

LEAN FLOW

Um guia rápido para transformar
com o lean digital

ELISA GRANHA LIRA

© 2023, Elisa Granha Lira. Todos os direitos reservados e protegidos pela Lei 9.610 de 19/02/1998. Nenhuma parte deste livro, sem autorização prévia por escrito da autora, poderá ser reproduzida ou transmitida sejam quais forem os meios propagados: eletrônicos, mecânicos, fotográficos, gravação ou quaisquer outros.

FICHA TÉCNICA

Capa

Elisa Granha Lira
Rodrigo Cabido

Projeto Gráfico

Rodrigo Cabido

Ilustrações

Elisa Granha Lira

Texto, conteúdo e diagramação

Elisa Granha Lira

Dados Internacionais de Catalogação na Publicação (CIP) (Câmara Brasileira do Livro, SP, Brasil)

Lira, Elisa Granha

Lean flow [livro eletrônico] : um guia rápido para transformar com o lean digital / Elisa Granha Lira. -- 1. ed. -- Belo Horizonte, MG : Ed. da Autora, 2023.

ePub

Bibliografia.

ISBN 978-65-00-69710-0

1. Administração - Estudo e ensino 2. Lean Digital 3. Manufaturas - Inovações tecnológicas 4. Produtividade 5. Transformação digital I. Título.

23-155822

CDD-658.11

Índices para catálogo sistemático:

1. Metodologia Lean : Administração 658.11

Henrique Ribeiro Soares - Bibliotecário - CRB-8/9314

Este livro é uma obra independente. Muito zelo e técnica foram empregados em sua edição pela autora. No entanto, podem ocorrer erros de digitação, impressão ou dúvida conceitual. Em qualquer dessas hipóteses, solicitamos a comunicação da questão para que possamos esclarecê-la.

A autora se empenhou em citar adequadamente e dar o devido crédito a todos os detentores de direitos autorais dos materiais que serviram de base para a escrita deste livro. Se necessário, acertos posteriores serão realizados caso se verifique, inadvertida e involuntariamente, a omissão de algum deles.

Aos processos e às pessoas que nos transformam e nos fazem fluir. Em especial, ao meu marido Rafa e ao meu filho Eduardo.

**A TARTARUGA CAUSA MENOS
DESPERDÍCIO E É MUITO MAIS
DESEJÁVEL QUE A RÁPIDA
LEBRE. O SISTEMA TOYOTA DE
PRODUÇÃO SÓ PODERÁ SER
REALIDADE QUANDO TODOS
FOREM TARTARUGA.**

TAIICHI OHNO

SOBRE A AUTORA



Elisa Granha Lira é professora do curso de Engenharia de Produção da Universidade Federal do Rio de Janeiro (UFRJ) em Macaé. Conheceu o *lean* a partir de suas experiências como engenheira na indústria e acabou se apaixonando. Hoje, na academia trabalha para que a filosofia *lean* se estenda para além da universidade e transforme pessoas e organizações.

Outras obras:



Estudo de tempos e movimentos:
uma abordagem *lean* para aumentar a
eficiência de processos físicos e digitais.



Site oficial de seus livros:
www.TemposEficientes.com.br

PREFÁCIO

Antes de mais nada, acredito ser importante pontuar o objetivo deste livro: atingir pessoas das mais diversas formações em linguagem acessível. Afinal, todos fazemos parte do processo de transformação digital que, de forma direta ou indireta, vem revolucionando o trabalho em organizações tanto públicas quanto privadas.

E o que é *lean manufacturing*, produção enxuta ou simplesmente *lean*, para os íntimos? Ou, indo mais ao cerne da questão: o que é *lean digital*? Realço essas questões porque imagino que alguns leitores possam estar se perguntando isso agora.

Portanto, antes de qualquer coisa é importante deixar claro: este livro não é um dicionário, um passo a passo, um discricionário, etc. Em um mundo onde a tecnologia faz o trabalho braçal, basta pesquisar no Google, perguntar ao chat-GPT, que seguramente se obterão respostas melhores a essas questões superficiais do conhecimento.

Meu interesse é trabalhar a parte submersa do *iceberg*, a parte que, pela minha experiência teórica e prática, é o estado da arte acessível a poucas pessoas. Assim, convido desde já o leitor para um espaço de construção do conhecimento em que desenvolveremos esses conceitos em um processo didático ao longo do livro.

Começando, então, vamos trabalhar o conceito de *lean*. Se você perguntasse a experts de *lean* ao redor do mundo “o que é lean”, provavelmente receberia diferentes respostas. Por exemplo:

“É uma filosofia de redução de desperdícios baseada no Sistema Toyota de Produção.”

“É um sistema de gestão que busca aumentar a eficiência e a produtividade reduzindo erros na produção industrial.”

“É um sistema de produção em que as necessidades do cliente puxam aquilo que deve ser produzido.”

Essas respostas não estão erradas, mas são recortes de diferentes ângulos do que é o *lean manufacturing*.

O *lean* envolve basicamente três níveis de abstração (Figura 1). Em um nível de abstração maior, pode ser definido como uma filosofia de produção, isto é, como uma forma de ser e de pensar. Trata-se, afinal, de uma filosofia com vistas a desenvolver uma cultura e valores em que naturalmente todos os funcionários colaborem na melhoria contínua em seu dia a dia de trabalho através da eliminação de desperdícios.

Em um nível de abstração intermediário, podemos ver o *lean* como um método de planejamento e controle: sistema de produção que se preconiza (programação puxada), ou seja, a necessidade do cliente deve puxar aquilo que deverá ser produzido; controle *kanban*; nivelamento de produção; e sincronização de fluxo.

Finalmente, em um nível mais basal, *lean* também é uma caixa de ferramentas, as quais não necessariamente são exclusivas dessa filosofia gerencial.

Acho importante frisar a existência dos três níveis para deixar claro que, neste livro, trabalharemos o nível mais alto de abstração, ou seja, a definição filosófica do *lean*, que objetiva transformar nosso jeito de pensar, agir e ser. Os outros níveis mais tangíveis podem ser facilmente encontrados em livros de referência, na internet ou com a ajuda de nossos amigos inteligentes artificialmente como o chatGPT.

Em outras palavras, a meu ver qualquer leitor pode, por exemplo, construir um mapa de fluxo de valor, mas são poucos que efetivamente vão utilizar a ferramenta para gerar mudanças reais e sustentáveis em uma organização e em seus integrantes. Por isso prefiro focar nesta questão filosófica: afinal, quantos experts em *lean*, com anos de experiência em diversas empresas e lindos rótulos no LinkedIn, são realmente *lean* quando adentramos as camadas mais profundas de seu ser?

Lean para mim é praticamente uma terapia e envolve muita (mas muita mesmo) autocrítica, humildade e coletividade. Já começa aí o desafio em um mundo cada vez mais superficial e individual.

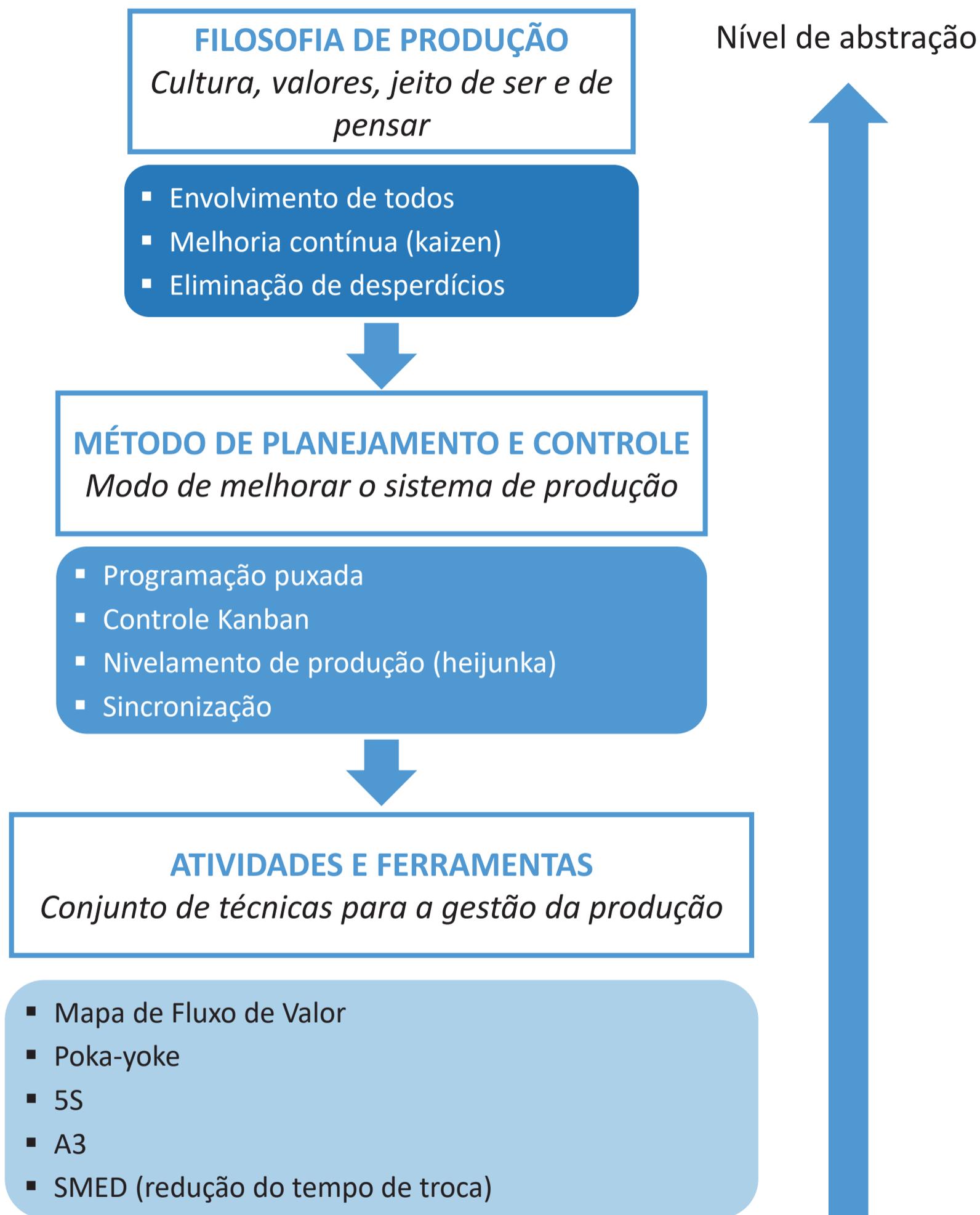


Figura 1 – Os diferentes níveis de abstração de *lean manufacturing*

Finalizemos com esta reflexão do livro “Sistema Toyota de Produção”, de Taiichi Ohno, em que o autor faz alusão à parábola da tartaruga e da lebre:

**“A TARTARUGA CAUSA MENOS
DESPERDÍCIO E É MUITO MAIS DESEJÁVEL
QUE A RÁPIDA LEBRE. O SISTEMA TOYOTA
DE PRODUÇÃO SÓ PODERÁ SER REALIDADE
QUANDO TODOS FOREM TARTARUGA.”**

Vivemos em um mundo cada vez mais acelerado, mas será que cada vez mais eficiente e produtivo? Taiichi Ohno, gênio por trás do Sistema Toyota de Produção (STP) e que deu origem ao *lean*, afirmava que o STP só daria certo quando todos os seus funcionários fossem tartarugas. Isso reforça que devemos objetivar o aumento de produtividade, mas não apenas de forma quantitativa. Vale reforçar que devemos sempre considerar o ser humano por inteiro, aproveitando de forma saudável todas as suas potencialidades e capacidades. Fica, assim, o convite: vamos juntos nessa jornada de transformação?

Começando então, a melhor forma de entender o *lean* é a partir da compreensão do fluxo nos processos. Niklas Modig e Pär Åhlström fazem isso muito bem em seu livro “Isto é *lean*: resolvendo o paradoxo da eficiência” (por sinal, super recomendo a leitura). Assim, usarei como base o *mindset* desses autores para definir o que é e o que NÃO é *lean*, e em seguida, ampliar os conceitos e reflexões para o mundo do *lean digital*.

Entrando em detalhes, primeiramente, irei apresentar o que é *lean*, que em resumo significa ter um processo enxuto e focar em satisfazer as necessidades da unidade de fluxo. Na sequência, serão apresentadas as diretrizes para fazer um processo fluir – afinal, elas são a base para tornar um processo cada vez mais *lean*. No terceiro capítulo será tratado o que NÃO é *lean*: os desperdícios do foco nos recursos e o paradigma da produtividade. Por último, será apresentado um estudo de caso de transformação digital para que esses conceitos se tornem mais tangíveis e possamos começar a adentrar o mundo do *lean digital*.

Elisa Granha Lira

CONTEÚDO

O QUE É LEAN?	11
<i>É ser enxuto</i>	11
<i>É focar no fluxo</i>	15
COMO FAZER UM PROCESSO FLUIR?	18
<i>A Lei de Little</i>	18
<i>A Teoria das Restrições</i>	21
<i>A variabilidade</i>	22
<i>Criando um processo enxuto a partir dos três princípios</i>	23
O QUE NÃO É LEAN?	26
<i>Os problemas de não enxergar o todo</i>	26
<i>O paradigma da produtividade</i>	30
O QUE É LEAN DIGITAL?	32
<i>Estudo de caso: o que NÃO é lean digital?</i>	32
<i>Análise do caso</i>	34
RUMO A UM MUNDO LEAN DIGITAL	44
REFERÊNCIAS BIBLIOGRÁFICAS	48

O QUE É LEAN?

Neste capítulo definiremos o que é *lean*. Resumindo, quando focamos em satisfazer as necessidades da unidade de fluxo, naturalmente tornamos nosso processo o mais enxuto possível.

É ser enxuto

Para se entender o fluxo de um processo, primeiro é importante compreender o que flui. Chamaremos, pois, esses objetos em transformação de unidades de fluxo. Essas unidades podem ser materiais, informações ou pessoas:

- **Materiais:** em indústrias, materiais são processados por máquinas e pessoas até se produzir um produto (carro, geladeira, etc.);
- **Informações:** funções e serviços administrativos processam informações como requerimentos, processos e outros tipos de documentos – podendo ser físicos ou digitais;
- **Pessoas:** em serviços, geralmente as unidades de fluxo podem ser pessoas que vão sendo “transformadas” pelas atividades do processo. Por exemplo: em um salão de beleza, uma pessoa tem sua estética transformada pelo cabeleireiro e produtos de beleza. Já em um hospital, um paciente é curado de eventuais doenças por médicos e outros profissionais de saúde.

Importante chamar atenção para isso a fim de mostrar como o *lean* é universal independentemente do contexto. Muitas vezes esbarramos com os termos *lean healthcare*, *lean office*, *lean digital*, mas a essência é sempre a mesma.

Outra questão a se pontuar é que ser *lean* não é um atributo: você ou um processo é ou não *lean*. Na verdade, trata-se de uma variável contínua, em que sempre podemos nos tornar cada vez mais *lean*!

