ARCOGI

INSIGHTS

MODELO OPERACIONAL

# Reconfigurando a jornada do dado ao valor para a era agêntica

A forma como agentes de IA começam a ser usados na operação de dados, decisões e ações é prenúncio de mudanças mais amplas no modelo operacional.

por
Matias Rein
Arcogi

Setembro 2026

# Às 9h, o trabalho da noite já está na fila

Às 9h, uma responsável por produtos de dados abre sua fila de revisão. Durante a noite, agentes enriqueceram os requisitos de um novo produto de dados, verificaram contratos e classificações, executaram testes e prepararam um contexto decisório para o negócio. Cada execução carrega seus recibos: o que foi feito, por qual agente, sob qual autorização, com qual resultado.

Nada foi publicado. Nada precisou ser desfeito. As exceções estão sinalizadas, com evidências anexadas.

No meio da manhã, a equipe revisa saídas, resolve ambiguidades e refina os limites dentro dos quais os agentes operam. No fim da tarde, os próximos pacotes de trabalho estruturado entram na fila para mais um ciclo noturno.

Ninguém virou a noite. Agentes trabalharam — dentro de limites explícitos, sob autoridade humana, deixando rastro verificável de tudo o que fizeram.

Esse modelo de trabalho contínuo deixou de ser teórico. A McKinsey relata que organizações que redesenharam a entrega de software em torno de agentes já observam ganhos expressivos de produtividade com equipes substancialmente menores. A observação importante não é o número — é a causa: **os ganhos não vêm de implantar agentes, vêm de reconfigurar o modelo operacional para que pessoas e agentes colaborem 24 horas por dia.**

| 3–5× | −60% | 24h |
|---|---|---|
| ganho de produtividade observado na entrega de software redesenhada para agentes | redução média no tamanho das equipes nas implementações relatadas | de operação contínua: execução agêntica à noite, julgamento humano durante o dia |

Fonte: McKinsey Technology, "Rewiring software delivery for the agentic era", mai/2026 — referência externa; não são medições da Arcogi.

O que a entrega de software está aprendendo primeiro, a jornada do dado ao valor aprenderá em seguida. Preparar dados, qualificar produtos, compor contextos, decidir, executar ações e conciliar resultados é, estruturalmente, o mesmo tipo de trabalho: intenção humana traduzida em execução, com julgamento concentrado nos pontos em que ele é indispensável.

E produtos de dados são o caso visível, não o limite: o mesmo desenho governa decisões, políticas, execuções e a própria operação dos agentes — qualquer trabalho em que a intenção humana precise virar efeito verificável.

> *Os ganhos não vêm de implantar agentes. Vêm de reconfigurar o modelo operacional para que pessoas e agentes colaborem 24 horas por dia.*

## O outro lado do número

Os ganhos são reais — e raros. A Deloitte observa que o acesso a ferramentas de IA sancionadas pelas empresas subiu cerca de 50% em um ano, enquanto o uso no fluxo diário de trabalho permaneceu parado. A ferramenta chega; o trabalho continua igual. O piloto impressiona na demonstração de sexta-feira e é abandonado na segunda.

| | | |
|---|---|---|
| **95%** dos pilotos de IA generativa sem retorno mensurável — o que trava é a integração ao fluxo (MIT) | **>80%** dos projetos de IA falham, por estimativas amplamente citadas — a causa raiz quase nunca é técnica (RAND) | **~6%** das empresas extraem impacto real no lucro; o fator-chave é redesenhar o processo (McKinsey) |

Fontes: MIT NANDA (2025), RAND (2024 — estimativa citada, não medição própria), McKinsey State of AI (2025); dado de acesso: Deloitte State of AI in the Enterprise (2026) — leituras de pesquisa e consultoria; não são medições da Arcogi.

A causa está na planilha, não na tecnologia. As empresas destinam cerca de 93% do orçamento de IA à ferramenta e apenas ~7% às pessoas, aos processos e ao redesenho da forma de trabalhar — a "segunda linha" que transforma capacidade comprada em capacidade instalada. A linha de tecnologia é fácil de defender em comitê: tem fornecedor, cotação, contrato e data de entrega. A linha do redesenho não cabe numa cotação — e por isso fica de fora. A organização compra o que sabe comprar; e colhe pouco porque investe pouco na frente que decide o retorno.

