

INTRODUZINDO A CULTURA DE INOVAÇÃO TECNOLÓGICA NO PODER JUDICIÁRIO:

A EXPERIÊNCIA DO IJUSPLAB

PARTE 1 – A EQUIPE DE GESTÃO DE DADOS

INTRODUCING THE CULTURE OF TECHNOLOGICAL INNOVATION IN THE JUDICIARY:

THE IJUSPLAB EXPERIENCE

PART 1 – THE DATA MANAGEMENT TEAM

*Caio Moysés de Lima**

RESUMO

Este capítulo aborda os desafios relacionados à gestão de dados na administração da Justiça Federal de São Paulo, descrevendo o processo de criação e consolidação da Equipe de Gestão de Dados do iJusPLab, sua estrutura, princípios e métodos de trabalho.

Palavras-chave: Tecnologia; Inovação; Experiência.

ABSTRACT

This chapter addresses some of the challenges related to data management in the administrative body of the Federal Courts of São Paulo, describing the process of creating and consolidating iJusPLab's Data Management Team as well as its structure, principles and work methods.

* Juiz Federal e Vice-diretor do Foro da Seção Judiciária de São Paulo. Mestre em Ética e Filosofia Política e Bacharel em Direito e Filosofia. Estudante de Ciência da Computação. Email: <cmlima@trf3.jus.br>.

- Inovação no Judiciário

Keywords: Technology; Innovation; Data Science; Data Governance; Data Management; Business Intelligence; Digital Transformation.

INTRODUÇÃO

A qualidade das decisões tomadas pelos gestores públicos depende em boa medida da qualidade da informação disponível. Para ser útil, a informação precisa ser atualizada, fidedigna e facilmente acessível. Se a informação demora a chegar, está desatualizada ou não é confiável, a decisão acaba tendo de ser adiada ou, quando urgente, precisa ser tomada com base em conhecimento insuficiente ou em suposições que podem não corresponder à realidade.

Ter boas fontes de informação é também importante para inovar¹. Qualquer nova oportunidade de criação de valor em uma organização precisa ser submetida a testes empíricos antes de ser implementada em larga escala² e muitas vezes não há outra forma de testar uma solução inovadora senão por meio da produção de métricas, o que também requer informação de qualidade.

A Justiça Federal de São Paulo produz quantidades enormes de dados todos os dias. A cada ato praticado por seus servidores e juízes, criam-se dados novos, valiosos, que poderiam ser preservados e convertidos em informação útil para os órgãos administrativos. Infelizmente, no entanto, esses dados são muitas vezes descartados e esquecidos ou ficam estacionados nas mãos de quem os produziu, aguardando que alguém os solicite.

Ainda hoje na Justiça Federal de São Paulo algumas informações precisam ser solicitadas por correio eletrônico e compiladas manualmente, num processo lento que se repete a cada novo pedido. Essas solicitações frequentemente precisam

¹ Segundo ROGERS (2017, p. 3665-3667), “Podemos definir inovação como qualquer mudança no produto, serviço ou processo de um negócio que agrega valor. Essa mudança pode variar desde uma melhoria incremental até a criação de algo totalmente novo e sem precedentes”.

² Essa técnica, bastante utilizada na metodologia *lean*, consiste em aplicar o método científico a modelos de negócio: “Um experimento verdadeiro segue o método científico. Começa com uma hipótese clara, que prognostica o que pode acontecer. Em seguida, testam-se tais prognósticos de forma empírica. Da mesma forma que a experimentação científica é permeada pela teoria, a experimentação da *startup* é orientada pela visão da *startup*. O objetivo de todo experimento associado à *startup* é descobrir como desenvolver um negócio sustentável em torno daquela visão”. (RIES, 2011, p. 1037-1047)

percorrer toda a hierarquia administrativa até encontrarem a pessoa detentora dos dados e esses dados precisam depois seguir todo o caminho inverso até chegarem às mãos de quem os solicitou. É um procedimento lento, oneroso e suscetível aos mais diversos tipos de erro. Oneroso, porque torna necessário envolver muitas pessoas para atender a um único pedido, consumindo tempo que poderia ser despendido em outras atividades. Lento, porque depende da transmissão do pedido de pessoa para pessoa e da consolidação manual dos dados. Suscetível a erros, porque quanto mais pessoas e etapas manuais envolvidas no processo de solicitação e produção da informação maior a chance de distorções e ruídos de comunicação.