Portanto, quando queremos definir em que grau um processo é *lean*, estamos interessados em saber o tempo durante o qual a unidade que flui recebe valor em relação ao tempo total que ela gasta durante todo o processo.

$$\text{Porcentagem de valor agregado} = \frac{\text{Tempo das atividades que agregam valor}}{\text{Tempo de atravessamento (lead time)}}$$

Ou seja, o quanto um processo é *lean* pode ser avaliado a partir do cálculo da porcentagem de valor agregado, que analisa a quantidade de tempo que se leva desde a identificação de determinada necessidade até se satisfazê-la. Desse modo, quanto maior é a porcentagem de valor agregado (%VA), mais *lean* é um processo.

Vale salientar ainda que uma atividade agrega valor quando atende as necessidades da unidade de fluxo, enquanto outra que não agrega valor é um desperdício que consome tempo do processo sem transferir valor à unidade de fluxo.

Já o tempo de atravessamento, também conhecido como *lead time*, refere-se ao tempo necessário para se realizar um processo do início ao fim. Assim, durante o tempo total de atravessamento de uma unidade de fluxo, ela será transformada por atividades que agregam valor e permanecerá parte desse tempo sendo transportada ou em estoque, ou seja, em atividades que não agregam valor (Figura 2).



Figura 2 – *Lead time* subdividido em atividades que agregam ou não valor

Repare que, para aumentar a porcentagem de valor agregado, tentar aumentar o tempo das atividades que agregam valor ao processo não é boa estratégia. Afinal, esse tempo é relativamente pequeno em relação ao tempo de atravessamento. Na verdade, a proporção de transferência de valor deve ser maximizada ao se eliminarem os desperdícios.

As atividades que não agregam valor são apresentadas no *lean* como os 8 tipos de desperdícios (Figura 3), os quais consomem tempo, recursos ou espaço, mas não necessariamente contribuem para satisfazer a necessidade da unidade de fluxo.

Segundo Taiichi Ohno, essas oito categorias de desperdícios são:

- 1. Superprodução:** esse desperdício ocorre quando a empresa produz mais do que precisa para atender o cliente. É considerado o pior tipo de desperdício, a mãe de todos os outros, uma vez que quando se produz em excesso são gerados estoques elevados, movimentações desnecessárias, e recursos são utilizados desnecessariamente, consumindo capacidade produtiva que poderia ser utilizada para atender as necessidades do cliente. A causa da superprodução é a falta de previsibilidade de seus processos internos ou de suas relações externas com clientes e fornecedores. Por essa razão, o controle da variabilidade deve ser sempre um objetivo a ser alcançado;
- 2. Espera:** quanto mais contínuo e fluido o fluxo produtivo, maior a eficiência e a rapidez com que podemos entregar o pedido a um cliente. Assim, esperas estão relacionadas a pessoas, máquinas ou informações “paradas”;
- 3. Estoques:** são problemáticos, uma vez que escondem problemas ao retardar, por exemplo, a detecção de defeitos, gerando muitas vezes retrabalhos em grandes lotes. Além disso, representam capital imobilizado, isto é, um dinheiro que é investido em insumos e cujo retorno só ocorrerá quando o cliente receber e pagar pelo produto ou serviço adquirido. Logo, deve-se definir a quantidade necessária de estoque a ser utilizada para amortecer possíveis variabilidades no processo, mas sem criar custos excessivos;
- 4. Transporte:** trata-se do desperdício relacionado a movimentos desnecessários. Esse desperdício pode ser fruto de fluxos truncados, estoques intermediários distantes das linhas de produção e esquemas de abastecimento ineficientes;
- 5. Movimentação:** o desperdício de movimentação envolve movimentos desnecessários. Por exemplo, a procura desnecessária por materiais ou informações é perda de tempo. Assim, as estações de trabalho, os *layouts* e os estoques devem ser planejados de forma a se otimizarem os movimentos que deverão ser realizados por seus funcionários. As movimentações desnecessárias não se restringem aos recursos físicos; as informações localizadas em computadores e outros bancos de dados devem ser organizadas a fim de otimizar a gestão do conhecimento da empresa e permitir que seus funcionários possam rapidamente acessá-las;

- 6. Processamento desnecessário:** trata do desperdício de se realizarem operações que, se fossem eliminadas, não fariam a menor diferença para o cliente. Por exemplo, inspeções de qualidade excessivas podem gerar mais custos que benefícios. Nesse sentido, o processamento excessivo de informação também é um desperdício, e muitas empresas e pessoas não sabem lidar com isso. Assim, engenheiros recebem altos salários para na maior parte do seu tempo ler e responder e-mails, elaborar e assistir a apresentações, atualizar indicadores, preencher relatórios e participar de reuniões. Obviamente, essas atividades são necessárias, mas não deveriam consumir mais do que 30% do tempo de trabalho de uma pessoa. A verdade é que engenheiros, por exemplo, hoje em dia gastam até 90% de seu tempo com essas atividades que muitas vezes não agregam valor. Ao contrário, eles deveriam permanecer a maior parte de seu tempo enxergando e investigando os problemas *in loco* no gembá. Este é um termo japonês que significa “local real” ou “local em que as coisas acontecem”, normalmente utilizado para a área de produção de uma indústria ou qualquer lugar em que ocorre o trabalho que agrega valor a um produto, a uma pessoa ou a uma informação;
- 7. Defeitos:** o ideal é sempre “produzir certo da primeira vez”. Esse desperdício diz respeito, pois, a comprometer recursos financeiros, humanos e tempo para refazer, corrigir ou retrabalhar o que foi feito de forma errada;
- 8. Recursos humanos:** esse oitavo desperdício não foi originalmente descrito por Ohno, mas nem por isso deve ser negligenciado. Trata-se do desperdício intelectual das pessoas que trabalham para uma empresa. Não utilizar o potencial das pessoas de uma empresa é desperdício muito sério. A ideia não é pensar os funcionários como máquinas que devem ter a maior quantidade de projetos e atividades possíveis alocadas a si. Isso pode ser “genial” a curto prazo, mas a médio e longo prazo pode ser um suicídio. Se seu funcionário for bom e proativo, provavelmente irá para outra empresa, ou se for acomodado é possível que venha a ter problemas de saúde física e mental que poderão comprometer sua empregabilidade. Assim, as empresas devem contratar e gerenciar seus recursos humanos de forma inteligente, estimulando um ambiente de trabalho que permita que o funcionário desenvolva e pratique suas capacidades e habilidades de forma plena como ser humano.



Figura 3 – Os 8 tipos de desperdícios

É focar no fluxo

A porcentagem de valor agregado evidencia o foco de um processo. Um processo com alta porcentagem de valor agregado é um processo com foco em satisfazer a necessidade da unidade de fluxo (Processo A – Figura 4). Já um processo com baixa porcentagem de valor agregado é um processo cujo foco está nos recursos (Processo B – Figura 4), isto é, o foco está em se utilizarem ao máximo os recursos dessa organização, como equipamentos e pessoas.

Ainda hoje, a maioria das organizações adota tal perspectiva. Existe um relevante motivo para essa abordagem dominar o mundo empresarial atual: aparentemente, à primeira vista, faz sentido, sob uma perspectiva econômica, focar na eficiência interna dos recursos pensando no custo de oportunidade. A questão é que, se focamos demasiadamente nos recursos, alcançamos apenas um ótimo local, que dificilmente será traduzido em ótimo global para uma organização.

SE FOCAMOS DEMASIADAMENTE NOS RECURSOS, ALCANÇAMOS APENAS UM ÓTIMO LOCAL, QUE DIFICILMENTE SERÁ TRADUZIDO EM ÓTIMO GLOBAL PARA UMA ORGANIZAÇÃO.

Bom exemplo prático são os projetos de redução de custo que muitas vezes as empresas promovem. Quantas ações de melhoria desses projetos têm seus ganhos financeiros departamentais diretamente evidenciados em ganhos para a empresa como um todo?

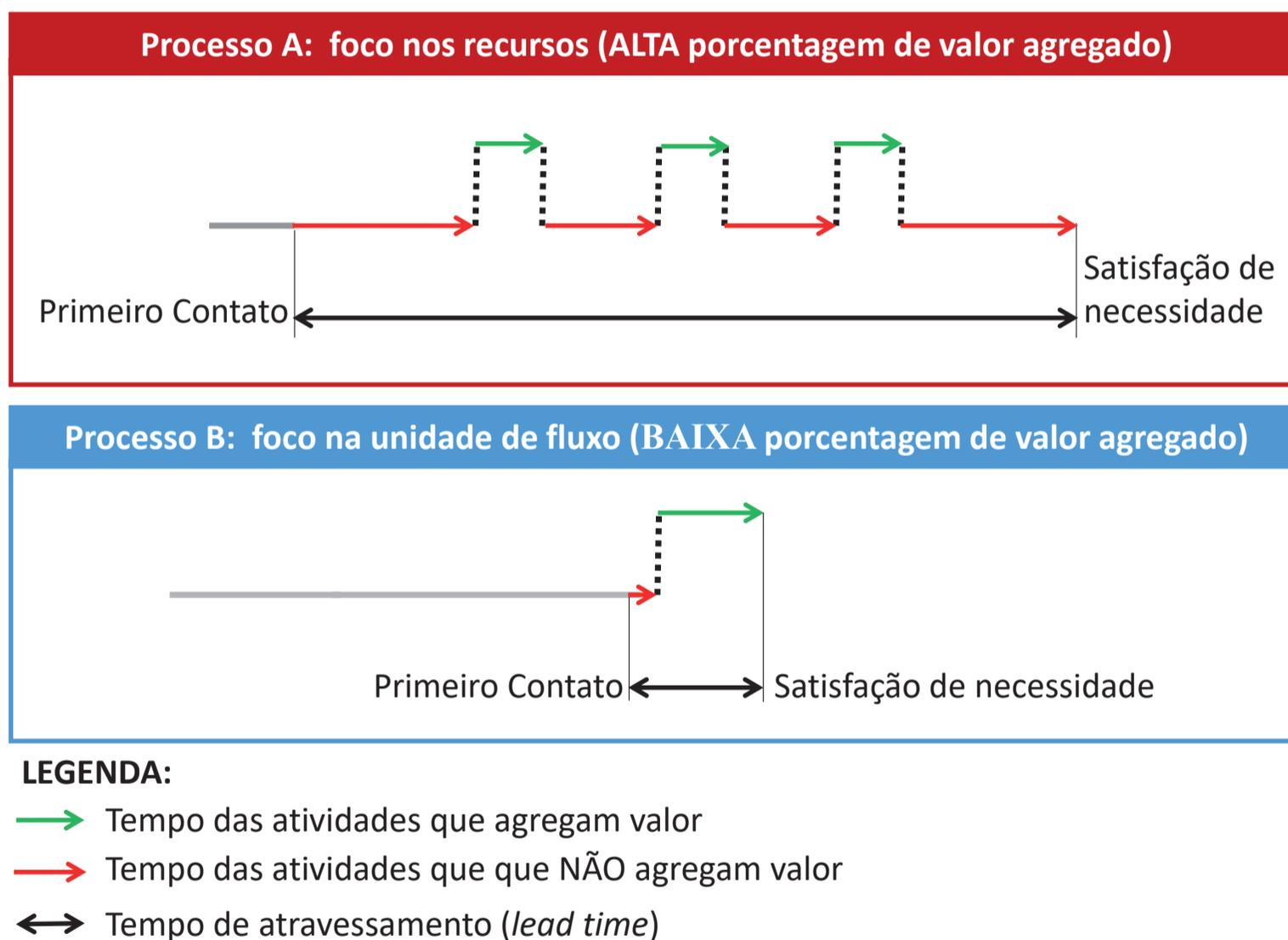


Figura 4 – Processos com foco nos recursos e na unidade de fluxo

Concluindo, enquanto em um processo com foco nos recursos a unidade de fluxo se adapta à organização, em um processo com foco na unidade de fluxo é a organização que se adapta à unidade que está sendo transformada de forma a satisfazer sua necessidade o mais rapidamente possível (Figura 5).

Por exemplo: em um processo de diagnóstico de câncer de mama de forma descentralizada (com várias organizações independentes sendo responsáveis pelas diferentes etapas do processo: consulta inicial, exame de imagem, biópsia, análise laboratorial e conclusão), se formos a uma clínica de imagem, veremos um equipamento de ultrassom (recurso) sempre ocupado com o intuito de processar (realizar exames) a maior quantidade possível de pacientes. Isto é, trata-se de um processo com foco nos recursos. Entretanto, existem centros especializados, principalmente fora do Brasil, em que uma única organização centraliza todas essas atividades. Dessa maneira, um paciente pode realizar todas as atividades em um único dia, otimizando o tempo necessário para o diagnóstico e os resultados de seu tratamento. Ou seja, nesse segundo caso o foco está em satisfazer a necessidade da unidade de fluxo (diagnóstico de câncer de mama em um paciente), e não em ocupar ao máximo um recurso.

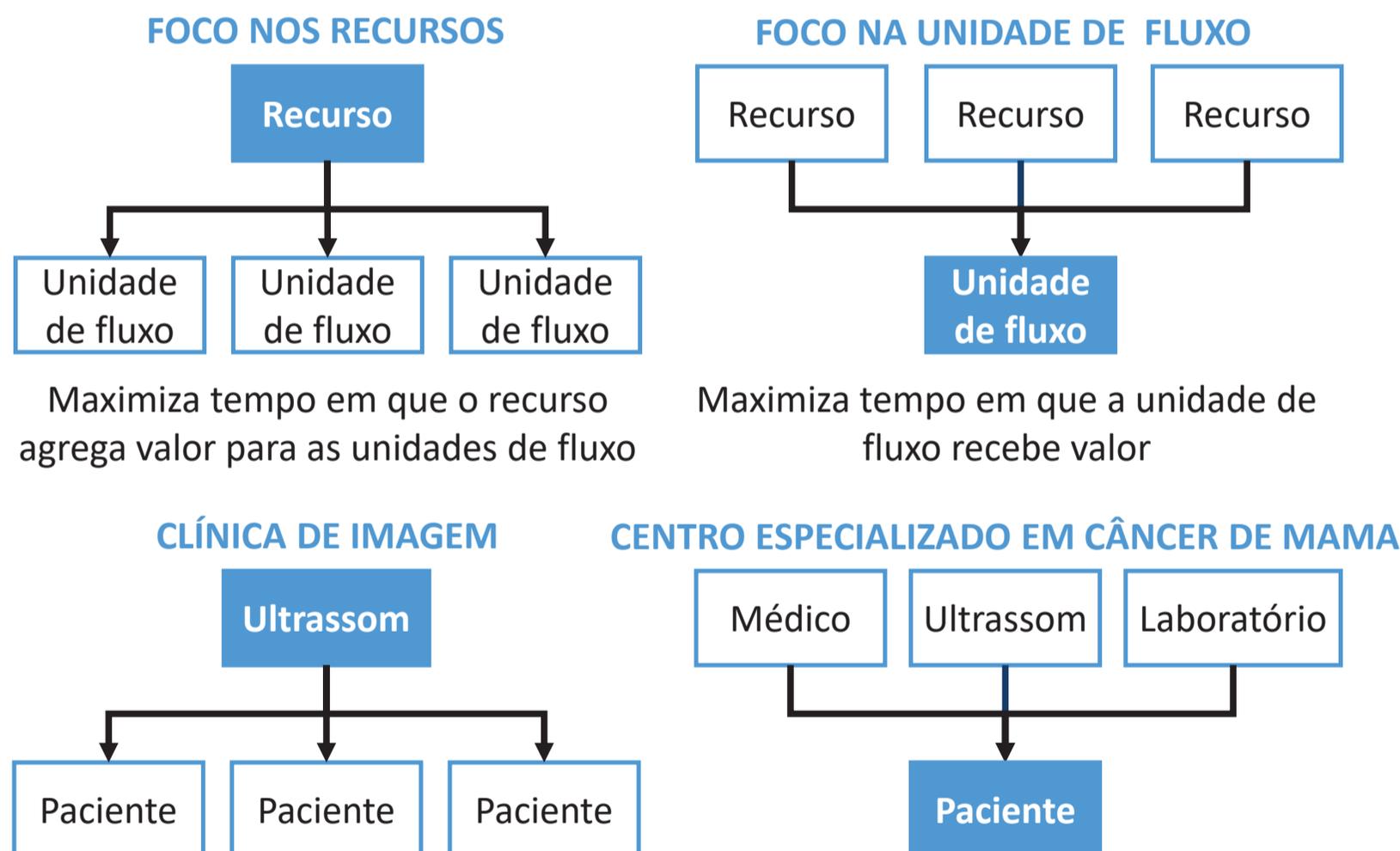


Figura 5 – Processos com foco nos recursos e na unidade de fluxo (exemplo diagnóstico de câncer de mama)

A seguir serão tratadas os princípios por trás dos processos, já que eles são cruciais para entender por que, ao se focar nos recursos, compromete-se a porcentagem de valor agregado de todo o processo.

