtsdestotrotz steht dieselbe Idee dahinter: Einen Wert für den Kunden zu schaffen, ist wichtiger als ein hyperproduktives Team. Genau hier setzt Flight Level 3 an.

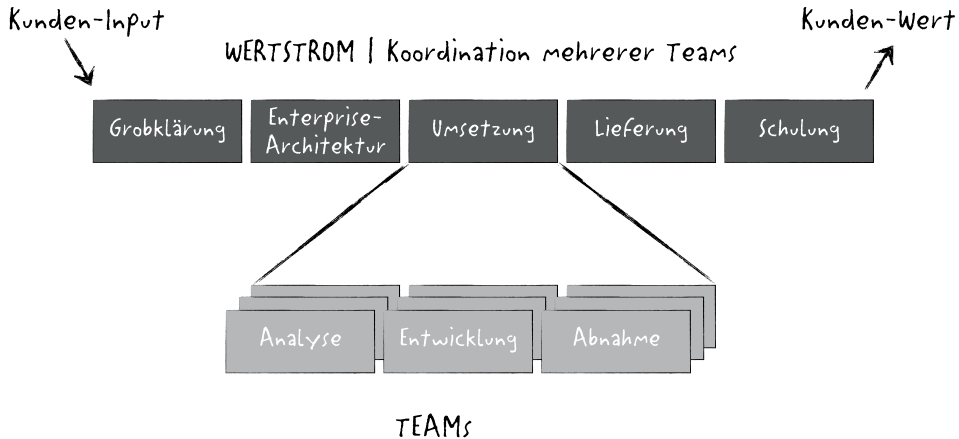


Bild 1.7 Flight Level 3: Betrachtung der Programmebene



Ist das ein Wasserfall?

Wenn man sich das Bild ansieht, kann man zur Auffassung kommen, dass Flight Level 3 die Manifestation einer wasserfallartigen Umsetzung ist. Was bedeutet Wasserfallentwicklung aber? Nur weil Dinge hintereinander abgearbeitet werden, bedeutet das nicht, dass man sich automatisch in einem Wasserfallentwicklungsprozess befindet (übrigens hat selbst Winston Royce das Wasserfallmodell als iterativen Prozess beschrieben – siehe Royce 1970). Es ist aber ein gutes Muster, zum Beispiel die Funktionalität einer Software zuerst zu programmieren, bevor man sie ausliefert.

Tatsächlich nach dem Wasserfallprinzip zu arbeiten, bedeutet, dass die Arbeitseinheiten, die durch den Arbeitsfluss gezogen werden, sehr groß sind, zum Beispiel ein komplettes Projekt. Es ist nicht zielführend, ein komplettes Projekt zuerst zu analysieren, dann zu entwickeln und schließlich zu implementieren. Davon spreche ich hier aber nicht. Die Arbeitseinheiten sollen möglichst klein sein, damit man so schnell wie möglich verifizieren kann, ob ein Ansatz richtig ist. Dazu gehört, so schnell wie möglich Feedback zu bekommen, um daraus lernen zu können.

Auf Flight Level 3 wird die Interaktion der Teams optimiert. Um in der Metapher der Tastatur zu bleiben, wird auf Flight Level 3 sichergestellt, dass das richtige Team zur richtigen Zeit die richtige Taste drückt, also an der richtigen Arbeit zur richtigen Zeit arbeitet. Das

Kanban-Board koordiniert somit die Arbeit mehrerer Teams, die gemeinsam daran beteiligt sind, Kundenwünsche zu erfüllen. Ziel ist es, den Arbeitsfluss über die Teamgrenzen hinweg zu optimieren. Kanban bringt auf Flight Level 3 massive Performance-Steigerungen, hauptsächlich aus zwei Gründen:

1. Die Mitarbeiter arbeiten an den richtigen Dingen zum richtigen Zeitpunkt, denn das Befüllen der Input Queues wird koordiniert.
2. Die Anzahl der Arbeiten wird durchgehend limitiert und somit kann das Arbeitssystem insgesamt optimiert werden.