**CASO · KLARNA**

A fintech anunciou que sua IA executava o trabalho de centenas de atendentes e economizava dezenas de milhões. Um ano depois, a própria companhia reconheceu que a decisão pesou custo demais, que a qualidade caiu — e voltou a contratar, num modelo híbrido. A lição não é que a IA não funciona. É que medir velocidade e custo médio sem redesenhar o serviço não é estratégia — é atalho.

A conclusão honesta também não é gastar menos em tecnologia: ferramenta boa é condição necessária, e o custo de acesso à capacidade nunca foi tão baixo. O erro é gastar *só* nela. Os ganhos existem e são grandes — mas ficam do outro lado de um redesenho que quase ninguém orçou.

Quatro movimentos separam a demonstração impressionante da operação confiável. Cada um depende do anterior — não há autonomia confiável sem conhecimento estruturado, nem fluxo contínuo sem handoffs legíveis por máquina.



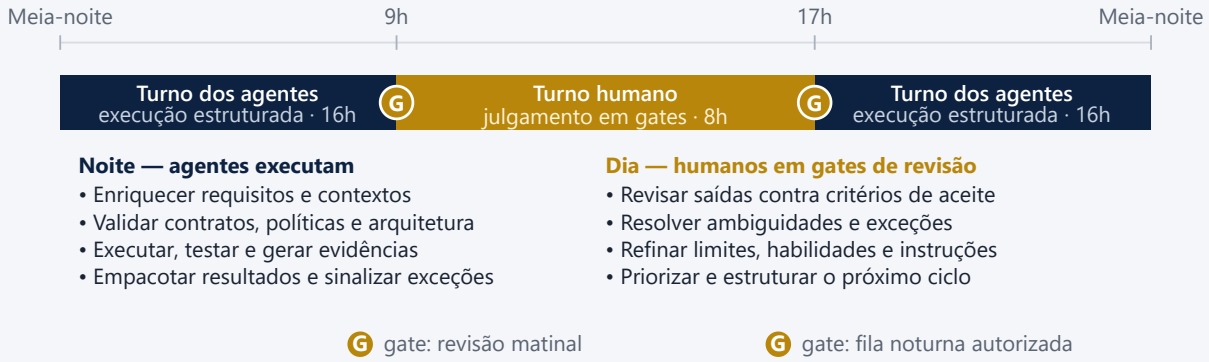

# 1 Projetar a operação para fluxo contínuo

A maioria das organizações ainda opera no ritmo das reuniões: o trabalho avança quando pessoas se encontram. No modelo agêntico, o ritmo se inverte. Durante o turno humano, a equipe revisa saídas, resolve o que exige julgamento, fortalece os limites de operação e prepara insumos estruturados. Durante o turno dos agentes, o trabalho executável acontece em escala: enriquecer requisitos, validar arquitetura e contratos, gerar e testar artefatos, empacotar resultados com evidências para revisão.

O papel humano muda de natureza: menos produzir artefatos, mais supervisionar e melhorar o sistema que os produz.

EXIBIÇÃO 1 DE 5 — O MODELO OPERACIONAL EM CADÊNCIA DIÁRIA

**O ciclo operacional de 24 horas (ilustrativo)**




Modelo de referência: McKinsey (mai/2026), adaptado à jornada dado→decisão→ação. A cadência é um padrão de desenho operacional — não uma alegação de feature específica.

Esse ciclo só funciona sobre algumas fundações práticas. O negócio precisa de uma visão clara do que deve ser produzido — um roadmap, um padrão, uma definição de pronto — para avaliar a qualidade do que os agentes geram. O ambiente precisa ser consistente, com convenções e arquiteturas modulares que permitam reuso em vez de reinvenção. O caminho da intenção à execução precisa seguir uma estrutura padronizada, para que agentes interpretem insumos e produzam saídas previsíveis. E os mesmos stakeholders precisam permanecer engajados ao longo do fluxo, sob pena de desalinhamento e retrabalho contínuos.

Sem esse nível de consistência, a produção dos agentes é fragmentada e difícil de confiar.

**MENSAGEM-CHAVE**

A operação contínua é alcançável, mas exige disciplina arquitetural e fluxos padronizados para que agentes operem de forma confiável em escala.