COMO FAZER UM PROCESSO FLUIR?

Os processos de qualquer organização são formados por atividades que criam a rota que transformará a unidade de fluxo, objetivando satisfazer sua necessidade. Desse modo, para compreender como tornar um processo cada vez mais *lean*, é necessário entender os princípios que regem os processos. Essas “leis” são universais, uma vez que são aplicáveis a qualquer tipo de unidade de fluxo (materiais, informações ou pessoas).

Como visto anteriormente, podemos avaliar o quanto um processo é *lean* por meio da fórmula de porcentagem de valor agregado, a qual depende do tempo de atravessamento. Assim, estamos interessados em entender o que afeta o tempo necessário para se atravessar um processo.

A Lei de Little

O primeiro princípio é a Lei de Little, que é importante na medida em que, quando se gerenciam os fluxos de um processo, três importantes questões devem ser respondidas:

- Quantas unidades de fluxo passam pelo processo dada uma unidade de tempo (taxa de fluxo)?
- Quanto tempo uma unidade de fluxo permanece, na média, dentro dos limites do processo (tempo de atravessamento/*lead time*)?
- Quantas unidades de fluxo estão presentes nos limites do processo (estoque/fila)?

Essa lei relaciona os três conceitos a partir da seguinte fórmula:

$$\text{Tempo de atravessamento} = \frac{\text{Estoque}}{\text{Taxa de fluxo}}$$

O tempo de atravessamento, como já apresentado previamente, diz respeito ao tempo necessário para realizar um processo do início ao fim, a partir dos limites definidos para o sistema.

As unidades de fluxo em processo são todas as unidades que iniciaram o processo, mas ainda não o finalizaram. Isto é, trata-se do estoque de unidades de fluxo dentro dos limites do processo.

A taxa de fluxo se refere ao ritmo em que as unidades de fluxo fluem no processo por unidade de tempo, a qual é dada por unidades de fluxo por unidade de tempo.

A Figura 6 ilustra a lógica por trás da Lei de Little.

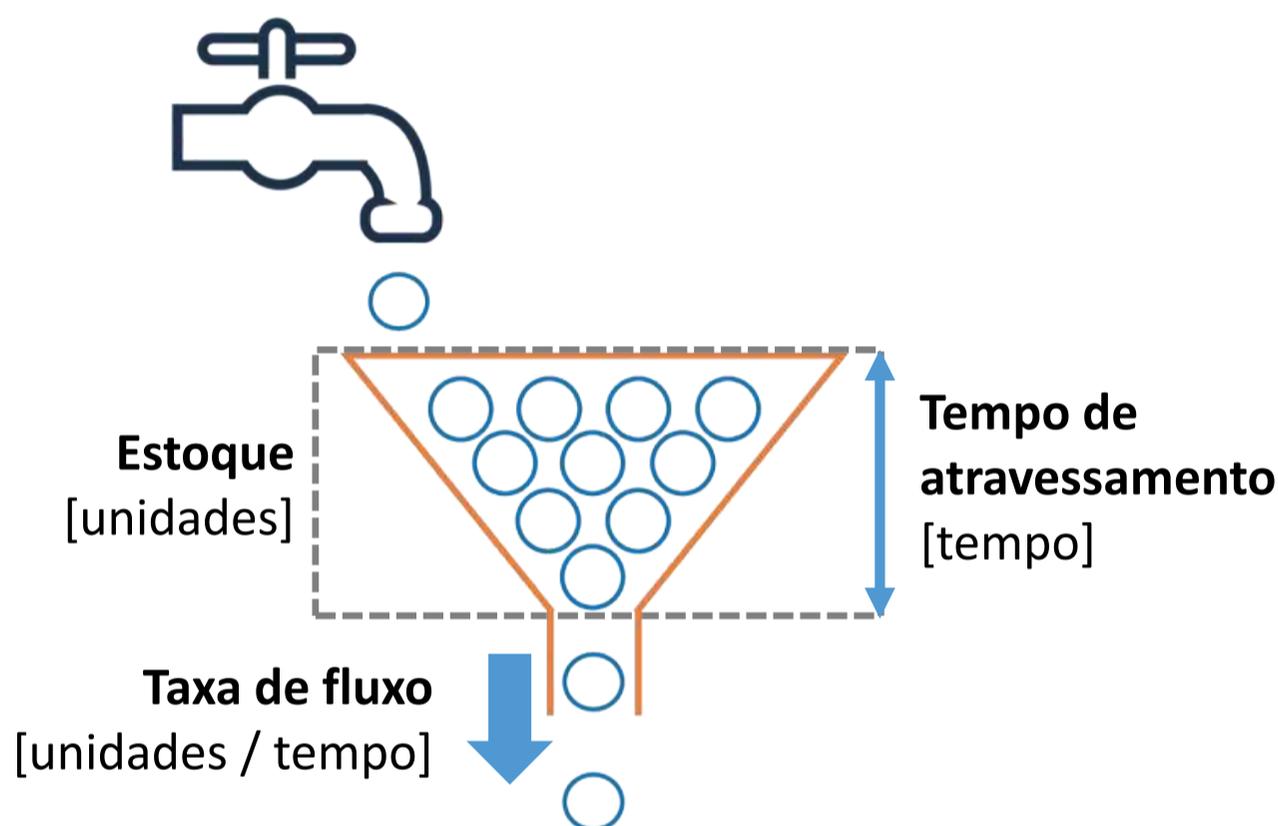


Figura 6 – Lei de Little

Vamos ilustrar a Lei de Little a partir do exemplo de um banco. Se dez pessoas estiverem em uma fila (estoque) nesse banco e uma pessoa for atendida a cada dois minutos pelo caixa (2 minutos por pessoa = 0,5 pessoa por minuto), o tempo de atravessamento será de 20 minutos:

$$\text{Tempo de atravessamento} = \frac{10 \text{ pessoas}}{0,5 \text{ pessoa por minuto}} = 20 \text{ minutos}$$

Ou seja, a décima pessoa dessa fila aguardará, em média, 20 minutos até ser atendida por um caixa (Figura 7).

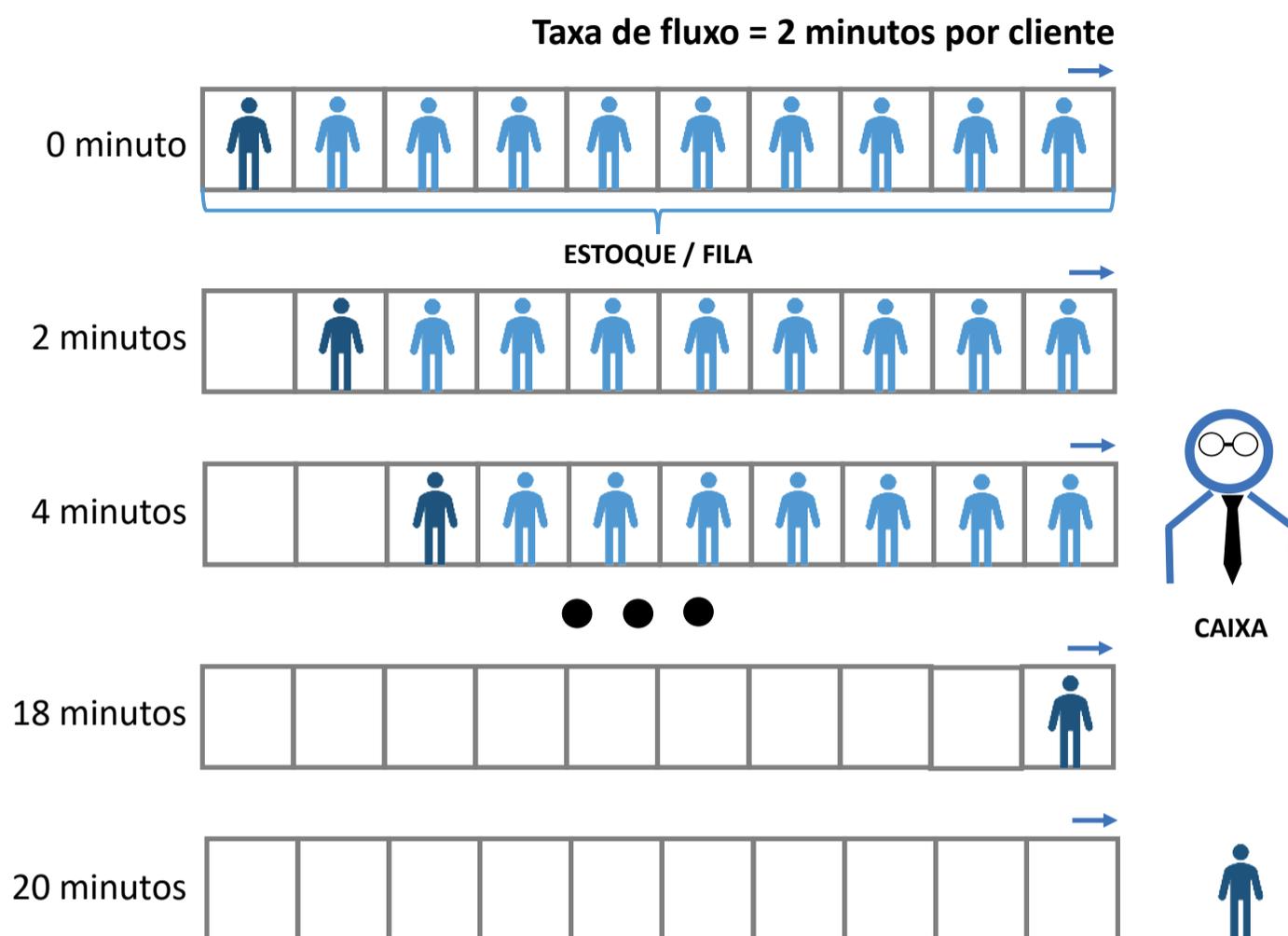


Figura 7 – Lei de Little (Exemplo de um banco)

A Lei de Little demonstra que o tempo de atravessamento é afetado por duas variáveis: o número de unidades de fluxo em processo (estoque) e a taxa de fluxo. Consequentemente, quanto maior o número de unidades de fluxo em processo e menor a taxa de fluxo, maior o tempo de atravessamento.

Isso explica por que o foco nos recursos compromete o fluxo do processo como um todo. Objetivando-se utilizar os recursos ao máximo, deve-se criar uma “fila” de unidades de fluxo antes desses recursos, para que o recurso esteja sempre ocupado. A lógica é que quem deve esperar é a unidade de fluxo e não os recursos, o que, por sua vez, acaba aumentando o tempo de atravessamento e reduzindo a porcentagem de valor agregado. Outro ponto importante é que o recurso está tão preocupado em se ocupar, que acaba não tendo sequer tempo de pensar (desperdício de capital humano) e refletir criticamente sobre eventuais retrabalhos que essa lógica gera em sua rotina de trabalho.

Por exemplo, em um banco, frequentemente nos deparamos com uma fila de clientes (unidade de fluxo) aguardando o atendimento por um gerente (recurso). Assim, a lógica de um processo com foco em recursos é que esses clientes devem esperar, enquanto o gerente deve estar sempre ocupado. Esse excesso de ocupação muitas vezes gera análises com problemas de qualidade e retrabalhos para os gerentes, já que o foco está em atender o maior número de clientes possível e não em resolver de forma assertiva as necessidades das pessoas. Sem contar que o mundo digital permite otimizar os trabalhos das instituições financeiras, automatizando suas atividades e atendendo mais prontamente as necessidades dos clientes. No Brasil, um bom exemplo são as instituições que oferecem contas 100% digitais aos clientes e atendem suas necessidades de forma assertiva. Enquanto o atendimento feito por essas empresas foca na unidade de fluxo, o atendimento realizado em um banco físico com uma fila de clientes foca nos recursos.

A Teoria das Restrições

Os gargalos são estágios do processo ou atividades que, de forma semelhante a um gargalo de garrafa, restringem o fluxo de unidades. Ao restringir esse fluxo, impactam o tempo de atravessamento, já que o estágio com menor taxa de fluxo limita os demais recursos.

Gargalos são um fenômeno natural dos processos, dadas a execução sequencial e a variação na taxa de fluxo de seus estágios ou subprocessos. No livro “A meta: um processo de melhoria contínua”, Eliyahu Goldratt e Jeff Cox ilustram bem como identificar um gargalo a partir dos seguintes sintomas (Figura 8):

- Antes de um gargalo há sempre uma fila/estoque: quando a unidade de fluxo é física, isso é prontamente visualizado. Mas, quando ela é digital, o gargalo pode ser camuflado e passar despercebido;
- Após o gargalo, verificam-se estágios do processo ou atividades que esperam para ser utilizados, ou seja, que possuem capacidade ociosa.

Logo, como os gargalos geram filas de unidades de fluxo a serem processadas, a Teoria das Restrições demonstra, portanto, que os gargalos do processo aumentam o tempo de atravessamento e reduzem a porcentagem de valor agregado.