Da der Arbeitsfluss über den gesamten Wertstrom optimiert wird, verringern sich die Wartezeiten an den Schnittstellen und besonders wichtig: Engpässe werden deutlich sichtbar. Wer Organisationen auf diesem Flight Level erlebt hat, kann nur mehr lächeln, wenn High-Performance-Teams als Erfolgsgeheimnis verkauft werden.

Betrachten wir Kanban Flight Level 3 aus der Sicht des Change Management, sind Initiativen auf dieser Flughöhe – auch wenn damit mehrere Hundert Menschen koordiniert werden können – manchmal einfacher als auf Flight Level 2. Die Entscheidung, teamübergreifend mit Kanban zu arbeiten, fällt hier nämlich im höheren Management. Das bedeutet auch, dass das Management mit gutem Beispiel vorangeht und die gewünschte Veränderung vorlebt. Gleichzeitig muss nicht zwingend in jedem einzelnen Team Kanban eingeführt werden, niemand muss seine Arbeitsweise sofort ändern. Ob ein Team nun Kanban oder Scrum macht oder einfach völlig normal vor sich hin arbeitet, ist völlig irrelevant. Die einzige Veränderung, wenn man es so nennen will, ist, dass die Teams ihre Arbeit geregelt durch die Spice Girls entgegennehmen und dass sich Delegierte aus den Teams in eigenen Meetings – meistens in regelmäßigen Standups – koordinieren.

1.4.4 Flight Level 4: Optimierung von Strategie und Portfolio

Üblicherweise arbeiten Organisationen nicht nur an einem einzigen Projekt oder einem einzigen Produkt. Das Portfolio eines Unternehmens setzt sich aus einer Vielzahl von Projekten und Produkten sowie strategischen Initiativen zusammen, die das Unternehmen fit für die Zukunft machen. Genau dieser Mix wird auf Flight Level 4 gemanagt. Man will den Überblick über das Geschehen im Unternehmen gewinnen, man will wissen, welche Projekte und strategischen Initiativen sich auf welche Art und Weise beeinflussen und wie weit die Umsetzung gediehen ist.

Grundsätzlich ist es ein gutes Problem, mehr Nachfrage als Möglichkeiten zur Umsetzung zu haben, denn sonst muss ein Unternehmen Mitarbeiter abbauen. Dadurch entsteht auf Portfolio-Ebene allerdings eine Konkurrenzsituation zwischen den Optionen. Dieses Missverhältnis zwischen Optionen und Umsetzungsmöglichkeiten muss explizit gemacht werden, da sonst der Eindruck entsteht, es stünden unendlich viele Ressourcen zur Verfügung. Dem ist aber nicht so und genau darum geht es bei Kanban auf Flight Level 4: Um die kluge Auswahl und Kombination von Projekten, zu entwickelnden Produkten und strategischen Initiativen, um Abhängigkeiten zu erkennen und den Fluss durch die Wertschöpfungskette angesichts der tatsächlich verfügbaren Ressourcen zu optimieren.

Bei den Aufgabentypen handelt es sich auf Flight Level 4 um große Arbeitseinheiten, zum Beispiel „Markteintritt in Ungarn“ oder „weniger Automotive-, mehr Aviation-Business“. Auf dieser Flughöhe konkurrieren diese großen Funktionalitäten miteinander, also ist die Organisation gezwungen, überlegte und bewusste Entscheidungen darüber zu treffen, was als Nächstes abgeschlossen werden soll. Im Mittelpunkt stehen dabei nicht mehr die Ziele der einzelnen Projekte, sondern das Gesamtergebnis der Organisation. Bedarf und Möglichkeiten müssen genau abgewogen werden.