Para ajudar a Justiça Federal de São Paulo a superar essas dificuldades e disseminar por toda a instituição uma cultura de gestão de dados, foi criada, no âmbito do iJuspLab, uma equipe interdepartamental para funcionar como uma espécie de órgão consultivo e desenvolvedor de projetos nessa área. Os integrantes da equipe recebem capacitação específica e atuam sem prejuízo de suas atribuições normais. Seu objetivo é ajudar a Administração a encontrar meios mais eficientes de coletar, armazenar e disponibilizar os dados produzidos pela instituição como um todo e de convertê-los em informação útil para os gestores.

Este texto conta a história da Equipe de Gestão de Dados do iJuspLab, explicando qual o seu propósito, como se deu a seleção de seus integrantes, quais seus métodos, estratégias, ferramentas e práticas de trabalho, a fim de que a nossa experiência possa orientar ou inspirar outros órgãos que pretendam seguir um caminho semelhante.

FASE DE DIAGNÓSTICO

A causa principal da falta de informações ou da dificuldade de obter informações na Justiça Federal não é o atraso tecnológico, mas a ausência de uma cultura organizacional de gestão de dados. Com uma simples planilha de cálculo pode-se fazer muita coisa quando as metodologias e os fluxos de trabalho são adequados.

Apesar disso, a tendência mais comum é pensar que somente um sistema informatizado é capaz de resolver esse tipo de problema. Raramente se leva em conta que a criação de um sistema informatizado é um empreendimento caro, mesmo com o uso de tecnologias *open source* e de metodologias de desenvolvimento ágil. Há um custo enorme em horas de trabalho tanto na fase de desenvolvimento

•• Inovação no Judiciário

quanto na de manutenção. Por isso, a decisão de criar um novo sistema nunca deve ser tomada de modo apressado. É preciso avaliar primeiro se não há outro modo de atingir o mesmo objetivo a um custo menor.

Adquirir um sistema existente no mercado é outra opção que exige cautela. Um sistema pronto pode trazer uma solução imediata para o problema, mas cria um vínculo de dependência entre a instituição e o proprietário do *software*. Além disso, essa estratégia pode atrasar a disseminação da cultura de gestão de dados, adiando um problema que terá de ser enfrentado algum dia.

Por questões de restrição orçamentária, na Justiça Federal de São Paulo nunca se chegou a cogitar em adquirir um sistema ou contratar uma empresa para a gestão de dados³. Desde o início decidimos seguir logo o caminho da transformação cultural e desenhar estratégias para conseguir fazer o máximo possível com o mínimo de recursos humanos e materiais.

Para nortear as nossas ações, fixamos dois grandes objetivos:

- 1) Ter um dia a possibilidade de saber o custo do processo judicial em cada unidade jurisdicional.
- 2) Ter um conjunto de *dashboards* que nos mostrasse esses dados de forma clara e intuitiva⁴.

Esses objetivos são ambiciosos e por isso não fixamos prazo para realizá-los. Preferimos tratá-los como grandes ideais para o futuro, em torno dos quais poderíamos organizar e planejar melhor as nossas pequenas ações do dia a dia. Isso nos deu um fio condutor geral e nos ajudou a ver cada novo projeto de gestão de dados como um passo no desenvolvimento daquele projeto mais ambicioso. Essa forma de estruturar as nossas ações deu sentido ao conjunto e permitiu que evoluíssemos organicamente.

³ A contratação de um profissional de fora, como explicado anteriormente, seria, em tese, viável, porque temos cargos comissionados e, portanto, disponibilidade orçamentária para esse fim. Para celebrar um contrato de aquisição de software ou de prestação de serviços, precisaríamos comprometer novos recursos orçamentários.