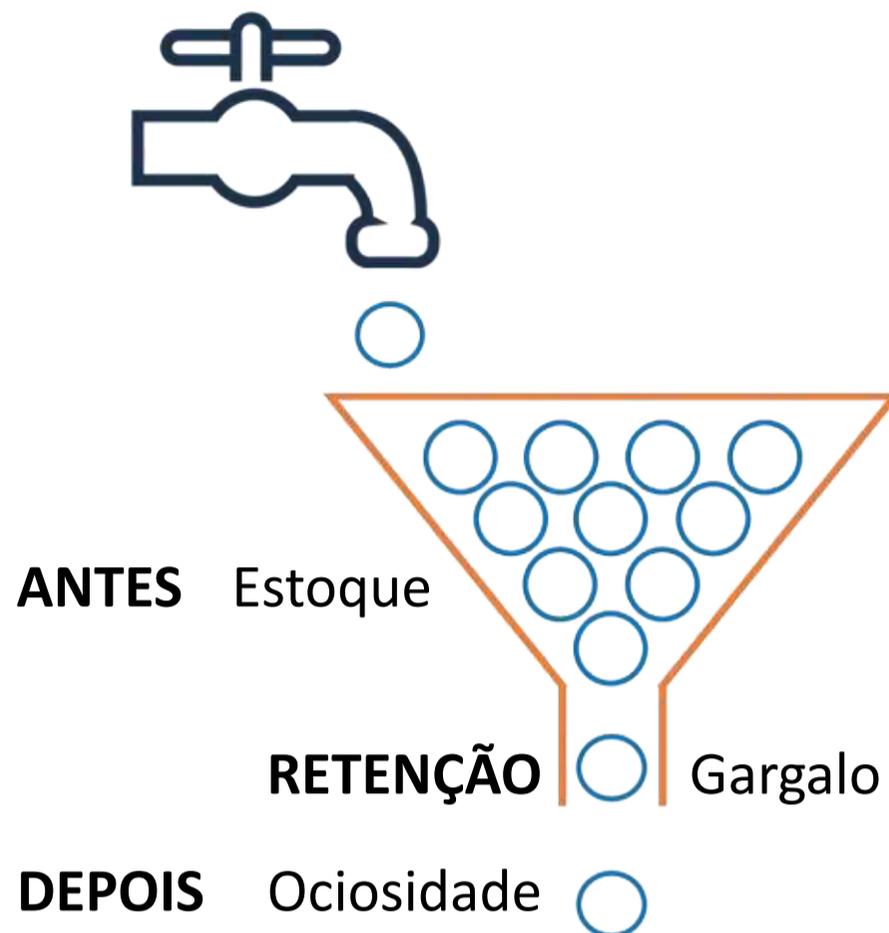


Figura 8 – Teoria das Restrições

A variabilidade

A palavra japonesa *mura* pode ser traduzida como “falta de equilíbrio” ou “variabilidade” e é uma grande fonte de desperdícios, uma vez que gera tanto *muda* (palavra japonesa que pode ser traduzida como “inútil”) quanto *muri* (palavra japonesa que pode ser traduzida como “desproporcional” ou “sobrecarga”). Essas três fontes de desperdício (*muda*, *mura* e *muri*) são conhecidas no *lean* como “3 Ms” (Figura 9), em razão de sua letra inicial.

O ideal seria que os processos fossem respeitados em sua capacidade ideal, sem *muda*, *muri* e *mura*. Ou seja, sem sobrecargas, subutilizações e com a menor variabilidade possível. Consequentemente, os custos seriam otimizados e os riscos de problemas de qualidade e de segurança se reduziriam.

A variabilidade também é inerente aos processos e afeta o tempo de atravessamento. Bom exemplo é o tráfego de veículos nas grandes cidades brasileiras. Se todos os carros em uma grande avenida se mantivessem na faixa adequada e com a mesma velocidade, problemas de trânsito seriam evitados. Mas a realidade é que as avenidas são muitas vezes caóticas: a velocidade de cada veículo varia, os veículos estão constantemente mudando de faixa, sem falar

que situações inesperadas com pedestres, bicicletas ou animais podem ocorrer. Essas variabilidades acabam limitando o fluxo de veículos e gerando filas nas vias de circulação.

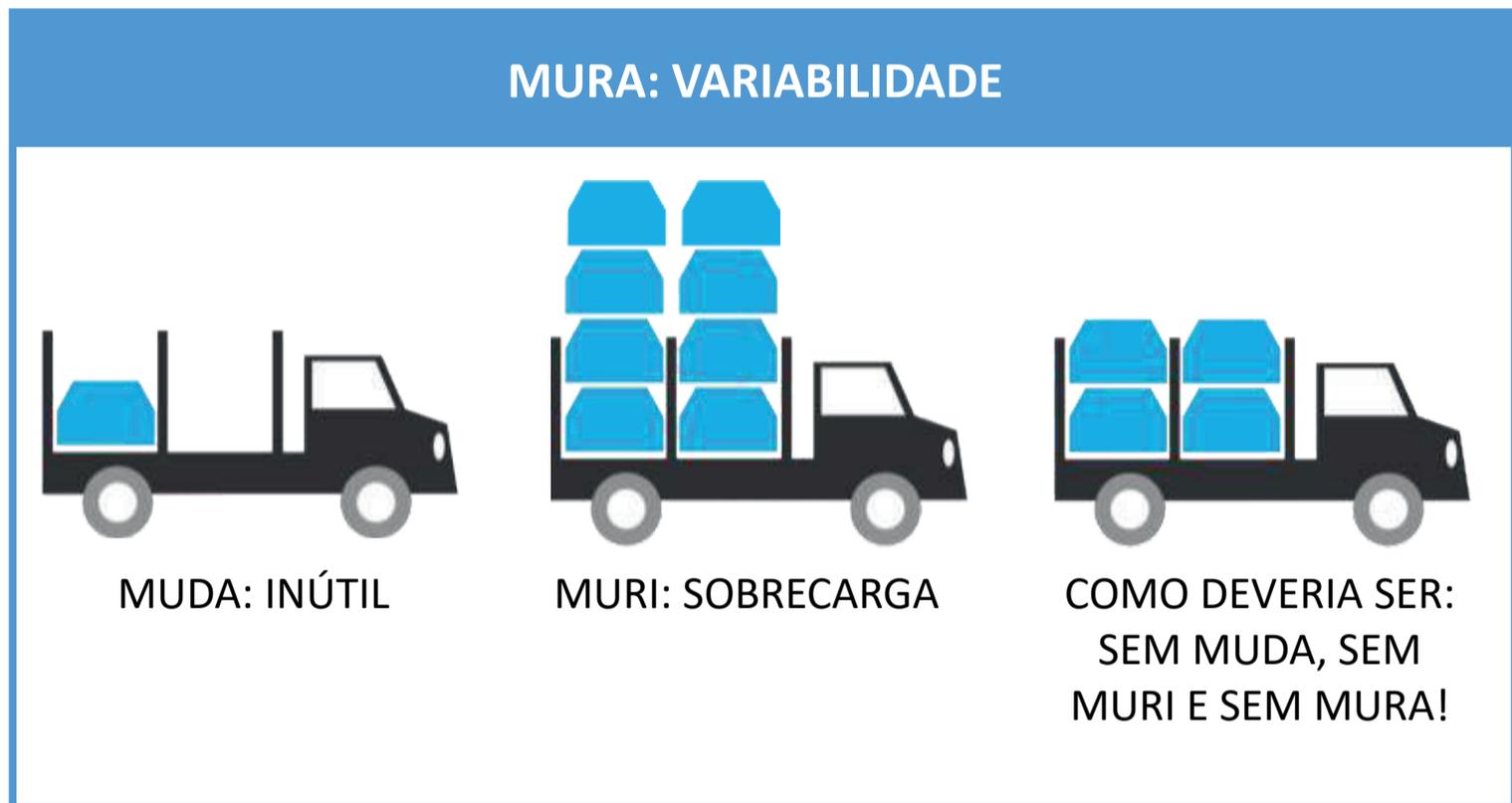


Figura 9 – A variabilidade (*mura*)

Assim, o tempo de atravessamento aumenta quanto maior for a variação do processo. Por sua vez, quando o tempo de atravessamento aumenta, o processo se torna menos enxuto.

Criando um processo enxuto a partir dos três princípios

Portanto, com base nas três diretrizes apresentadas, conclui-se que o tempo de atravessamento é afetado pelas seguintes variáveis do processo: quantidade de unidades de fluxo (estoque); taxa de fluxo; gargalos; e grau de variação (variabilidade).

Podemos, pois, tornar um processo mais enxuto a partir das seguintes ações:

- Redução da quantidade de unidades de fluxo em processo (filas/estoques);

- Aumento da taxa de fluxo, isto é, processar mais unidades de fluxo por unidade de tempo;
- Identificação e eliminação de gargalos;
- Controlar as fontes de variabilidade do processo.

Esses princípios também demonstram a impossibilidade de se focar simultaneamente na unidade de fluxo e nos recursos. Imagine que um gestor objetive alcançar a utilização de 100% de seus equipamentos, ou seja, o gestor quer que os equipamentos estejam 100% do tempo produzindo peças boas.

Para que isso seja alcançado, a lógica natural é trabalhar com uma fila de unidades de fluxo a serem processadas. Quanto maior a variação do processo, maior esse estoque necessário para amortecer a variabilidade e manter os recursos na capacidade máxima. Além disso, gargalos também criam filas adicionais de unidades de fluxo.

Segundo a Lei de Little, uma maior quantidade de unidades de fluxo dentro dos limites do processo (estoque) aumenta o tempo de atravessamento e reduz a porcentagem de valor agregado. Assim, quando se foca em aumentar a utilização dos recursos, compromete-se o fluxo como um todo.

Outro problema do foco na utilização dos recursos é que trabalhos supérfluos, que não agregam valor ao processo, são gerados a partir da criação de necessidades secundárias. Isso é conhecido como paradigma da produtividade:

**ACREDITA-SE QUE, POR SE UTILIZAR
AO MÁXIMO UM RECURSO, ESTE É
PRODUTIVO; MAS, COMO MUITAS DE SUAS
ATIVIDADES SÃO DESNECESSÁRIAS OU
EVITÁVEIS, NA VERDADE SE TRATA DE UM
RECURSO POUCO EFICIENTE. É A FAMOSA
“FAZEÇÃO” OU PRODUTIVISMO POP:
QUANTO MAIS EU FAÇO, MAIS ME OCUPO,
MAIS EFICIENTE EU SOU.**

Afinal, quem é mais produtivo: um funcionário que trabalha dia e noite para bater suas metas ou outro que entrega os mesmos resultados dentro da sua jornada de trabalho regular?

Acredite ou não, ser *lean* começa em nós mesmos. Sem contar os potenciais malefícios dessa sobrealocação de recursos tanto em equipamentos quanto em pessoas no médio/longo prazo. Citando mais uma vez Taiichi Ohno: “A velocidade não tem sentido sem continuidade. Basta lembrar-se da tartaruga e da lebre. Além do mais, não podemos deixar de perceber que máquinas que não são projetadas para resistir a altas velocidades terão sua vida útil encurtada se nós as acelerarmos”.

O QUE NÃO É LEAN?

Como vimos, quando se foca apenas nos recursos sem olhá-los em relação ao todo, a eficiência de todo o fluxo que perpassa esses recursos é afetada. Isso porque ações primárias geram necessidades secundárias, as quais, apesar de aparentemente se mostrarem necessárias e agregadoras de valor, são em verdade, muitas vezes, desnecessárias e representam desperdício financeiro e temporal para as organizações. Desse modo, embora muitas organizações preconizem o aumento da utilização de seus recursos como objetivo primário, muito desse trabalho seria desnecessário se a organização focasse em atender as necessidades da unidade de fluxo de imediato. Neste capítulo serão apresentados os problemas do foco nos recursos e o paradigma da produtividade.

Os problemas de não enxergar o todo

O foco exagerado nos recursos pode ser facilmente diagnosticado a partir de três sintomas, conforme o que foi visto sobre os princípios para fazer um processo fluir:

- Longo tempo de atravessamento (*lead time*);
- Muitas unidades de fluxo (estoque);
- Muitos reinícios por unidade de fluxo.

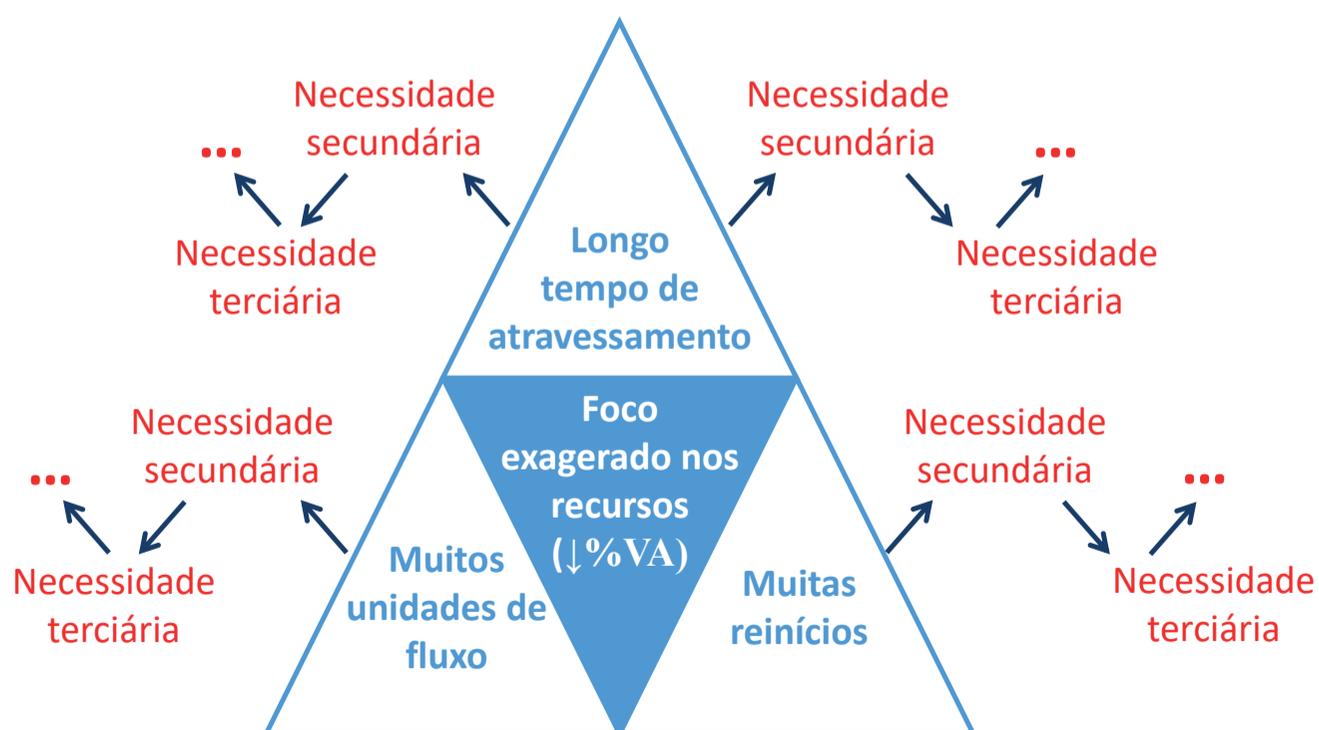


Figura 10 – Efeito cascata de desperdícios

Esses três sintomas, por sua vez, criam no sistema desperdícios em efeito cascata que comprometem a eficiência do fluxo (Figura 10), uma vez que consomem recursos sem agregar valor para a unidade de fluxo.

A Tabela 1 apresenta os problemas que surgem em um processo que foca nos recursos sem olhar o todo, a partir de cada um de seus três sintomas.

SINTOMAS DO FOCO NOS RECURSOS
LONGOS TEMPOS DE ATRAVESSAMENTO
Gera necessidades secundárias
Fecha janelas de oportunidade
MUITAS UNIDADES DE FLUXO
Alto estoque demanda recursos adicionais
Gera necessidades secundárias
Perde-se visão geral do processo
Escondem-se problemas
Criam estímulos produtivistas (fazeção)
Gera necessidades secundárias
Gera estresse e desmotivação
Estimula perda de controle
MUITOS REINÍCIOS POR UNIDADE DE FLUXO
Gera necessidades secundárias
Demanda tempo de preparação mental
Risco de transferência de responsabilidade

Tabela 1 – Os problemas do foco nos recursos

Ilustraremos as ineficiências geradas por um processo com foco em recursos a partir de diferentes contextos para cada um de seus sintomas.