Bild 1.8 fasst die Flight Levels auf einen Blick zusammen. Flight Level 4 ist das strategische Herz der Organisation. Hier laufen die Projekte und Initiativen des Unternehmens zusammen und hier findet auch das strategische Management statt. Flight Level 4 ist mit mehreren Systemen auf Flight Level 3 und Flight Level 2 verknüpft, wo die Arbeit detaillierter gemanagt wird. Das Beispiel „Markteintritt in Ungarn“ besteht auf Flight Level 4 vielleicht nur aus den zwei Tickets „Produktgruppe X bereitstellen“ und „Produktgruppe Y bereitstellen“, auf Flight Level 3 werden diese groben Arbeitseinheiten feiner zerlegt und für die Teams auf Flight Level 2 zur Verfügung gestellt.

Es ist auch häufig der Fall, dass Arbeit von Flight Level 4 direkt zu einem Team auf Flight Level 2 fließt. Nehmen wir an, bei unserem Unternehmen handelt es sich um einen Autoteilezulieferer und die Strategie auf Flight Level 4 lautet „mehr Aviation-Business“. Es gibt ein Produkt, das derzeit für den Automotive-Markt entwickelt wird, und es steht die Hypothese im Raum, dass dieses Produkt mit leichten Modifikationen auch für den Aviation-Markt interessant sein kann. Ein Spezialistenteam kümmert sich darum, diese Hypothese zu überprüfen, die Arbeitseinheit fließt also von Flight Level 4 direkt auf Flight Level 2 zum Spezialistenteam.

Wie bereits am Beginn dieses Abschnitts erwähnt, sind die Flight Levels kein Maturitätsmodell. Im Idealfall sind alle Flight Level von 2 bis 4 in einer Organisation vorhanden. Flight Level 4 ist also nicht „besser“ als Flight Level 2, es löst einfach ein anderes Problem. Die Unternehmensstrategie ist kein Fall für die Teamebene und umgekehrt können Teamprobleme nicht mit einem noch so ausgefeilten Portfolio-Management behoben werden. Natürlich können auf einem Board auch mehrere Flight Levels dargestellt werden. Das sieht man häufig in kleinen Unternehmen, wo zum Beispiel nur zwei Teams an einem Projekt oder Produkt arbeiten. Hier wird der gesamte Wertstrom auf einem Flight-Level-3-Board explizit gemacht und zusätzlich werden am selben Board auch die Aufgaben der Teams visualisiert.

Ganz allgemein verstehe ich die Flight Levels als Kommunikationstool, mit dem man herausfinden kann, wo Kanban in einem Unternehmen eingesetzt werden kann und wo es im konkreten Fall die größte Hebelwirkung entfalten kann. Flight Level 1 ist in Bild 1.8 verblasst dargestellt. Damit will ich zum Ausdruck bringen, dass man versuchen sollte, immer höher als auf Flight Level 1 zu starten, da die Hebelwirkung in der Organisation sonst zu gering ist.

Kanban Flight Levels Wo wird Kanban im Unternehmen eingesetzt?