⁴ Quando a Luciana Ortiz me chamou para participar da administração e contribuir especialmente na área de gestão de dados, ela me pediu um *dashboard* por meio do qual pudéssemos acessar com apenas um clique no mapa do Estado de São Paulo todas as informações relevantes sobre cada um dos fóruns da Justiça Federal. Esse maravilhoso *dashboard* vive desde então nos nossos sonhos e serve de símbolo e inspiração para todos os nossos projetos em gestão de dados.

Paralelamente a isso, realizamos um diagnóstico global dos problemas de criação e circulação de informações na Justiça Federal de São Paulo. Ao longo de três meses, foram feitas reuniões com todos os servidores da Administração Central, um total de cerca de 400 pessoas. Nessas reuniões, ouvimos as queixas, anotamos as demandas por “sistemas” e começamos a explicar o significado e a importância da gestão de dados. Com isso demos também início à circulação das novas ideias entre os servidores.

Em seguida, compilamos todas as queixas e demandas e procuramos agrupá-las por afinidade. Esse mapeamento tornou possível separar três tipos de problemas:

- 1) Os que podiam ser resolvidos sem necessidade de novos sistemas ou de alteração dos sistemas atuais.
- 2) Os que podiam ser resolvidos por meio de alterações simples nos sistemas existentes.
- 3) Os que demandariam a criação de novos sistemas ou de funcionalidades completamente novas nos sistemas existentes.

Felizmente, a maioria dos problemas se concentravam nos grupos (1) e (2). Identificamos a necessidade de criação de apenas dois novos sistemas (um sistema de administração predial e outro de gestão de contratos e licitações) e de uma nova funcionalidade no sistema de RH (módulo para organização e atendimento dos pedidos de movimentação de servidores).

As demandas por alterações nos sistemas existentes foram encaminhadas ao nosso departamento de TI (na Justiça Federal de São Paulo, esse órgão está vinculado ao Tribunal) e o desenvolvimento dos dois novos sistemas foi assumido pela Incubadora de Soluções Tecnológicas do iJuspLab.

Com isso, estava traçada a estratégia principal para coletar, armazenar e disponibilizar a maior parte dos dados produzidos na esfera administrativa da Justiça Federal de São Paulo. Faltava ainda selecionar e capacitar um grupo de pessoas que pudesse entender como gerenciar esses dados de forma integrada aos fluxos de trabalho, consolidá-los num repositório adequado (um *data warehouse* ou armazém de dados) e convertê-los em informação útil para os gestores.

FORMAÇÃO DA EQUIPE DE GESTÃO DE DADOS

Inicialmente, pensamos em trazer um especialista de fora para conduzir a gestão de dados da Justiça Federal de São Paulo, mediante a oferta de um dos poucos

•• Inovação no Judiciário

cargos comissionados que tínhamos à disposição, mas essa alternativa não recebeu apoio institucional.

Assim, decidimos criar uma equipe interna e capacitá-la. Havia, contudo, um problema: o nosso quadro de pessoal, em constante redução e sem perspectiva de reposição, não nos permitiria criar uma equipe dedicada à nova atividade. Os que aceitassem integrar a equipe teriam de acumular as novas atribuições com o trabalho que desempenhavam, sem nada receber a mais por isso. Portanto, precisaríamos de voluntários.

Para nossa sorte, muitos servidores se ofereceram. As motivações principais eram as seguintes: (i) oportunidade de obter uma qualificação especial, custeada pela Administração; (ii) oportunidade de participar de projetos de grande relevância, aumentando a experiência profissional; e (iii) oportunidade de começar a contribuir desde logo para minorar os efeitos da progressiva redução do quadro de servidores, problema que teriam de enfrentar de qualquer modo no futuro. Compreenderam que a gestão de dados poderia reduzir a burocracia e, por conseguinte, a carga de trabalho.