O longo tempo de atravessamento gera necessidades secundárias e fecha janelas de oportunidade. Vamos pegar como exemplo a típica atividade de organizar recibos durante uma viagem, para futuro pedido de reembolso. Trata-se de realidade comum para muitos de nós: vamos juntando os recibos de forma desorganizada na carteira e geralmente, quando já voltamos da viagem e temos tempo livre, aproveitamos para preencher uma planilha com essas informações, digitalizamos os recibos e entramos com o pedido de reembolso no sistema.

A questão é que surgem muitas necessidades secundárias nesse processo, desde organizar os recibos a encontrá-los todos, por exemplo. Veja: se assim que almoçássemos já tirássemos foto do recibo e déssemos entrada do valor em um sistema de reembolso via aplicativo, por exemplo, não seria necessário preencher planilha, organizar os recibos, buscá-los, e não existiria risco de os perdermos. Por isso não dar entrada de imediato, além de criar necessidades secundárias, faz com que percamos a janela de oportunidade de fazer certo da primeira vez.

O sintoma “muitas unidades de fluxo no processo” é prejudicial por gerar necessidades secundárias, esconder problemas e contribuir para se perder a visão global do processo. Assim, o alto estoque acaba demandando recursos adicionais para suprir essas ineficiências.

Exemplo disso é o excesso de informação a que somos submetidos nos dias atuais por e-mail, mensagens de aplicativos e redes sociais. Será que o avanço na tecnologia da informação nos tornou mais eficientes? Altos volumes de informação fazem com que tenhamos dificuldade de ter uma visão geral do que é prioridade e do que é importante. Dessa forma, acabamos perdendo muito tempo na procura e organização das informações (necessidades secundárias). Sem contar que o elevado estoque tende a esconder problemas.

Além do que, essa segunda fonte de ineficiência (alto estoque) cria estímulos produtivistas (“fazeção”) e sobrecarrega os recursos existentes. Fazer muitas coisas ao mesmo tempo aumenta o risco de perda de controle, o que deixa as pessoas frustradas e estressadas, gera problemas de comunicação e cria necessidades secundárias (reuniões, busca, organização e planos de ação). Se até as máquinas são recursos limitados, por que os seres humanos não seriam?

Voltando ao exemplo do excesso de informações a que estamos submetidos, é nítido seu efeito negativo na saúde física e mental das pessoas. Estamos sempre ocupados, somos cada vez mais bombardeados com novas demandas e,

em consequência, acabamos por realizar uma série de retrabalhos porque ao invés de pensar e planejar simplesmente estamos em *looping* no ato de executar. Como se diria antigamente, quando a comunicação ainda se dava por carta, algumas vezes é importante se perguntar: será que vale a pena?

Ao final, muitas estruturas e rotinas acabam sendo criadas pelas organizações para lidar com os problemas oriundos da alta quantidade de unidades de fluxo em processo, as quais não seriam necessárias caso a organização focasse em atender a real necessidade da unidade de fluxo.

A terceira fonte de ineficiência – muitos reinícios por unidade de fluxo – é o famoso “arranca e para”. Quando reiniciamos uma tarefa, é necessário um tempo de preparação mental. Além disso, quando mudamos o foco de uma tarefa para outra repetidamente, nos cansamos mais mentalmente, o que pode acarretar aumento contínuo do tempo necessário para preparação mental.

Por exemplo: quando lidamos com grande quantidade de e-mails, com frequência temos de ler o mesmo e-mail mais de uma vez, sobretudo quando a mensagem tem muitos detalhes e maior nível de complexidade. Assim, acabamos preferindo adiar essa tarefa para um momento mais oportuno.

Os reinícios geram, por conseguinte, necessidades secundárias de releitura, busca e organização desses e-mails, além do risco de esquecimento e retrabalho. Aumenta-se também o risco de transferência de responsabilidade em caso de férias, licença-maternidade, afastamento por motivos de saúde e rotatividade de empregados.

Se necessitamos, nós mesmos, de tempo para lembrar e nos preparar mentalmente para uma atividade que reiniciamos, imagine quando é uma nova pessoa que será responsável por ela... Esse “telefone sem fio” gera distorções na informação e problemas de comunicação que, por sua vez, podem acarretar defeitos e problemas de qualidade. Ou seja, essa fonte de ineficiência acaba sendo convertida em retrabalho e em impactos negativos na satisfação do cliente.

Como pode ser visto, a ineficiência se propaga em efeito dominó, no qual cada uma dessas três fontes de ineficiência influencia as demais. Tal qual discutido previamente, a estratégia de focar no ótimo local, priorizando os recursos em detrimento da unidade de fluxo, possui como sintomas muitas unidades de fluxo em processo e longos tempos de atravessamento, os quais estão diretamente ligados pela Lei de Little. Esses dois sintomas acabam por gerar a terceira fonte de necessidades secundárias: a necessidade de reiniciar continuamente a mesma tarefa.

O paradigma da produtividade

Como vimos, o foco nos recursos cria ilhas “produtivistas”, em que as necessidades das unidades de fluxo são subdivididas em subetapas que são executadas por recursos com alta utilização. Entretanto, cada ilha enxerga apenas a si mesma e não possui visão do todo. Conseqüentemente, o processo é subotimizado com vários ótimos locais, enquanto o ótimo global é comprometido com essa abordagem, criando necessidades secundárias e reduzindo sua porcentagem de valor agregado. Trata-se do que chamamos de paradigma da produtividade, isto é, a ilusão de acreditar ser eficiente, quando na verdade não é, uma vez que atividades que aparentemente “agregam valor” são desperdícios camuflados.

Vale ressaltar que esse paradigma da produtividade existe tanto em nível individual quanto organizacional em todos os contextos: em serviços privados, em indústrias e até em repartições públicas.

Por exemplo, imagine uma empresa siderúrgica, em cuja planta temos diversos departamentos como aciaria, laminação e trefilaria. Esses departamentos também são subdivididos em subprocessos com gestores ou coordenadores específicos. Se a organização não tiver preocupação e não criar um esforço efetivo na melhoria global, o gestor de cada departamento naturalmente vai favorecer interesses locais.

Assim, para se otimizar um sistema, todos os envolvidos devem entender o objetivo geral do sistema e começar a pensar diferente, na medida em que interesses locais nem sempre condizem com interesses globais. Além disso, é crucial contar com uma liderança engajada e com a utilização de métricas de performance compartilhadas entre as funções.

No âmbito individual, esse desperdício com trabalhos supérfluos também passa muitas vezes despercebido: achamos que estamos agregando valor por estarmos continuamente ocupados, mas a realidade é que grande parte de nosso trabalho é desperdício que poderia ser evitado. O trabalho de um engenheiro, por exemplo, muitas vezes envolve tempo considerável com leitura de e-mails, coleta de informações para indicadores de performance, preenchimento de relatórios e reuniões.

É óbvio que algumas dessas atividades são necessárias, mas será que todo esse trabalho agrega valor para a organização? Por isso é importante sempre nos indagarmos:

QUANTO TEMPO DO NOSSO TRABALHO OU DOS NOSSOS COLABORADORES/ EQUIPAMENTOS É GASTO COM ATIVIDADES QUE NÃO NECESSARIAMENTE AGREGAM VALOR E QUE FORAM CRIADAS A PARTIR DE NECESSIDADES SECUNDÁRIAS?

Pode ter certeza de que sempre existem desperdícios camuflados passíveis de ser eliminados, e o autoquestionamento é o primeiro passo para lidar com o paradigma da produtividade. Ao nos questionarmos sobre esse paradoxo de sermos produtivos apenas em nível micro, criamos o desafio de olhar para o fluxo em todo o processo e almejar ganhos globais. Em um sistema cujo foco está em satisfazer as necessidades da unidade de fluxo, o tempo de atravessamento é curto e poucas unidades de fluxo se encontram dentro dos limites do processo (estoque pequeno). Isso acaba por evitar reinícios desnecessários. Necessidades secundárias e trabalhos supérfluos são, por conseguinte, evitados. Literalmente, tudo flui.

O *lean* é uma estratégia que visa a otimizar a eficiência e eficácia do fluxo como um todo. Basta lembrarmos de sua filosofia apresentada na Figura 1: busca-se eliminar desperdícios ao se adotar uma visão global em que todos são responsáveis pela melhoria contínua, em detrimento de pensamentos e estímulos individuais, a partir das necessidades reais dos clientes.

O QUE É LEAN DIGITAL?

Neste capítulo vamos ilustrar os conceitos apresentados previamente em um estudo de caso de transformação digital. Vale ressaltar que, como o próprio nome diz, trata-se de uma transformação em que um processo físico vai gradualmente incorporando tecnologias digitais. Nesse sentido, o *lean digital* nada mais é do que utilizar essas tecnologias digitais para tornar o processo cada vez mais enxuto.

Estudo de caso: o que NÃO é lean digital?

Certa organização criou um projeto com o objetivo de digitalizar seus processos. A ideia é que no futuro os analistas possam realizar seu trabalho remotamente, o que é chamado de teletrabalho. Assim, os atendimentos antes prestados de forma presencial passaram a ser realizados a distância. A mudança, segundo os envolvidos, representará considerável ganho em produtividade para a organização, na medida em que será necessária menor estrutura física. Além disso, com um ambiente de trabalho mais “controlado”, com menos interferências, espera-se que esses analistas consigam analisar maior quantidade de requerimentos por dia.

Com o intuito de criar metas de produtividade para os analistas, mapearam-se suas atividades e realizaram-se cronoanálises a fim de avaliar o tempo necessário para se realizá-las. Avaliou-se, pois, o processo e mediu-se o tempo antes e após sua digitalização.

A Figura 11 representa como o processo era realizado antes de sua digitalização: o solicitante agendava o atendimento em uma das unidades da organização por telefone ou internet. Ao chegar ao local na data agendada, aguardava ser chamado. Quando chegava sua vez, era atendido diretamente por um analista, o qual já realizava a análise do caso na presença do usuário. Caso se detectasse alguma pendência em relação à documentação, o analista requeria o documento ao solicitante. Na maioria das vezes o solicitante tinha o documento em mãos e, assim, já o entregava para o analista dar prosseguimento ao processo. No caso cronoanalisado, a pessoa que solicitou o serviço gastou 28 minutos em atividades que “agregavam valor” (processo de atendimento, resolução de pendências e finalização) em um total de 68 minutos, que foi o tempo total que a pessoa permaneceu nesse local. Os 40 minutos restantes, a pessoa gastou esperando atendimento.

Processo de atendimento, resolução de pendências e finalização (deferimento ou indeferimento)

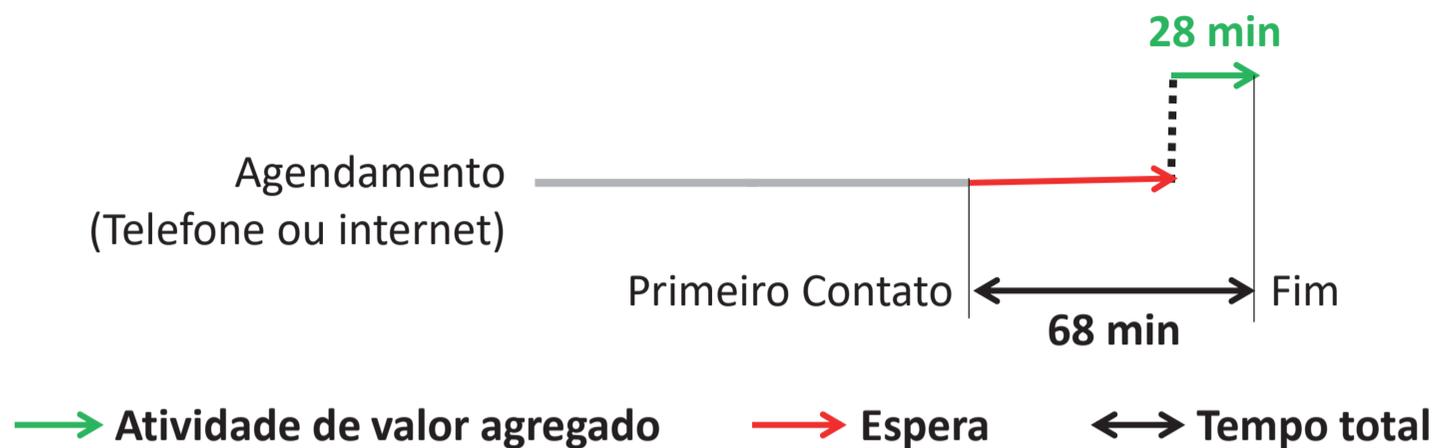


Figura 11 – Processo ANTES de sua digitalização

A Figura 12 ilustra o processo após a digitalização. No futuro, o solicitante também deverá agendar seu atendimento por telefone ou internet, mas na data agendada será atendido por um estagiário, que será responsável por atendê-lo e digitalizar todos os documentos necessários à análise do seu caso. Estabeleceu-se um tempo predeterminado de 20 minutos para cada agendamento. Após esse atendimento inicial, o requerimento entrará então em uma fila virtual de requerimentos a serem analisados. Na Figura 12, as setas verdes representam as atividades que os analistas realizaram e que agregam valor ao processo; já as setas vermelhas, os momentos em que o requerimento ficou parado no sistema aguardando a documentação pendente e a próxima análise. Neste segundo caso, foi necessário que o analista solicitasse a resolução de pendências duas vezes. Assim, o tempo total de trabalho do analista foi de 90 minutos em relação ao tempo total desde sua primeira análise realizada até a data de finalização do processo (18 dias).

Vale ressaltar que, quando falta algum documento necessário, o analista faz uma exigência para o solicitante, a qual deve ser atendida para que a análise do requerimento possa prosseguir. Essa é uma atividade muito curiosa também. Assim que se gera pendência em um processo, o analista escreve em um caderno dicas e informações relevantes desse processo para futura reanálise. Outro ponto que chama atenção é que não existe nem gestão visual, nem interface amigável nesse sistema que ele utiliza. Dessa forma, ele não consegue distinguir facilmente quais processos estão aguardando uma primeira análise ou já voltaram de uma resolução de pendência. Logo, o analista sempre destina par-

te do dia para, a partir de suas anotações do caderno, buscar processos (através do famoso Ctrl+F) que eventualmente tenham voltado para uma segunda ou terceira análise.