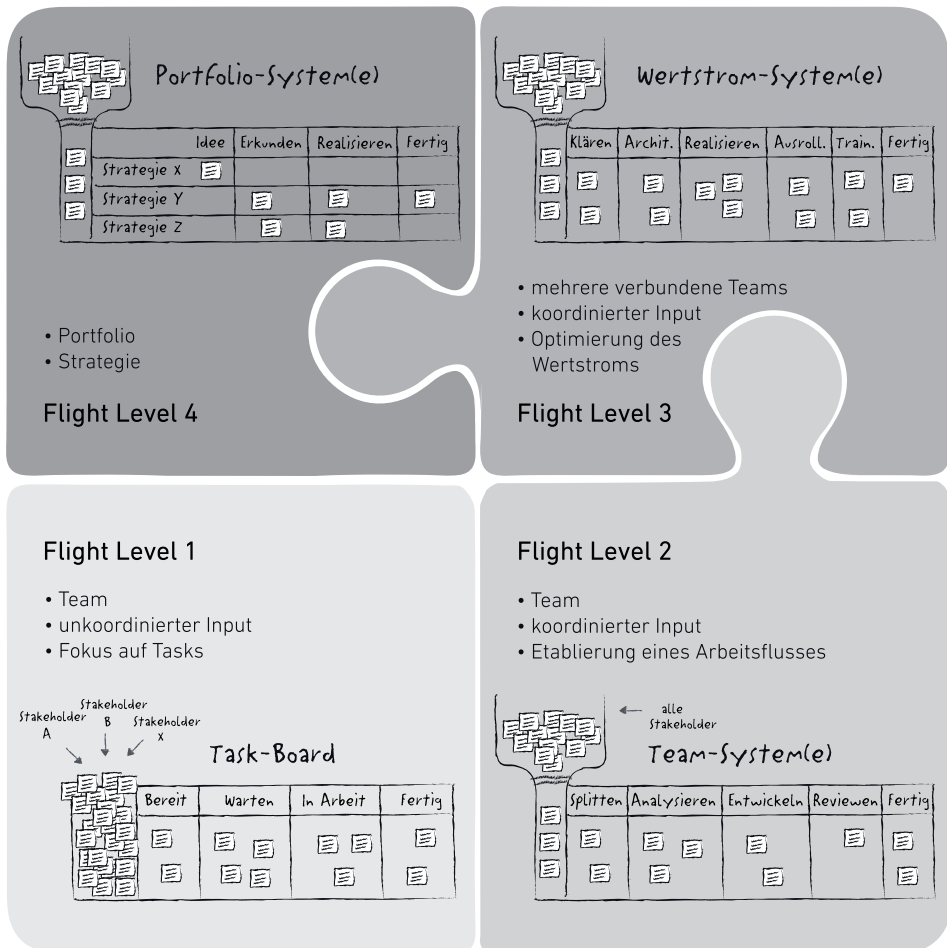


Bild 1.8 Die Flight Levels im Überblick

Das Praktische an Kanban ist: Obwohl es auf mehreren Ebenen eingesetzt werden kann, reichen immer die sechs vorgeschlagenen Praktiken (siehe Abschnitt 1.2) aus, um den Prozess der evolutionären Verbesserung in Gang zu setzen und ein flussbasiertes Arbeitssystem entstehen zu lassen. Sehen wir uns nun an, wie Kanban-Systeme auf Basis dieser Praktiken im Betrieb funktionieren.



Zusammenfassung

Soll Kanban in einem Unternehmen eingesetzt werden, lautet die wichtigste Frage zunächst: Warum wollen wir das überhaupt und was erhoffen wir uns davon? Idealerweise geht es darum, einen Wert zu schaffen, der den Kunden eines Unternehmens wichtig ist und darum lautet die eigentliche Frage: Was ist dem Kunden wichtig? Dementsprechend sollte der Arbeitsfluss gestaltet werden. Geht man von dieser Frage aus, wird auch deutlich, dass ein erfolgreiches Kanban-System nicht darauf aufbaut, einfach nur die sechs Praktiken umzusetzen. Erfolgreich ist man mit Kanban, wenn man die Erkenntnisse, Prinzipien und Werte hinter diesen Praktiken verstanden hat.

Bei genauerer Betrachtung entfalten sich in Kanban durch die sechs Praktiken von Kanban Wirkungen, die nicht immer intuitiv sind:

- Wenn wir schneller arbeiten, wird trotzdem nicht mehr Arbeit fertig.
- Wir haben genug Zeit für Aufgaben, für die wir nie Zeit haben.
- Wir werden verlässlicher, wenn wir uns Grenzen setzen.
- Wenn alles Priorität hat, hat nichts Priorität.
- Je später wir anfangen, desto besser für den Kunden.
- Lokale Optimierung führt zu globaler Suboptimierung.