## 2 Estender a automação para eliminar handoffs manuais

A automação tradicional concentrou-se no fim do pipeline: testes, publicação, monitoramento. A maior parte do esforço — de requisitos a decisões — permaneceu manual e intensiva em interpretação. É aí que o atrito se acumula e o valor estagna.

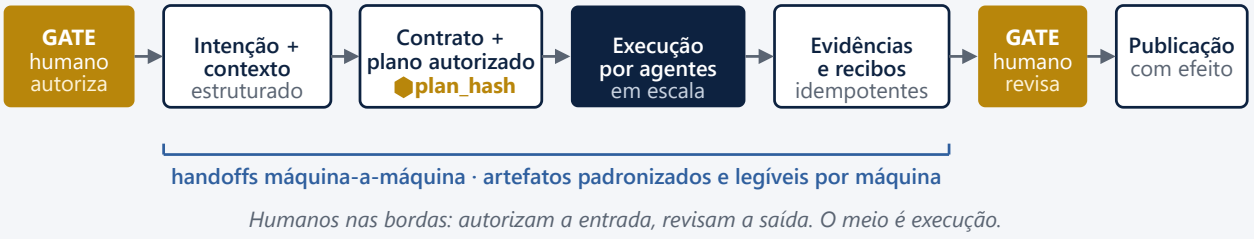
Na maioria das organizações, requisitos, padrões, políticas e especificações vivem em documentos e ferramentas desconectados. Cada transição exige que alguém traduza intenção de um artefato para outro. Cada tradução introduz ambiguidade.

O modelo agêntico remove esse atrito estruturando artefatos para handoffs máquina-a-máquina. Descrições funcionais, requisitos não funcionais, políticas, contratos e especificações passam a existir em formatos padronizados e legíveis por máquina. O pipeline roda de ponta a ponta em horas, e humanos intervêm em gates de revisão definidos — não como intermediários.

Nesse desenho, detalhes de engenharia viram mecanismos de governança. Um plano de execução pode ser amarrado à autorização que o aprovou por um hash: altere qualquer campo e a autorização anterior deixa de valer. Toda mutação relevante gera um recibo: replay não reexecuta efeitos, divergência de conteúdo é rejeitada. Publicar exige revisão por humano distinto de quem submeteu. A separação de funções deixa de ser política em um PDF e passa a ser invariante no caminho da execução.

EXIBIÇÃO 2 DE 5 — PIPELINE AGÊNTICO PONTA A PONTA

**Da intenção ao efeito, sem tradução manual**




*Humanos nas bordas: autorizam a entrada, revisam a saída. O meio é execução.*

Mecanismos com lastro na arquitetura EDAVEN: plano amarrado à autorização por hash, recibo por mutação com idempotência (a repetição não duplica efeitos), revisão por humano distinto do solicitante.

**MENSAGEM-CHAVE**

Escalar IA exige aplicar práticas de engenharia ao próprio sistema de trabalho — tornar o processo repetível e automatizar os handoffs.

## 3 Criar infraestrutura de conhecimento para desbloquear autonomia

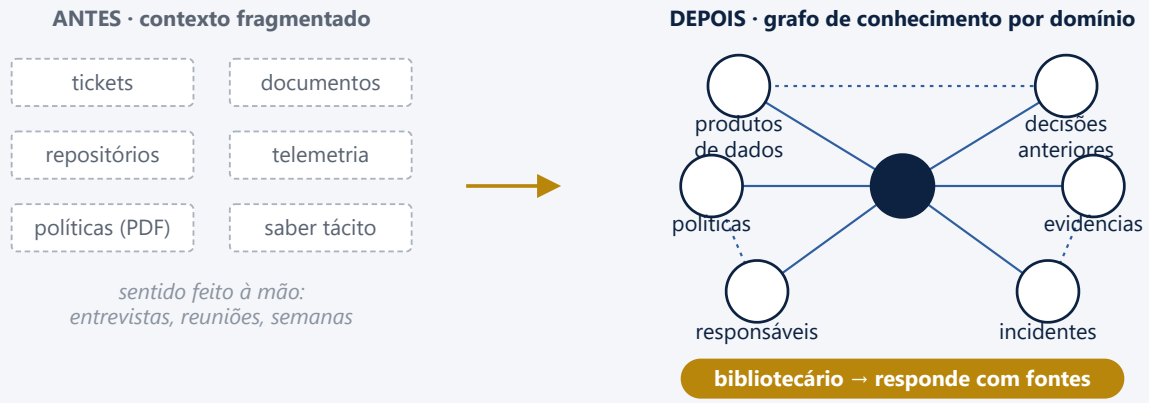
Agentes produzem resultados corretos quando têm contexto e memória organizacional. Sem isso, cada tarefa recomeça do zero — ou recomeça errado.