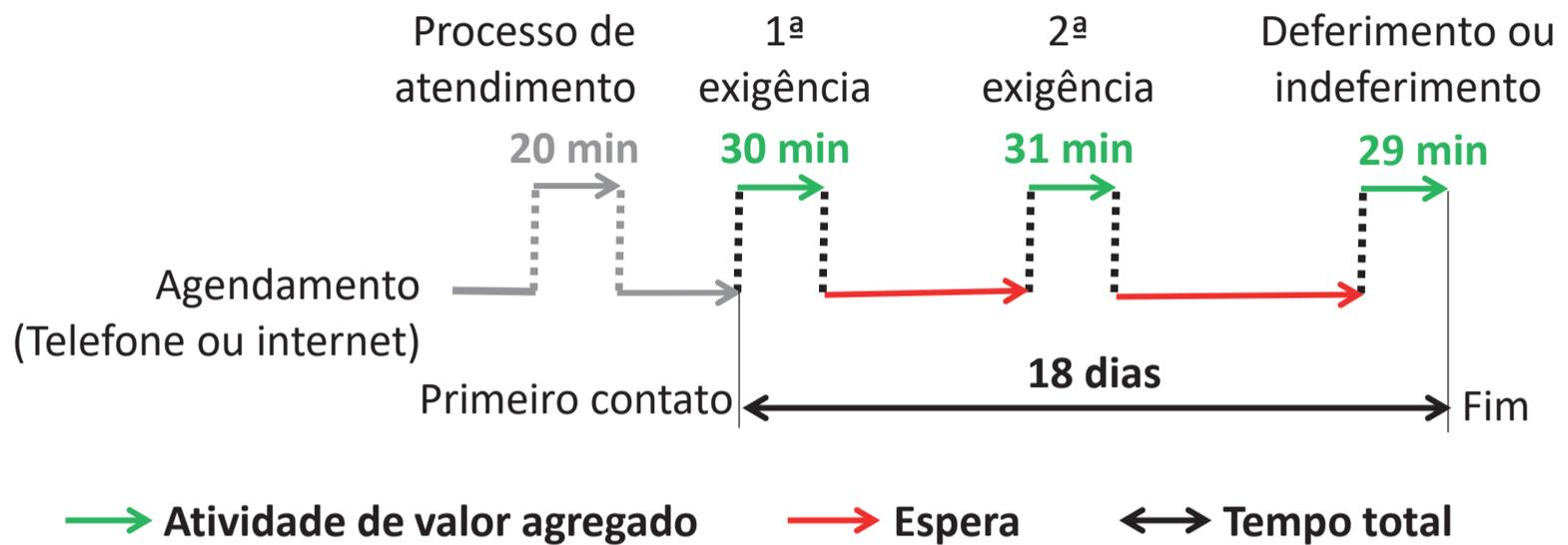


Figura 12 – Processo APÓS sua digitalização

Análise do caso

Primeiramente, é interessante comparar os dois casos em relação à porcentagem de valor agregado (% VA) desde o primeiro contato do recurso analista até a finalização de sua análise:

- Cálculo da porcentagem de valor agregado ANTES da digitalização do processo:

$$\% VA_{ANTES} = \frac{28 \text{ minutos}}{68 \text{ minutos}} \cong 0,41 \rightarrow \mathbf{41\%}$$

- Cálculo da porcentagem de valor agregado DEPOIS da digitalização:

$$\% VA_{DEPOIS} = \frac{30 + 31 + 29 \text{ minutos}}{18 \text{ dias} \times 24 \frac{\text{horas}}{\text{dia}} \times 60 \frac{\text{minutos}}{\text{hora}}} =$$

$$\frac{90 \text{ minutos}}{25920 \text{ minutos}} \cong 0,0035 \rightarrow \mathbf{0,35\%}$$

Esse é um exemplo curioso, porque percebemos que houve redução na porcentagem de valor agregado do processo. Ou seja, percebemos que o processo antes da digitalização tinha foco bem maior na unidade de fluxo em comparação com a situação após as mudanças realizadas. Já após a digitalização, verificou-se baixa porcentagem de valor agregado (< 1%), o que caracteriza um processo que foca na utilização de seus recursos. Isto é, se avaliássemos seus recursos, no caso os analistas, concluiríamos que eles analisam muitos requerimentos por dia, mas finalizam poucos no mesmo período, o que significa que são eficientes retrabalhadores. Esse foco exagerado em ocupar os recursos pode ser evidenciado na forma dos três sintomas que um processo não enxuto apresenta:

- Longos tempos de atravessamento: tempo elevado de análise de um requerimento desde sua entrada no sistema até seu deferimento ou indeferimento;
- Muitas unidades de fluxo: longa fila virtual de requerimentos a serem analisados;
- Muitos reinícios por unidade de fluxo: a falta de documentação impede que ocorra uma única análise pelo analista e cria retomadas em série. Por exemplo: toda vez que um requerimento retorna com uma pendência resolvida, o analista gasta tempo considerável reestudando-o, na medida em que, devido à alta quantidade de requerimentos analisados por dia, ele já não se lembra muito desse que acabou de retornar.

A Tabela 2 apresenta em detalhes as necessidades secundárias e outros problemas que cada um desses sintomas gera nesse processo.

SINTOMAS DO FOCO NOS RECURSOS	
LONGOS TEMPOS PARA ANÁLISE DO PROCESSO	
Gera necessidades secundárias: escrever e-mail solicitando novos documentos	
Fecha janelas de oportunidade: antes poderia solicitar documentos faltantes diretamente para solicitante	
FILA VIRTUAL CRESCENTE DE PROCESSOS AGUARDANDO ANÁLISE	
Alto estoque demanda recursos adicionais (analistas)	
Gera necessidades secundárias: escrever no caderno dados sobre processos interrompidos e dicas para agilizar próxima análise	
Perde-se visão geral do processo: analista tem dificuldade de saber quais processos voltaram para serem reanalisados	
Escondem-se problemas: geram-se problemas camuflados de qualidade e até éticos no processo de análise	
Cria estímulos produtivistas (fazeção)	
Gera necessidades secundárias: reanálises sucessivas	
Gera estresse: analistas sobrecarregados em retrabalhar	
Estimula perda de controle: alguns processos voltam com os documentos que faltavam, mas o analista demora semanas até perceber esse retorno	
CONSTANTES INTERRUPTÕES DOS ANALISTAS PARA SOLICITAR DOCUMENTOS ADICIONAIS	
Gera necessidades secundárias: tira parte do dia para buscar processos que voltaram para serem reanalisados	
Demanda tempo de preparação mental: analista precisa relembrar caso que analisou meses antes	
Risco de transferência de responsabilidade: analista pode pedir licença-maternidade, afastamento por motivos de saúde, trocar de setor ou empresa e até se aposentar	

Tabela 2 – Os problemas do foco nos recursos (estudo de caso)

Esse caso ilustra bem as ineficiências de um processo não enxuto: o longo tempo de atravessamento fecha a janela de oportunidade de o analista poder solicitar os documentos faltantes diretamente ao requerente, o que facilitaria a comunicação e até permitiria sanar no primeiro encontro o problema – afinal, quando isso acontecia antes da digitalização, muitas vezes os cidadãos os tinham em mãos. Ou seja, a própria atividade de solicitação de documentos adicionais pode ser vista como necessidade secundária.

Assim, percebe-se que a demora em atender uma necessidade primária criou reação em cadeia de necessidades secundárias. As várias solicitações de documentação faltante contribuíram para aumentar o tempo de atravessamento, já que nem sempre o requerente conseguia entregar de pronto os documentos pendentes. Toda vez que se abre novo ponto de decisão no processo, a variabilidade é convidada a atuar. Isto é, toda vez que “a bola é passada” para o requerente apresentar os documentos faltantes, muitas coisas podem acontecer até que ele efetivamente consiga solucionar a pendência. Esses efeitos do longo tempo de atravessamento acabam gerando trabalho supérfluo que consome os recursos existentes e ainda demanda a aquisição de novos recursos.

TODA VEZ QUE SE ABRE NOVO PONTO DE DECISÃO NO PROCESSO, A VARIABILIDADE É CONVIDADA A ATUAR.

Exemplificando os problemas gerados a partir do sintoma “muitas unidades de fluxo no processo” nesse estudo de caso, toda vez que solicitava a documentação pendente, o analista escrevia em um caderno informações sobre o respectivo requerimento (necessidade secundária), porque o sistema dificultava identificar visual e automaticamente os processos que tinham as pendências sanadas. Conclusão: muitas vezes uma pendência já estava sanada, mas o analista demoraria alguns dias até perceber isso e prosseguir com a análise do processo. Sem contar que ao final, após essa mudança no processo, o analista trabalha mais, ou seja, é submetido a maior estresse, sem ganho real de produtividade como contrapartida. Afinal, ele retrabalha mais e finaliza menos análises. Eis o paradigma da produtividade na prática. Quem vê o analista, muitas vezes acredita que ele tenha se tornado mais produtivo após a digitalização do processo, já que se tornou mais ocupado. Entretanto, a maior parte de seu trabalho são retrabalhos, isto é, desperdícios camuflados.

Voltando à parábola da tartaruga e da lebre, o analista nitidamente é a lebre que corre mais, mas perde a corrida. Afinal, está cada vez mais ocupado reanalisando processos em que falta documentação, mas finaliza menos.

Em relação às constantes reanálises, estas também geram necessidades secundárias como a atividade do analista de checar ao final do dia quais processos do caderno eventualmente poderiam ter retornado. Além do mais, quando o analista começa uma segunda análise, é como se ele a estivesse analisando pela primeira vez. Afinal, apesar das dicas no caderno, já transcorreu longo período desde que ele analisou aquele processo pela primeira vez. Por fim, mas não menos importante, existe o risco de transferência de responsabilidade. Isto é, o analista pode pedir licença do trabalho, aposentar ou trocar de empresa. Nesse caso uma eventual segunda análise será realizada por outra pessoa e, obviamente, o tempo de atravessamento será ainda maior.

↑ Tempo de atravessamento: Estoque, Gargalo e Variabilidade

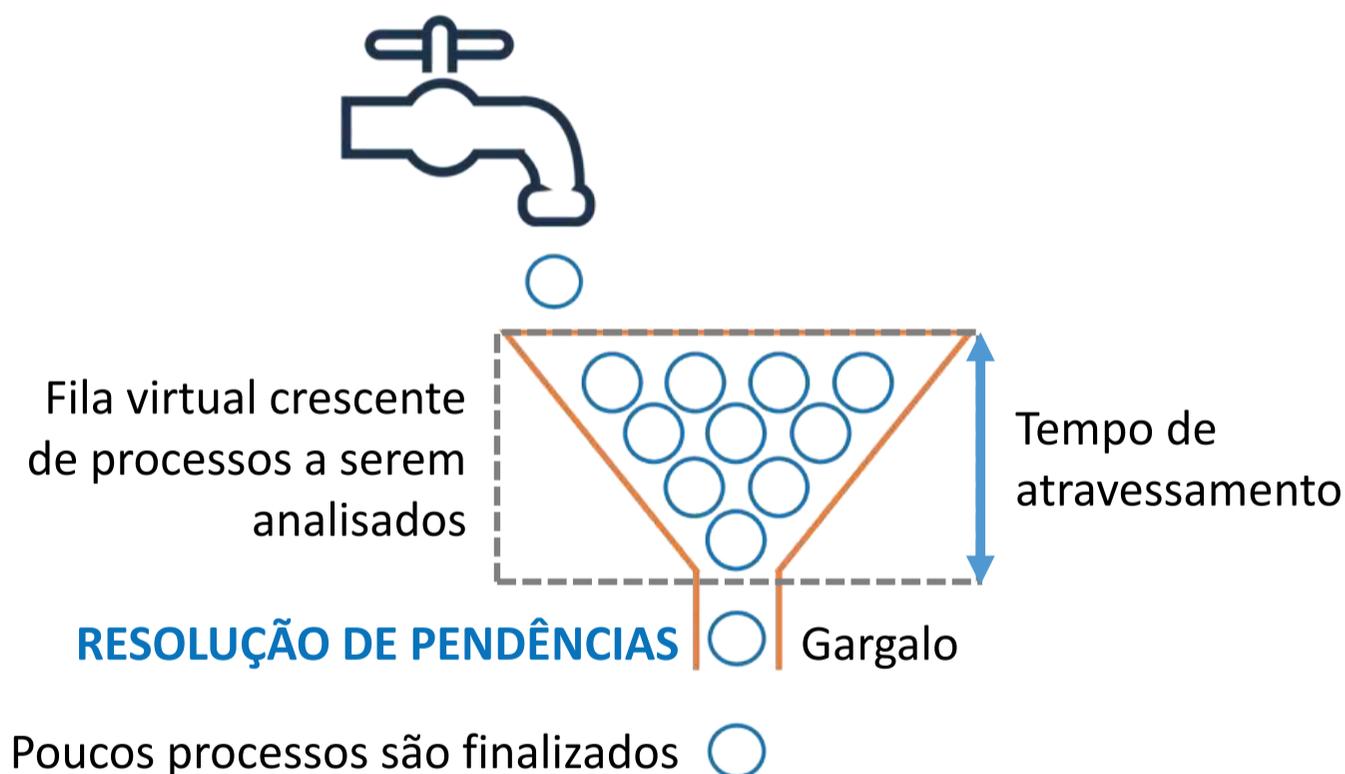


Figura 13 – Princípios dos processos aplicados nesse estudo de caso

Esse exemplo demonstra que uma transformação digital mal planejada pode ter o potencial de tornar um processo menos enxuto, isto é, menos *lean*. Com certeza, muitas evidências reais de que algo tinha piorado apareceram, como: a fila crescente de processos a serem analisados, o fato de os analistas finalizarem menos análises e o fato de o gargalo desse processo ser sempre a resolução de pendências (Figura 13). Essa, por sinal, é uma atividade que nunca deveria existir se a entrada de dados tivesse sido realizada corretamente – e que, por

não ter sido, acabou gerando essa série de retrabalhos em cadeia. Assim, vale frisar que essa atividade, assim como outros tipos de desperdícios, não devem ser incluídos no padrão de trabalho. Devemos realizar sempre o *kaizen* (melhoria contínua) já no papel.

Soa óbvio quando um exemplo desse é apresentado de forma mastigada e didática para as pessoas, mas será que é tão fácil assim enxergar o que está acontecendo quando não se tem a visão do todo, mas somente de parte do processo? Afinal, a fila virtual vai crescendo continuamente, o que estimula ações do tipo bombeiro, que resolvem apenas os sintomas e criam ainda outros problemas.

Por exemplo, o objetivo desse trabalho era criar metas de análises para os analistas: se o problema que surge é que está reduzindo o número de processos finalizados, por que não criar uma meta de número de processos a serem finalizados por dia? É a famosa “gestão” ingênua, que com certeza reduz a qualidade da análise e ainda pode gerar situações antiéticas ao estimular o indeferimento de determinado processo para simplesmente se bater a meta. Assim, vale reforçar:

**A ALTA DIREÇÃO É RESPONSÁVEL
PELO COMPROMETIMENTO DE TODA
A ORGANIZAÇÃO EM SATISFAZER AS
NECESSIDADES DA UNIDADE DE FLUXO, DE
FORMA QUE A GOVERNANÇA CORPORATIVA
ESTEJA VOLTADA PARA ESSE ÚNICO PROPÓSITO.**

Agora vamos investigar a causa-raiz do problema para pensar em ações que poderiam facilmente contornar a situação, garantindo que ganhos reais de produtividade fossem alcançados nesse processo de transformação digital.

Como visto anteriormente, a baixa porcentagem de valor agregado está relacionada às esperas e ao retrabalho que surgem por não se fazer certo da primeira vez. O gargalo do processo se encontra na atividade de resolução de pendências, já que, antes dela, cria-se uma fila virtual crescente de processos aguardando ser analisados e, na sequência, verifica-se que poucos processos são finalizados. Se o processo fosse feito “certo desde a primeira vez”, essa atividade não deveria sequer existir.

Assim, esse gargalo, além de aumentar consideravelmente o tempo de atravessamento (*lead time*) do processo, cria necessidades secundárias que repre-

sentam “desperdício” de tempo para o analista. Consequentemente, a má qualidade de entrada dos dados desencadeia retrabalhos sucessivos, os quais por sua vez geram alto estoque de requerimentos aguardando análise (longa fila virtual), longo tempo de atravessamento e recursos sobrecarregados que são altamente eficientes em retrabalhar (Figura 14).