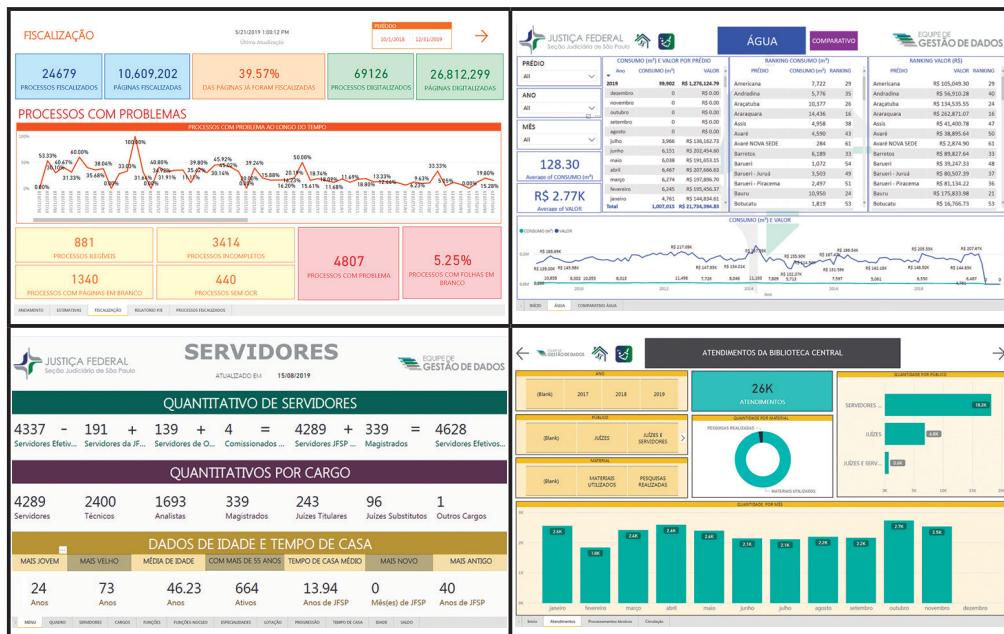
A composição da equipe sofreu inicialmente alguma oscilação, mas o grupo rapidamente se estabilizou em 15 pessoas. Logo percebemos as vantagens de reunir uma equipe interdepartamental de voluntários. Havia muita iniciativa e empenho por parte de todos os membros e a equipe apresentava, ainda, uma saudável diversidade de visões, aliada à vivência concreta dos problemas da Administração.

Essas características do grupo tornaram possível que os primeiros projetos fossem concluídos rapidamente e tivessem um alto nível de qualidade. Logo após o primeiro ciclo de capacitação, que envolveu nivelamento em Excel e treinamento em ferramentas de *business intelligence*, a equipe foi capaz de iniciar mais de 20 projetos em gestão de dados e finalizar a maior parte deles, incluindo sete conjuntos de relatórios gerenciais, cada um composto de diversos *dashboards*, num período de apenas três meses.

Até a data em que escrevo este texto (abril de 2019), tínhamos disponíveis no portal de relatórios gerenciais da Justiça Federal de São Paulo os seguintes *dashboards*: (i) histórico das despesas de água de toda a Seção Judiciária de São Paulo nos últimos dez anos; (ii) histórico das despesas de energia elétrica de toda a Seção Judiciária de São Paulo nos últimos dez anos; (iii) métricas das atividades

da biblioteca central; (iv) resumo da situação das contratações em andamento e dos contratos em vigor; (v) dados da execução do contrato de digitalização do acervo processual; (vi) perfil dos servidores em atividade, incluindo informações sobre faixa etária, lotação, sexo, especialidade, distribuição de funções e saldo total anual de ingressos e exonerações; (vii) dados sobre estoque, solicitações e valores de materiais permanentes; (viii) histórico de solicitações de assistência judiciária gratuita em todo a Seção Judiciária, incluindo as oriundas das Varas Estaduais de competência delegada, discriminadas por valor, por profissional e por tipo.

Figura 1. Alguns exemplos de dashboards criados pela equipe de gestão de dados.



Fonte: Acervo da autoria.