Hoje, o contexto de qualquer decisão relevante está espalhado: feedback de clientes em uma ferramenta, decisões de arquitetura em outra, tickets, repositórios, relatórios de incidentes, políticas resumidas e, sobretudo, o conhecimento tácito de quem está há anos na operação. Dar sentido a isso exige especialistas e longas discussões.

Organizações na fronteira estão construindo uma camada de conhecimento conectada — um grafo que relaciona produtos de dados, decisões anteriores, políticas, evidências e responsáveis dentro de cada domínio, em formato que agentes conseguem navegar. O efeito é transformador: perguntas que exigiam semanas de entrevistas com especialistas passam a ser respondidas em minutos, por um agente "bibliotecário" que consulta a memória institucional estruturada — e responde com as fontes.

EXIBIÇÃO 3 DE 5 — CAMADA DE CONHECIMENTO

**Do contexto conectado à mão ao contexto navegável por agentes (ilustrativo)**




O grafo evolui por domínio prioritário, não por grande esforço ontológico top-down. Conhecimento vira infraestrutura de produção — não documentação estática.

Toda decisão se torna rastreável. Se alguém pergunta por que uma iniciativa foi despriorizada, a resposta aponta para a fonte — pesquisa de clientes, telemetria de uso, uma política. Conhecimento implícito vira explícito e explicável, reduzindo o tempo de integração de novos membros e fortalecendo a governança.

Importante: isso não começa com um grande esforço ontológico top-down. O grafo deve evoluir organicamente em torno dos domínios prioritários e dos programas vivos, compondo valor ao longo do tempo. Em escala, conhecimento deixa de ser documentação estática e vira infraestrutura de produção — e uma fonte durável de vantagem competitiva.

**MENSAGEM-CHAVE**

Conhecimento estruturado e conectado é a fundação da autonomia dos agentes. Trate a arquitetura de conhecimento como infraestrutura estratégica.

## 4 Capturar valor: redesenhar papéis, autoridade e evidência

A produtividade do modelo agêntico só se converte em valor quando a organização muda junto — pessoas, dados, regras e agentes operando como um único sistema de trabalho. Equipes grandes e generalistas dão lugar a times pequenos e multidisciplinares, de alta qualificação, supervisionando

execução agêntica — cronogramas comprimidos, custo menor ou capacidade ampliada.

Capturar esse valor exige três movimentos. O primeiro é requalificar pessoas: enquanto um time central desenvolve e mantém a "fábrica" de agentes — padrões, conformidade, boas práticas —, profissionais em toda a organização desenvolvem julgamento, revisão e supervisão. Papéis migram de coordenação manual para coerência de arquitetura, modelagem de domínio e supervisão de IA.

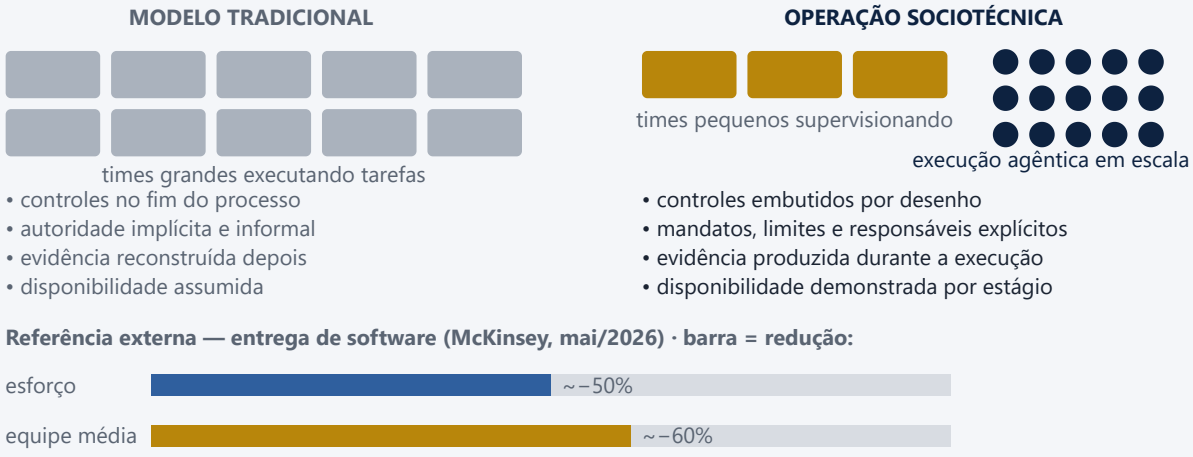
O segundo é trazer os papéis de controle externo ("outer loop") — risco, jurídico, conformidade, compras — para dentro do desenho, não como porteiros no fim da fila. Um ciclo mais rápido não acelera nada se os controles continuarem sendo gargalo. Política como código, aprovações multi-etapa e certificações embutidas no fluxo garantem que os controles sejam aplicados por desenho — com mais consistência e rastreabilidade, não menos.