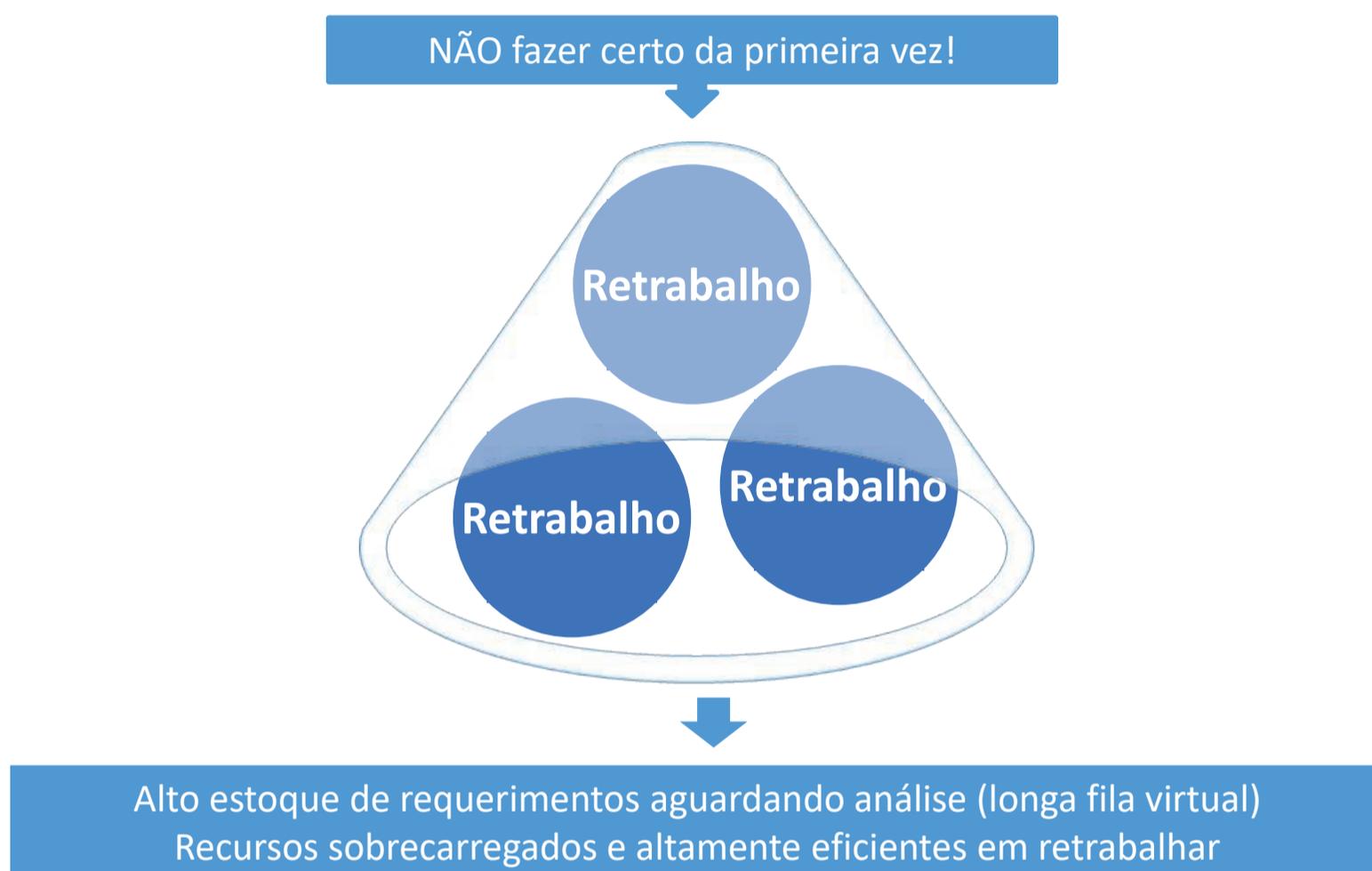


Figura 14 – Problemas gerados por não se fazer certo da primeira vez

Investigando-se a causa-raiz desse problema (não fazer certo da primeira vez) por meio da metodologia dos 5 porquês (Figura 15), percebem-se dois pontos importantes. Primeiro, muitas vezes as pessoas que realizam o atendimento não dispõem de *know-how* suficiente, uma vez que funcionários com menor experiência são alocados para essa função como estagiários, apesar de a lógica de exigência de documentação ser complexa. Assim, muitos problemas que poderiam ser previstos durante o atendimento passam despercebidos e acabam se transformando em pendências posteriores no processo.

Segundo, os agendamentos são realizados no sistema de 20 em 20 minutos, tempo já bem curto – mesmo para realizar as atividades básicas. Dessa maneira, com frequência o estagiário que faz o atendimento até sabe haver pendência e ser possível que o requerente naquele momento esteja com o documento

faltante. Entretanto, ele é cobrado apenas pela meta de não gastar mais que 20 minutos em um atendimento, porque do contrário será gerado atraso acumulativo nos atendimentos. Ou seja, logo se percebe que se trata de meta que prioriza um ótimo local, mas compromete o todo.

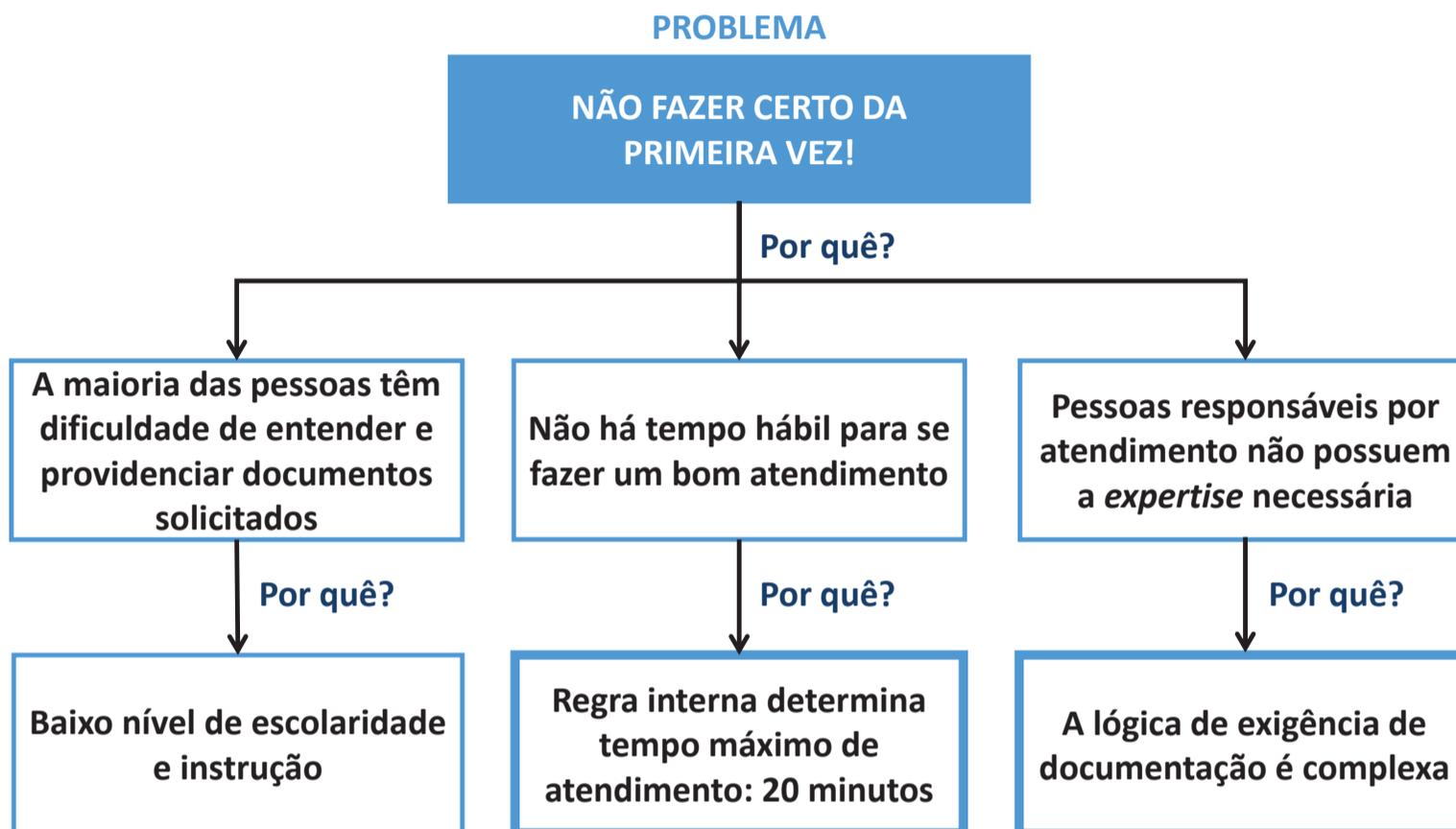


Figura 15 – Investigação via 5 porquês da causa-raiz do problema

Vale ressaltar que o baixo nível de escolaridade e instrução dos requerentes se trata de problema contextual maior. Assim, não abordaremos com profundidade essa questão nem proporemos ações de melhoria acerca do tema. Mas é importante frisar que o processo deve ser planejado com uma comunicação amigável para qualquer tipo de usuário.

Esse é um exemplo que ilustra bem o paradigma da produtividade. Um projeto mal planejado criou ilhas “eficientes”, em que as necessidades do cliente foram subdivididas em subetapas executadas por recursos com alta utilização. No caso, as ilhas seriam os analistas que estão sempre ocupados processando muitos requerimentos por dia, mas finalizando poucos. Assim, cada ilha enxerga apenas a si mesma e não tem visão do todo. O processo torna-se, pois, subotimizado com vários ótimos locais, enquanto o ótimo global é comprometido com essa abordagem, ao se criar necessidades secundárias e reduzir a porcentagem de valor agregado.

SE A GESTÃO DE DESEMPENHO PRIORIZA UM ÓTIMO LOCAL SEM OLHAR O TODO, O ÓTIMO GLOBAL É COMPROMETIDO!

E como a questão poderia ser solucionada pensando-se nessas duas causas-raiz? O problema obviamente não são os estagiários, mas sim o fato de eles estarem deixando passar processos com documentos faltantes – seja por falta de conhecimento, seja por metas que os estimulam a, por exemplo, priorizar atendimento de 20 minutos mesmo que a documentação esteja incompleta. Em relação ao tempo preestabelecido, a pergunta é simples: o que pode causar mais problemas, cumprir o tempo-padrão de 20 minutos ou extrapolar esse tempo quando necessário para evitar futuros retrabalhos? Conforme vimos acerca dos princípios dos processos, não fazer certo da primeira vez gera retrabalhos em série. Então, basicamente isso poderia ser revertido com uma reformulação no padrão operacional dos estagiários, orientando-os no sentido de que a documentação deve ser priorizada em relação ao tempo.

NÃO TEMOS TEMPO DE FAZER CERTO DA PRIMEIRA VEZ, MAS INCRIVELMENTE SEMPRE TEMOS TEMPO PARA O RETRABALHO QUE SURGE DESSA MÁ ESCOLHA.

Em relação à lógica complexa de exigência de documentação, seria possível mapear todos os documentos necessários em uma análise do processo e criar um *checklist* no sistema (Figura 16). Isso “obrigaria” os estagiários a só entrar com os dados caso toda a documentação fosse apresentada no atendimento inicial. Se faltassem documentos, os estagiários já orientariam os requerentes sobre a pendência e deixariam agendado novo atendimento, para que a entrada dos dados pudesse ser concluída. Ou seja, o atendimento inicial deve ser um filtro que só permite a entrada de processos com todos os documentos necessários.

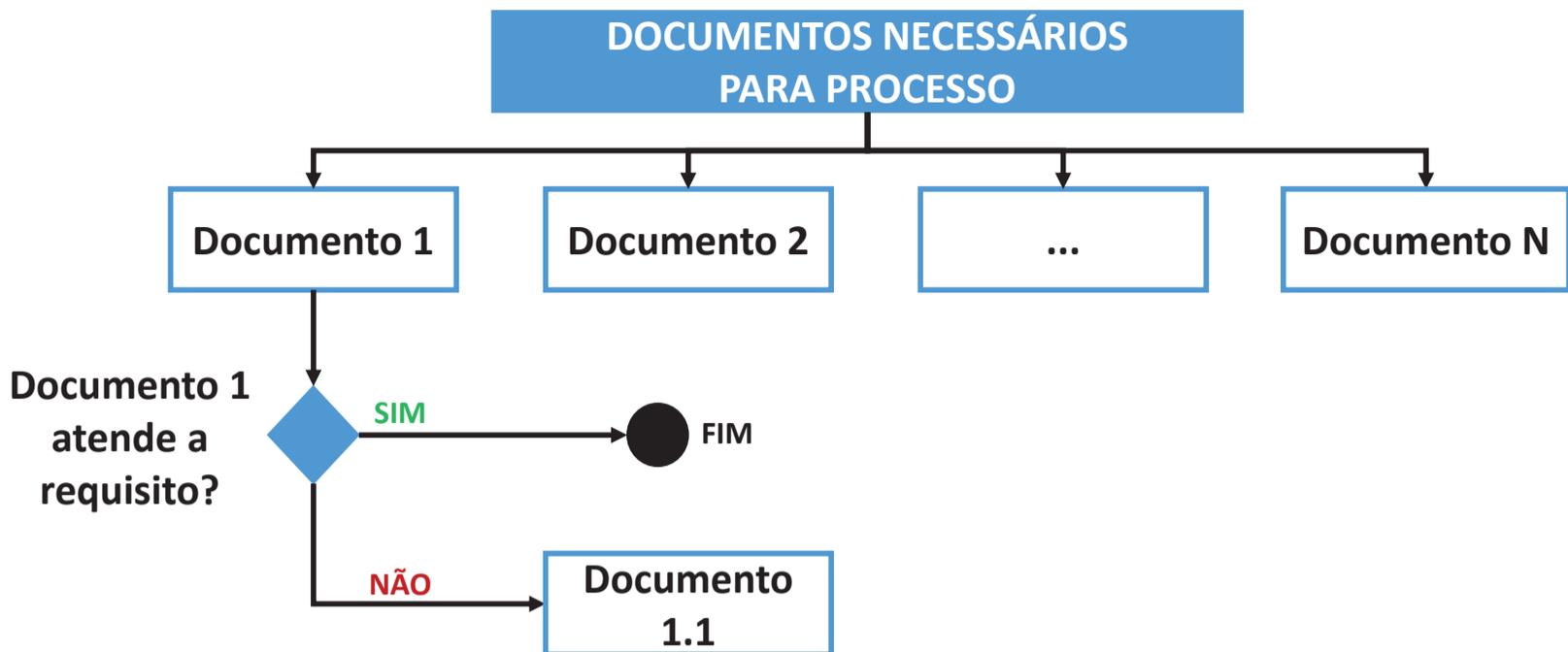


Figura 16 – Fluxograma hipotético de fatores críticos ou requisitos das entradas para um processo

Concluindo a partir desse estudo de caso, a transformação digital tem realmente o potencial de “enxugar” um processo, trazendo ganhos de produtividade. Entretanto, se mal planejada, pode acontecer como no caso em análise, em que a transformação trouxe outra mudança: a piora na qualidade dos dados de entrada. Caminhou-se, pois, no sentido contrário da famosa diretriz de se “fazer certo da primeira vez” e o processo foi inflado com desperdícios camuflados, aumentando o tempo total para que a unidade de fluxo suprisse sua necessidade.