ORGANIZAÇÃO DA EQUIPE DE GESTÃO DE DADOS

As atribuições da Equipe de Gestão de Dados do iJuspLab foram definidas pelo art. 55-B da Portaria n. 36, de 13 de agosto de 2018, da Diretoria do Foro:

Art. 55-B. Com o objetivo de promover a cultura de gestão de dados na Administração, fica instituída a Equipe de Gestão de Dados do iNovaJusp, a ser composta por um grupo de 5 (cinco) a 15 (quinze) servidores integrantes de áreas diversas, nomeados por meio de portaria específica, com as seguintes atribuições:

•• Inovação no Judiciário

I – Pesquisar, estudar e disseminar conhecimentos em gestão de dados, podendo, inclusive, propor e realizar cursos, palestras, seminários, encontros e outras atividades de treinamento e intercâmbio de ideias.

II – Prospectar, desenvolver e propor aos demais órgãos administrativos, técnicas e soluções tecnológicas para melhorar a gestão dos dados produzidos e armazenados pela Administração.

III – Auxiliar os demais órgãos da Administração a gerir seus dados de modo eficiente, segundo os valores da inovação, dando-lhes assessoria e treinamento nessas áreas, podendo, inclusive, nesse contexto, propor alterações nos fluxos de trabalho.

IV – Prospectar, desenvolver e propor aos demais órgãos da Administração técnicas e soluções tecnológicas para integração e análise de dados, a fim de melhorar a qualidade das informações destinadas a subsidiar as decisões dos gestores.

V – Identificar dados indisponíveis e propor meios de disponibilizá-los.

VI – Padronizar a forma como são armazenados, processados e disponibilizados os dados produzidos pela Administração, inclusive no tocante à identidade visual.

VII – Criar e manter o armazém de dados (data warehouse) da Diretoria do Foro em colaboração com a Secretaria de Tecnologia da Informação – SETI do Tribunal Regional Federal da 3ª Região.

VIII – Pesquisar, experimentar, desenvolver e implantar métodos de automação das atividades previstas nos incisos anteriores.

§ 1º Dentre os integrantes da Equipe de Gestão de Dados, será designado um coordenador operacional, que ficará responsável pelas tarefas previstas no art. 46 desta portaria, e um coordenador técnico, que distribuirá e orientará as tarefas, definirá o cronograma de atividades e as prioridades de ação, aprovará os projetos propostos e selecionará as boas práticas a serem seguidas, ouvidos sempre os demais integrantes da equipe.

§ 2º A Equipe de Gestão de Dados reunirá-se sempre que necessário, mediante convocação prévia pelo coordenador operacional, que poderá também solicitar a participação de outros servidores não integrantes da equipe quando necessário para o desenvolvimento de soluções relacionadas a suas áreas de trabalho.

A equipe conta atualmente com 15 integrantes e existem mais oito servidores que manifestaram interesse em participar da equipe, os quais se encontram ainda em fase inicial de capacitação. Não houve definição de percentual de representantes de cada área da Administração, até porque seria difícil conciliar tal medida com a natureza voluntária da adesão à equipe. Talvez por conta do trabalho de divulgação prévia e das conversas que tivemos na fase de diagnóstico, conseguimos obter uma composição razoavelmente diversificada.

Uma das questões abordadas na literatura especializada sobre a composição de equipes de gestão de dados diz respeito à conveniência de incorporar técnicos

em tecnologia da informação. Alguns projetos podem tornar necessário o desenvolvimento de *scripts* e aplicativos em VBA, Python ou R, por exemplo. Além disso, por vezes pode ser útil ter algum conhecimento em lógica de programação para lidar com procedimentos ou fórmulas mais complexas em SQL, Excel, DAX ou M (Power Query). Terceirizar essa parte dos projetos ao departamento de TI não seria uma boa solução, porque diminuiria a agilidade no desenvolvimento de projetos e retiraria da equipe o controle sobre parte da lógica e da implementação. Assim, desde o início nos pareceu que seria importante integrar à equipe pessoas com alguma formação na área de TI ou pelo menos com habilidades mínimas em lógica de programação, leitura de documentação técnica, linguagens e práticas de desenvolvimento de *software*, especialmente no que se refere a documentação e versionamento.

Felizmente conseguimos atrair voluntários com formação na área de TI. Alguns deles vieram da Incubadora e tinham, além da experiência em desenvolvimento de *software*, conhecimento das regras e práticas adotadas pelo nosso departamento de TI, com o qual estavam sempre em contato.