O terceiro é redesenhar a alocação de capacidade: a capacidade liberada precisa ser conscientemente reinvestida — acelerar roadmaps, modernizar plataformas, lançar produtos — ou o ganho evapora em eficiência invisível.

Há ainda um quarto elemento, menos discutido e decisivo: a disciplina de verdade. Em uma operação em que agentes executam, a confiança não se declara — se demonstra. Isso significa manter, no próprio sistema, a distinção entre o que foi especificado, o que foi simulado, o que foi implementado, o que foi testado e o que está homologado para operação. Um modelo operacional que não consegue dizer, a cada momento, o que é real e o que é intenção não merece autonomia.

EXIBIÇÃO 4 DE 5 — REDESENHO DE EQUIPE E CONTROLE

**Do modelo tradicional à operação sociotécnica (ilustrativo)**




Magnitude da redução conforme McKinsey (mai/2026), entrega de software — referência externa, ilustrativa; não é projeção para operações de dados nem medição Arcogi.

**MENSAGEM-CHAVE**

Ganhos de produtividade só viram mudança estrutural quando autoridade, evidência e responsabilidade são redesenhados junto com a capacidade.

# O modelo de negócio também muda: serviço na forma de software

Existe uma consequência comercial desse desenho que raramente aparece na discussão sobre agentes. Quando o modelo operacional é codificado — práticas, contratos, gates, evidências —, o serviço que o entrega deixa de ser um projeto e passa a ser uma capacidade contínua.

Chamamos isso de *serviço na forma de software* (services as software): o serviço não é um volume de horas de especialistas; é uma prática codificada em um ambiente operacional que permanece com o cliente. O corpo de conhecimento explica como praticar; o software operacionaliza o comportamento comprovado; pessoas e agentes executam dentro dele.

A diferença prática é profunda. No modelo tradicional de consultoria, o conhecimento entra com a equipe e sai com ela: cada projeto recomeça o contexto, reconstrói a evidência e depende da disponibilidade de poucos especialistas. No serviço-como-software, a memória fica no ambiente do cliente — decisões, critérios, exceções e evidências acumulam valor a cada ciclo, e cada ciclo tende a ser mais produtivo que o anterior.

É o mesmo princípio que vale para o julgamento: ele não pode ser episódico. A capacidade de operar com dados e agentes também não pode depender de uma equipe externa que vai embora. O que se contrata não são horas de terceiros — é a operação em si, com profissionais sêniores supervisionando e agentes executando.

A entrega é modular e soberana. O centro de operações é implantável por partes no ambiente do próprio cliente — começando por organização e agentes, e somando dados prontos para uso e decisões governadas no mesmo espaço de trabalho. E a autoridade não se transfere: configurações, aprovações e mandatos permanecem sob controle da organização. O fornecedor entrega capacidade — a operação continua sendo do cliente.

A "segunda linha" do orçamento que vimos acima — a que transforma capacidade comprada em capacidade instalada, e que a recomendação de referência da BCG coloca como a maior parte do esforço — é precisamente o que o serviço na forma de software entrega: codificada, contínua e sob autoridade do cliente, em vez de recomprada como consultoria a cada ciclo.