Wenn man von der Frage ausgeht, was für den Kunden Wert generiert, kann man Unternehmen nicht mehr als Konglomerat einzelner Teams und Abteilungen betrachten. Sinnvoll ist das Denken in Services, also in Teilschritten, die von der Auftragserteilung bis zur Lieferung an den Kunden Teile des Werts erzeugen. Services können unterschiedliche Spannweiten haben, daher ist es sinnvoll, sich zu Beginn einer Kanban-Initiative anzusehen, auf welcher Ebene – oder auf welchem „Flight Level“ – man ansetzen kann. Prinzipiell gibt es vier Flight Levels:

- Flight Level 1: Team mit unreguliertem Input und Task-Fokus
- Flight Level 2: Team mit koordiniertem Input, Etablierung eines Arbeitsflusses
- Flight Level 3: Optimierung des Wertstroms
- Flight Level 4: Optimierung von Strategie und Portfolio

Die größte Hebelwirkung entfaltet Kanban auf Flight Level 3, weil die Initiative auf dieser Ebene meistens vom Top-Management legitimiert ist und auf dieser Flughöhe die Interaktionen und Abhängigkeiten zwischen Teams und Abteilungen gemanagt werden können, um einen kundenorientierten Wertstrom entstehen zu lassen. Versuchen Sie also, mit Kanban so breit wie möglich zu starten.



Literatur

(Interviews Ackoff, Deming)

Interviews with Dr. Akoff (sic!) and Dr. Deming, Tapes #1A through #4A.
<http://ackoffcenter.blogs.com/files/dr-ackoff-dr-deming.pdf>

(Ackoff, Gharajedaghi 1984)

Russel Ackoff; Jamshid Gharajedaghi: Mechanisms, Organisms and Social System. In Strategic Management Journal, Vol. 5, 289 – 300 (1984).

(Anderson 2010)

David J. Anderson: Kanban. Successful evolutionary Change for your technology business. Blue Hole Press 2010.

(Anderson 2013)

David J. Anderson: The Kanban Lens. Blogbeitrag vom 18. Oktober 2013.
<http://www.djaa.com/kanban-lens>

(Benson, DeMaria Barry 2013)

Jim Benson; Tonianne DeMaria Barry: Personal Kanban. Visualisierung und Planung von Aufgaben, Projekten und Terminen mit dem Kanban-Board. dpunkt Verlag 2013.

(Deming 2000)

W. Edwards Deming: Out of the Crisis. The MIT Press 2000.

(Leopold, Kaltenecker 2013)

Klaus Leopold, Siegfried Kaltenecker: Kanban in der IT. Eine Kultur der kontinuierlichen Verbesserung schaffen. 2., überarb. Aufl., Hanser 2013.

Blogbeiträge von Klaus Leopold:

- Kanban and its Flight Levels. <http://bit.ly/1KwruTN>
- Why Kanban Flight Levels? <http://bit.ly/1zfKMvF>
- Optimize value creation and not teams. <http://bit.ly/1ysp471>

(Ohno, Bodek 1988)

Taiichi Ohno; Norman Bodek: Toyota Production System. Beyond Large-Scale Production. Productivity Press 1988.

(Rother 2010)

Mike Rother: Toyota Kata. Managing people for improvement, adaptiveness, and superior results. McGraw-Hill 2010.

(Royce 1970)

Winston W. Royce: Managing the development of large software systems. In: Proceedings, IEEE WESCON, August 1970, p. 1 – 9.

Index

A

Aktivität 40
Arbeit
– nicht geplante 54
Arbeitseinheit 40
Arbeitsfluss 22, 24, 40
Arbeitstyp 40, 49
Arbeitstypen 50
– hierarchische 52
– Veränderung 51

B

Batchsize 178
Blockade 39, 58, 114
– beeinflussbare 62
– durch Engpass 62
– externe 62
Blockadencluster 60
Blockadensticker 60
Blocker Clustering 58
Bug Fixing 67
Burn-up-Chart 148

C

Capacity Constrained Resource
 siehe Kapazitätsengpass
Change Management 29, 32
Cost of Delay *siehe* Verzögerungskosten