RUMO A UM MUNDO LEAN DIGITAL

A transformação digital, assim como “ser *lean*”, é um processo gradual de melhoria. Como o nome sugere, transformar é processo contínuo, em que tecnologias digitais vão sendo incorporadas a determinado processo. Por exemplo: na sequência do estudo de caso apresentado, poderiam ser criadas ações de melhorias – o próprio solicitante realizar o *upload* dos documentos necessários para a análise. Entretanto, deve-se ter o devido cuidado para que essas ações efetivamente tragam ganhos de produtividade e tornem o processo mais enxuto. O estudo de caso apresentado é bom exemplo disso. Trata-se de transformação digital, mas que tornou o processo menos enxuto. Daí a importância de submergir além dos conceitos superficiais de *lean* e de transformação digital, para realmente entender em profundidade o que é ser *lean digital*.

A TRANSFORMAÇÃO DIGITAL, ASSIM COMO “SER LEAN”, É UM PROCESSO GRADUAL DE MELHORIA CONTÍNUA.

Agora vamos dar um passo além, já que estamos trabalhando com processos em transformação contínua, nos quais são incorporadas cada vez mais tecnologias digitais. Será apresentado, pois, como seria esse mesmo estudo de caso se o processo digital possuísse 100% de valor agregado. Conceituamos *lean digital* a partir de uma contradefinição; nada mais justo que apresentar na sequência o extremo oposto.

Os bancos de dados com informações pessoais estão cada vez mais integrados e disponíveis *on-line*. Em breve, não serão necessários analistas para estudar as informações de um solicitante e chegar à conclusão de deferimento ou indeferimento. Será desnecessário telefonar, agendar, digitalizar documentos e esperar. Na verdade, quando alguém solicitar esse serviço ao clicar em um botão de um *site*, um algoritmo será acionado a fim de acessar os dados da pessoa e em frações infinitesimais de segundo dará o veredicto (Rotina 100% *lean digital* – Figura 17). Importante ressaltar que tecnologias de reconhecimento facial poderiam ser utilizadas, por exemplo, para garantir confiabilidade ao processo e evitar fraudes.

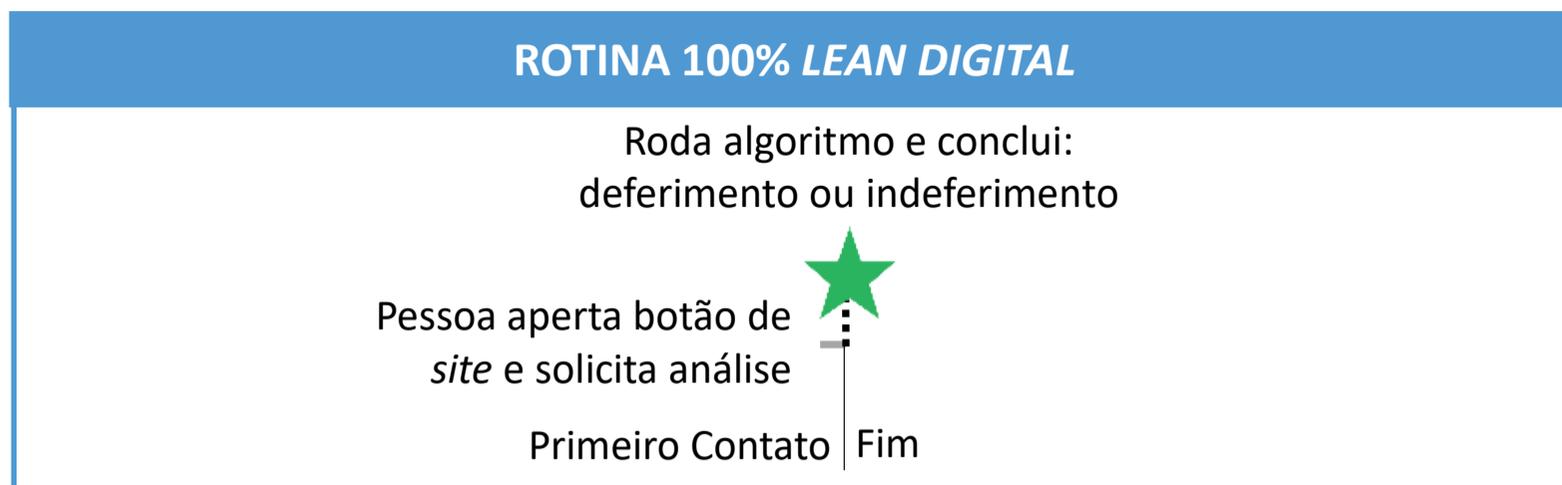


Figura 17 – Rotina 100% *lean digital*

É interessante refletir que, em um processo com 100% de valor agregado, o *lead time* pode ser em segundos, minutos, horas ou dias. No mundo físico, mesmo que enxugássemos ao máximo o processo, o *lead time* provavelmente ainda seria de alguns minutos. Já o mundo digital nos permite que as atividades que agregam valor sejam realizadas em segundos ou até frações de segundos. Esse é o mundo de possibilidades e agilidade que se abre.

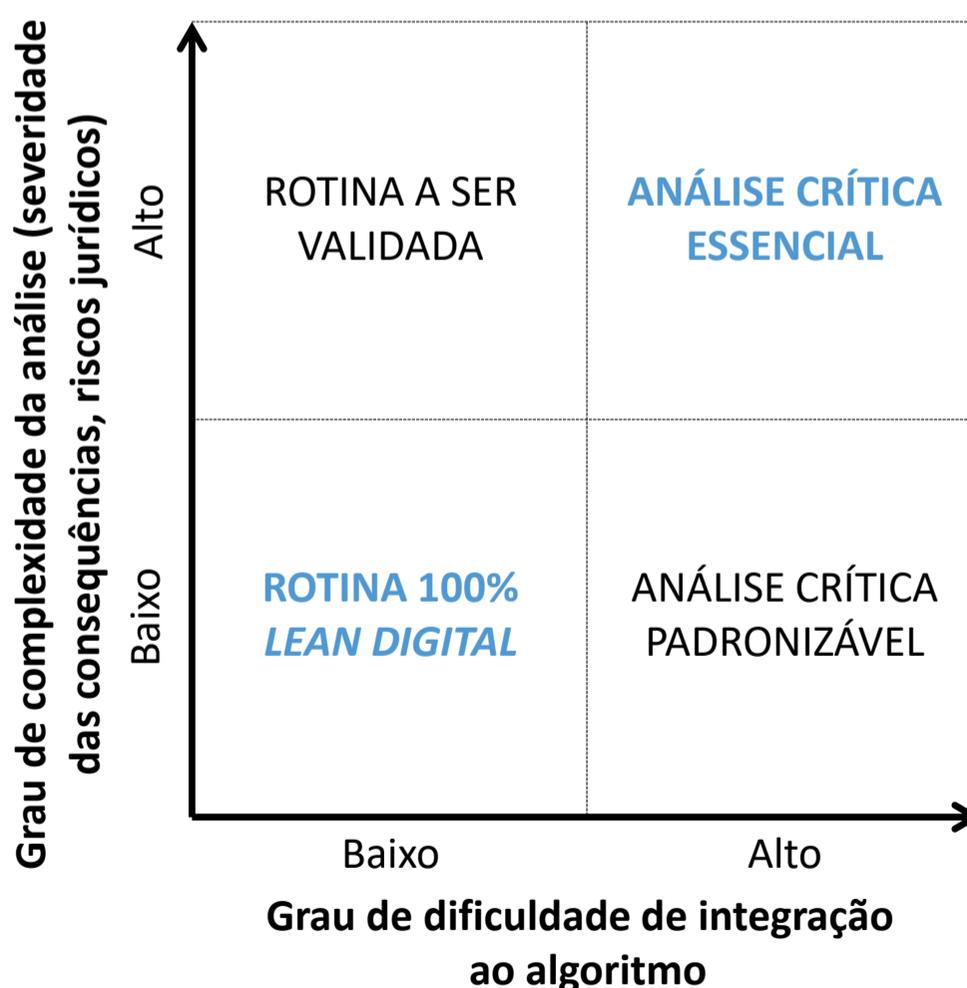


Figura 18 – Matriz para avaliar casos conforme seu nível de complexidade e o grau de dificuldade de integração ao algoritmo

Outro ponto merece reflexão: alguns casos mais complexos ainda vão demandar análise crítica por parte do analista – por terem consequências mais severas ou por envolverem maiores riscos jurídicos. Assim, conforme matriz apresentada na Figura 18, podemos avaliar os casos a serem analisados conforme seu nível de complexidade e o grau de dificuldade de integração ao algoritmo. Aqueles que apresentam consequências mais brandas e já estão previstos no algoritmo são o que apresentamos na Figura 17 como “Rotina 100% *lean digital*”. Já quando os riscos associados a algumas análises forem substanciais e se tratar de casos não previstos pelo algoritmo, é preferível que os analistas continuem responsáveis por essa (Análise crítica essencial – Figura 19).

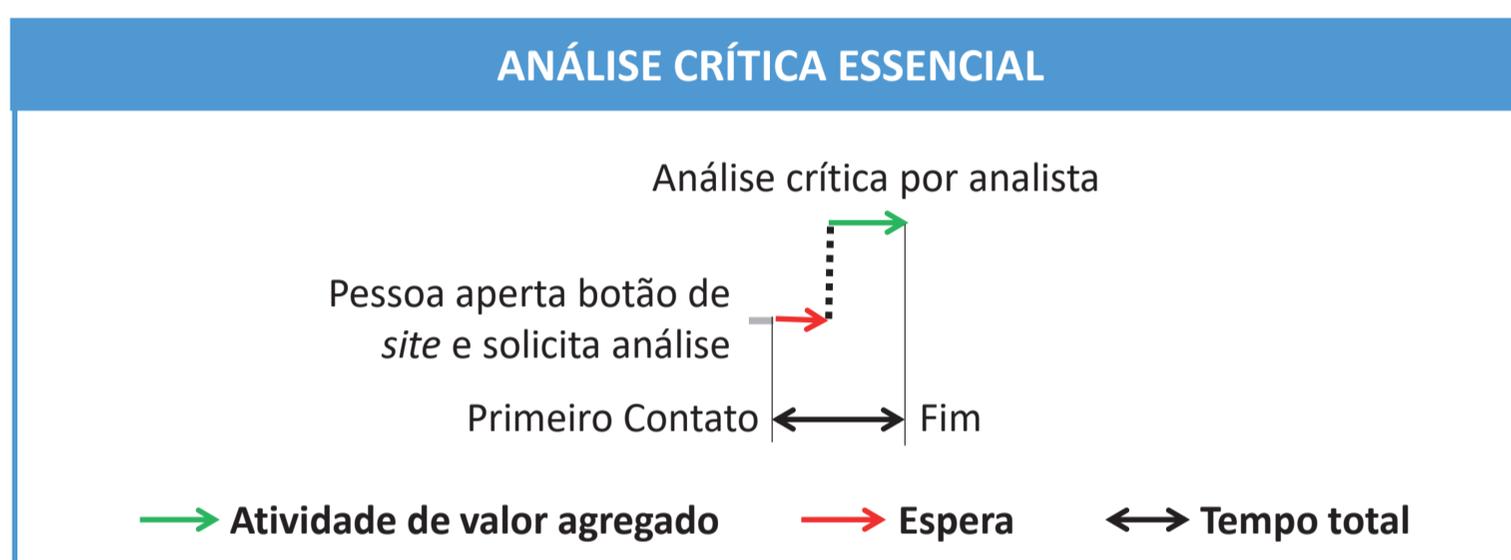


Figura 19 – Análise crítica essencial

Eventualmente pode haver casos analisáveis pelo algoritmo, mas, como envolvem consequências mais severas, poderiam ser rapidamente validados pelo analista para mitigar desgastes desnecessários (Rotina a ser validada).

Podem também surgir casos de maior complexidade, mas que poderiam ser aprendidos pelo algoritmo (Análise crítica padronizável). Assim, na primeira vez em que aparecessem, de fato teriam de ser realizados por um analista, mas na próxima poderiam ser integrados ao algoritmo existente, deixando-o cada vez mais robusto. Ou seja, uma análise crítica padronizável tende a tornar-se rotina 100% *lean digital*. Afinal, a inteligência artificial, aprendizado de máquina e tudo mais que a *Data Science* nos fornece estão aí para nos ajudar nesse processo de melhoria contínua. Então, um caso pode servir como aprendizado para outros casos que porventura surjam no futuro.

Concluindo, a evolução do trabalho acompanha as possibilidades que são criadas com a evolução da Tecnologia da Informação. Assim, com certeza surgem várias questões relativas à psicologia e sociologia do trabalho para o novo mundo digital. Embora este livro não desenvolva o assunto, vale reforçar que se trata de conhecimento muito importante e que deve ser estudado para que a transição ocorra da forma mais harmônica possível.

Outro ponto de atenção, verificado a partir desse exemplo, é a importância cada vez maior de uma educação voltada para o pensamento crítico em detrimento da habilidade de repetir e copiar. Afinal, não podemos controlar o futuro, mas, quanto mais tentarmos prever e nos adaptar a ele, maiores as chances de sermos bem-sucedidos.

Terminamos, enfim, esta jornada de transformação. Espero que agora tenham ficado mais claros os conceitos do que é ser *lean* e do que é ser *lean digital*.

DEIXO, ENFIM, O CONVITE PARA O LEITOR SE DESAFIAR A QUEBRAR ESSE PARADIGMA DA PRODUTIVIDADE: PRIMEIRO INTERNAMENTE; E, UMA VEZ QUE ESSA MUDANÇA OCORRA DENTRO DE NÓS, QUE ELA TRANSBORDE PARA TRANSFORMAR PROCESSOS, PESSOAS E ORGANIZAÇÕES EM UM MUNDO CADA VEZ MAIS DIGITAL.

REFERÊNCIAS BIBLIOGRÁFICAS

GOLDRATT, E. M.; COX, J. **A meta: um processo de melhoria contínua**. São Paulo: Nobel, 2014.

LIRA, Elisa Granha. **Estudo de tempos e movimentos: Uma abordagem lean para aumentar a eficiência de processos físicos e digitais**. Belo Horizonte: Ed. da Autora, 2020.

MODIG, N.; ÅHLSTRÖM, P. **Isto é lean: resolvendo o paradoxo da eficiência**. Estocolmo: Rheologica Publishing, 2015.

OHNO, Taiichi. **O sistema Toyota de produção: além da produção em larga escala**. Bookman, 1997.

ROTHER, M.; SHOOK, J. **Aprendendo a enxergar: mapeando o fluxo de valor para agregar valor e eliminar o desperdício**. São Paulo: Lean Institute Brasil, 1998.

ROTHER, Mike; HARRIS, Rick. **Criando fluxo contínuo**. São Paulo: Lean Institute Brasil, 2002.

SHOOK, J. et al. **Léxico lean**. São Paulo: Lean Institute Brasil, 2007.