Figura 2. Equipe de Gestão de Dados em uma de suas reuniões mensais no iJusLab.



Fonte: Acervo da autoria.

Tão ou mais importante, no entanto, foi poder contar com pessoal da área de negócios, visto que são estes os efetivos representantes “dos interesses coletivos dos

•• Inovação no Judiciário

produtores de dados e dos consumidores de informação”⁵. Aliás, o perfil ideal para os membros da Equipe de Gestão de Dados é o do servidor da área de negócios que conhece também o *stack* de ferramentas tecnológicas da equipe, um profissional capaz de dialogar bem com pessoas desses dois mundos. Nem sempre é fácil encontrar profissionais com esse perfil, mas é útil tê-lo em mente na fase de recrutamento, como forma de orientar a formação da equipe.

Os trabalhos da Equipe de Gestão de Dados do iJuspLab são coordenados por dois líderes, um coordenador operacional e outro técnico, ambos atuando em conjunto, mas preocupados com aspectos distintos do trabalho. O coordenador operacional (na verdade, a coordenadora, pois hoje o posto é ocupado por uma jovem servidora), tem as seguintes atribuições, fixadas no art. 46 da Portaria n. 36/2018, citada anteriormente:

- a) Reunir-se com os membros da equipe para organização dos trabalhos.
- b) Definir atribuições, metas e prazos.
- c) Lavrar atas de reunião e documentá-las em expediente eletrônico.
- d) Estabelecer calendário de reuniões.

O coordenador técnico (função hoje exercida por um jovem servidor, com formação em matemática⁶) é responsável por distribuir e orientar as tarefas, definir o cronograma de atividades e as prioridades de ação, aprovar os projetos propostos e selecionar as boas práticas a serem seguidas, depois de ouvidos os demais integrantes.

⁵ RÊGO, 2013, p. 1065-1067. Prossegue o autor: “Equipes de Gestão de Dados subordinadas ao departamento de TI têm sua atuação muito limitada às necessidades de tecnologia das empresas, sem muitas das vezes, considerar as reais necessidades do negócio. Equipes de Gestão de Dados estabelecidas fora do departamento de TI, preferencialmente nos níveis mais altos dentro da hierarquia da empresa, possuem uma visão de negócio muito mais abrangente e também aderente às reais necessidades das empresas, porém, de forma geral, faltam condições e conhecimento técnico para acompanhar as implementações tecnológicas, geralmente conduzidas pela área de TI. Uma terceira abordagem, mais aderente às necessidades atuais do mercado e alinhadas com os princípios da Gestão de Dados, sugere a adoção de um modelo de atuação híbrido no qual a função Gestão de Dados seria compartilhada entre as áreas de negócio e TI”. Idem, p. 1434-1450.

⁶ A formação em matemática não é essencial para o exercício da coordenação técnica, mas contribui muito para a qualidade dos trabalhos, que diversos conceitos e ferramentas em ciência de dados exigem conhecimento teórico em matemática.

Percebe-se que as atribuições dos dois coordenadores coincidem na parte relativa à distribuição de tarefas, metas e prazos. Por isso, cabe a ambos definir, de forma alinhada com os órgãos diretivos, as decisões estratégicas sobre o momento de (i) acelerar a entrega de projetos; (ii) investir em capacitação; ou (iii) concentrar esforços na melhoria das ferramentas e dos fluxos de trabalho da equipe.

É bastante importante que a equipe exerça essas três atividades de forma cíclica. Concentrar esforços apenas no desenvolvimento de projetos não é uma boa estratégia. Capacitar os membros da equipe e aprimorar suas ferramentas e seus fluxos de trabalho são atividades que asseguram a qualidade dos projetos e abreviam o tempo de desenvolvimento. Todavia, isso não significa que os ciclos de capacitação e de revisão dos instrumentos de trabalho devam ser conduzidos de forma isolada da atividade de desenvolvimento de projetos. Ao contrário, a aplicação prática enriquece a capacitação e é o que dá sentido aos instrumentos de trabalho.