D

Defekte 67
Definition of Done 55
Definition of Ready 57
Deming, William Edwards 45
Dienstgütevereinbarung *siehe* Service Level Agreement
Downstream 196
Dringlichkeitsprofil 182
Durchlaufzeit 45, 106, 131, 171
Durchsatz 10, 45, 132, 171
Durchsatz-Forecasting 147

E

Engpass 10, 32
Epic 52

F

Feedbackloops 23
Fitnesskriterien 5
Flow debt 167
Fluss 47
Flusseffizienz 40, 45, 84
Forecast
– für eine Arbeitseinheit 134
– für mehrere Arbeiten 139
Forecasting 127, 219

G

Generalisierung 176
Generalisten 82

H

High Speed Assessment 96
HiPPO-Verfahren 175, 208
Histogramm 138

I

Input Queue 38, 57, 89

K

Kanban
– im Großen 105
– Praktiken 21
– Prinzipien 19
Kanban Flight Level 1 27
Kanban Flight Level 2 28
Kanban Flight Level 3 30
Kanban Flight Level 4 32
Kanban Flight Levels 26
Kapazitäten
– Erhöhung der 170
Kapazitätsengpass 79
Knowledge Agility 84
Koordination 87
Kumulatives Flussdiagramm 160
Kundenvvalidierung 75

L

Late Commitment 14, 216
Limitierung 170
Little, John D.C. 45
Little's Law 45, 157

M

Market Role Model 205
Marktrisiko 199
Messdaten
– Gewinnung von 132

Modell 128
Moderation 90
Monte-Carlo-Simulation 143
Multitasking 44

N

Nachfrage
– Management der 174
Nachschubmeeting 88
Non Instant Availability
siehe Spezialistenengpass

O

Optimierung
– lokale 15, 30

P

Personal Kanban 52
Portfolio-Ebene 32
Priorisierung 169
Prioritäten 13
Produktivfehler 65
Produktlebenszyklus 206
Prozessregeln 22

Q

Qualität 11
Quantil 136

R

Regular Standup Meeting 87, 89
Release 52
Retrospektive 87, 91
– Ablauf 93
Risiko 198
– technisches 199
Risikoarten 199
Risikobetrachtung 197
Risikobewertung 169, 200, 204
– freie 201
– modellbasierte 204

Risikomanagement 199
Rückfluss 65

S

Schätzen 12
Schätzungen 127
Sequencing 187
Serviceeinheit 24
Service Level Agreement 137
Services
– Aggregation von 116
– geteilte 119
– Verbinden von 117
Skalierung 105, 114
Slack 12
Spezialisierung 176
Spezialisten 82
Spezialistenengpass 81
Spice Girls 28, 170
Spice Girls Meeting 38, 87, 193
 siehe Nachschubmeeting
Stabilität
– eines Systems 159
Stabilitätsmuster 161
Stakeholder 88
Staumanagement 170
Storypoints 130
Strategie 32
Streudiagramm 134
– Interpretation 137
Swimlane 50
System
– instabiles 161
– stabiles 164

T

Tasks 48 f., 52
Taskswitching 44

Ticket 49
Toyota Production System 18
T-Shaped-Engineers 84

U

Übergabepunkt 171
Unsicherheit 198
Upstream 194
User Story 52

V

Velocity 130
Verbesserungsarbeit 102
Verzögerungskosten 174, 204
– Bestimmen der 182
– Quantifizierung 179
Visualisierung 16, 21, 37, 217

W

Wahrscheinlichkeit 198
Wartezeit 106
Wasserfallprinzip 31
Wert 47, 176, 216
Wertgenerierung 24
Wertschöpfungskette 106
Wertstrom 47
WIP-Limits 40
Wissensarbeit 18
Wissenstransfer 79
Work in Progress 21, 130
Work Item 49

Z

Zeitstempel 132
Zufallsprinzip 140