Para o cliente, a economia vem de três lugares: menos horas cobradas, porque o executável roda em agentes; menos retrabalho, porque artefatos padronizados eliminam a tradução manual; e menos dependência, porque o conhecimento institucional permanece na própria organização — não na firma contratada. Para quem entrega, o mesmo desenho permite que equipes pequenas e sêniores atendam muito mais: a repetibilidade está no software, não na memória de quem executa.

EXIBIÇÃO 5 DE 5 MODELO DE ENTREGA

### Do projeto episódico à capacidade contínua (ilustrativo)




o conhecimento sai junto com a equipe

- custo por hora, cresce com o escopo
- evidência é o relatório no fim
- dependência do fornecedor
- continuidade episódica

a memória acumula valor a cada ciclo

- custo por operação, decresce por ciclo
- evidência produzida durante a execução
- autonomia crescente da organização
- continuidade operacional

**Economia compartilhada:**
cliente: menos horas, retrabalho e dependência · entregador: times sêniores, repetibilidade no software

Modelo de entrega em orientação (services as software) — a economia descrita é qualitativa; ganhos concretos dependem de escopo, programa e condições de cada cliente. A assimetria 93/7 (Deloitte Tech Trends, dez/2025) e a regra 10-20-70 (BCG, 2025) são leituras de consultoria, não medição auditada.

**MENSAGEM-CHAVE**

Quando o serviço é codificado em software, a capacidade fica com o cliente — e cada ciclo fica mais produtivo, mais verificável e mais econômico.

# Por onde começar

A transformação deve começar onde o impacto é maior. Na maioria das organizações, um número pequeno de programas concentra a maior parte do investimento — modernizações, reconstruções de sistemas legados, lançamentos de novos produtos de dados. Mirar essas iniciativas maximiza o impacto visível e acelera o aprendizado.

O primeiro movimento é pequeno por desenho: um domínio priorizado, um programa vivo e um responsável nomeado. Organiza-se primeiro quem trabalha — papéis, agentes, conhecimento e limites de autoridade — e os dados e as decisões entram no mesmo ambiente à medida que a confiança é demonstrada. A escala vem da repetição de ciclos verificáveis, não de uma implantação big bang. E quem ainda é pequeno tem vantagem real: menos processo cristalizado para desmontar, menos camada para atravessar — o redesenho sai mais barato quando se desenha certo desde o começo.

À medida que agentes assumem a execução em escala, os papéis humanos concentram-se em arquitetura, julgamento de produto e desenho de sistemas — e o conhecimento institucional e a coerência técnica tornam-se os diferenciais decisivos. As organizações que começarem a construir essas capacidades como parte de um esforço maior de reconfigurar o modelo operacional não vão apenas se mover mais rápido.

Vão redefinir como os dados criam valor.

> *Organizações que reconfiguram o modelo operacional não apenas se movem mais rápido — redefinem como os dados criam valor.*

---

*Este artigo descreve o modelo operacional que fundamenta a arquitetura EDAVEN — um centro de operações sociotécnico em que pessoas e agentes executam práticas do dado à decisão, com autoridade, evidência e disponibilidade demonstradas por estágio — e o modelo de entrega da Arcogi: serviço na forma de software, em que a capacidade construída permanece no ambiente do cliente.*



**SOBRE O AUTOR**

## Matias Rein

**Advisor executivo em inteligência decisória e governança sociotécnica, fundador da Arcogi.** Três décadas acompanhando a transformação de dados, analytics, nuvem, IA e agentes em ambientes corporativos — serviços financeiros, educação, indústria farmacêutica, varejo, indústria e telecomunicações — incluindo mais de uma década liderando consultoria.

**SOBRE A ARCOGI**

A Arcogi transforma essa prática em software. O EDAVEN é um centro de operações sociotécnico implantável por módulos no ambiente do cliente, no qual pessoas e agentes executam práticas do dado à decisão — com autoridade, evidência e memória acumuladas a cada ciclo.

ARCOGI

Este material descreve o modelo operacional que fundamenta a arquitetura EDAVEN. Os números de produtividade citados são da pesquisa McKinsey referenciada no texto; não são medições da Arcogi nem de clientes. Capacidades de produto mencionadas descrevem estágios progressivos de disponibilidade e não constituem promessa de resultado ou condição comercial.

Arcogi · Data to Value