Nessa linha de raciocínio, uma boa estratégia de capacitação pode incluir, por exemplo, temas relacionados a projetos em andamento ou que sejam de interesse da equipe. Quanto a esse ponto, convém distinguir dois tipos de capacitação: (i) aquela destinada à formação do repertório conceitual básico, não necessariamente relacionado a uma tecnologia específica; e (ii) aquela relacionada a tecnologias específicas, sujeitas a obsolescência. Os cursos do primeiro tipo podem ter duração mais prolongada e não precisam necessariamente ser atrelados a projetos. Refletem objetivos de longo prazo. Os do segundo tipo devem ter, de preferência, duração curta e abordar projetos reais em desenvolvimento ou que estejam nos planos da equipe. É muito importante também que os cursos sejam planejados de modo a não “competir” com os projetos em andamento ou com o trabalho ordinário dos integrantes da equipe, para não causar desestímulo.

Os projetos da Equipe de Gestão de Dados do iJuspLab raramente são conduzidos pelo grupo todo. Em geral, formam-se times de uma a quatro pessoas para conduzir cada projeto, sempre com a participação de um ou mais gestores beneficiários (*stakeholders*), salvo se houver *stakeholders* entre os membros do time de desenvolvimento. O modelo inspira-se na metodologia *lean* e nos valores da inovação. Times enxutos se comunicam melhor e evitam desperdício. Asseguram, ainda, que todos tenham papel relevante no projeto e possam, desse modo, ganhar experiência e conhecimento mais rapidamente. A participação de um ou mais *stakeholders* assegura papel de destaque aos usuários, que podem, assim, contribuir ativamente para o projeto, especialmente na definição dos requisitos.

Outro aspecto importante do método de trabalho da equipe tem sido tentar designar para os projetos pessoas com algum interesse concreto na solução do problema. No início, chegamos a trabalhar com a hipótese contrária, de que seria melhor um distanciamento do desenvolvedor em relação à solução pretendida. Achávamos que isso daria mais objetividade ao trabalho e possibilitaria um novo olhar sobre os problemas. Partindo dessa hipótese, designamos para o nosso primeiro projeto, que envolvia um problema de fluxo de trabalho entre a folha de pagamento e um dos setores da área de financeira, dois membros da equipe que atuavam na Subsecretaria de Materiais, Arquivo e Gestão Documental, responsável pela manutenção dos arquivos documentais e estoques de materiais permanentes e de consumo. O trabalho não evoluiu como esperado. Os desenvolvedores tiveram enorme dificuldade para entender o fluxo de trabalho e sofreram resistência por parte do gestor da área financeira, que interpretou algumas ações da equipe, especialmente os questionamentos sobre seus fluxos de trabalho, como interferência externa. Assim, em virtude das deficiências na comunicação e da falta de colaboração, o projeto acabou suspenso, sem previsão de ser retomado.

Desde então, temos adotado por regra (i) envolver nos projetos os membros da equipe que atuem diretamente nas áreas beneficiárias ou que de algum modo tenham familiaridade com os fluxos de trabalho envolvidos; ou (ii) convidar o gestor responsável a integrar formalmente a equipe de desenvolvimento e a participar das reuniões de trabalho, prestando-lhe esclarecimentos sobre o escopo do projeto e obtendo desde logo seu consentimento expresso (ainda que informal, verbal) com eventual revisão de seus fluxos de trabalho, se necessário.

Logo pudemos notar os efeitos benéficos da adesão voluntária. Os gestores passaram a ter participação ativa nos projetos e a divulgar espontaneamente o trabalho da equipe, a tal ponto que hoje raramente precisamos buscar oportunidades de atuação, que, em geral, nós é que somos procurados com propostas de novos projetos.

Para assegurar uma boa comunicação interna, a equipe se reúne mensalmente no iJuspLab para receber orientações gerais e compartilhar informações. A pauta é definida pela coordenadora operacional e divulgada antecipadamente a todos os integrantes. O tempo de duração das reuniões é limitado a duas ou três horas no máximo, dependendo da pauta, a fim de assegurar o bom uso do tempo e permitir que todos possam conciliar a participação nas reuniões com seu trabalho ordinário. Depois de cada reunião, a coordenadora operacional redige um pró-memória para registro e divulgação dos assuntos discutidos.

Além das reuniões mensais, há também reuniões de trabalho entre os times de cada projeto e os coordenadores técnico e operacional da equipe. Essas reuniões são divulgadas em calendário compartilhado publicamente, para que qualquer outro membro que queira também possa participar, mesmo que não esteja atuando no projeto. O objetivo é discutir qualquer assunto relacionado ao desenvolvimento de um projeto, como, por exemplo, quais ferramentas utilizar, como dividir as funcionalidades, definição de prazos e metas etc. Não há pauta nem duração pré-definidas para essas reuniões. O tempo de duração é ditado exclusivamente pela natureza dos temas discutidos e pela disponibilidade dos participantes.

PRINCÍPIOS DA GESTÃO DE DADOS

Na Portaria n. 36, de 13 de agosto de 2018, da Diretoria do Foro da Seção Judiciária de São Paulo, a gestão de dados é assim definida:

Art. 55-A. A gestão de dados consiste no esforço de coletar, integrar e analisar os dados produzidos ou utilizados no desempenho das atividades administrativas e judiciais, estruturados ou não, provenientes de fontes diversas e heterogêneas, a fim de torná-los prontamente disponíveis aos órgãos gestores que deles necessitarem, com a devida preservação do sigilo legal, quando for o caso, com o objetivo de:

- I – Subsidiar a tomada de decisões gerenciais, mediante análise situacional e identificação de causas e tendências.
- II – Automatizar a apuração de métricas de desempenho e outros indicadores numéricos importantes para a tomada de decisões gerenciais; e
- III – Automatizar e aprimorar fluxos de trabalho.

Parágrafo único A gestão de dados é responsabilidade de todos os órgãos da Administração.

Merece atenção especial a regra do parágrafo único: a gestão de dados não é atribuição exclusiva de uma só equipe, mas um dever de toda a Administração; a equipe apenas fornece o apoio necessário para o cumprimento desse dever. Assim, os princípios da gestão de dados aplicam-se a todos.

Quais são esses princípios?

Inicialmente, convém distinguir entre gestão de dados e ciência de dados. A gestão de dados recorre à ciência de dados para produzir informação e conhecimento⁷, mas seu escopo é bem mais amplo, porque envolve também os problemas

⁷ A ciência de dados é uma disciplina relativamente nova, cujas fronteiras ainda não estão bem definidas. Mas existe um consenso de que o seu núcleo está na criação de modelos preditivos

•• Inovação no Judiciário

éticos, jurídicos e administrativos relacionados à produção e utilização de dados pela organização⁸. A ciência de dados nos diz o que é possível fazer e como fazer. A gestão de dados nos diz se devemos fazer e por quê. Assim, os princípios da gestão de dados se aproximam muito mais do saber prático que do saber teórico.

Princípios de natureza prática admitem, em geral, diversas formulações diferentes. Por isso, ao oferecermos uma lista de princípios de gestão de dados, nossa intenção não é excluir ou negar a existência de outros princípios. A lista a seguir deve ser entendida como uma expressão condensada de nossa própria experiência nessa área, mesclada aos conceitos que aos poucos fomos assimilando. Ela não exaure o assunto. Deve ser tratada tão somente como um guia prático.

Dada essa premissa, nos trabalhos da Equipe de Gestão de Dados temos adotado os seguintes princípios de ação: (i) dar prevalência aos fluxos de trabalho sobre a tecnologia; (ii) definir a fonte única da verdade, sem impedir que haja múltiplas interpretações ou versões da verdade; (iii) assegurar a qualidade dos dados; (iv) observar as políticas de segurança da informação vigentes na Justiça Federal.

Prevalência dos fluxos de trabalho

Na fase inicial de implantação da gestão de dados na Justiça Federal de São Paulo, quando ainda fazíamos o diagnóstico da situação, era comum que servidores nos procurassem com a ideia de um novo sistema informatizado ou aplicativo que resolveria todos os seus problemas do trabalho. Esse tipo de atitude tem sido muito alimentado pela nova onda da inteligência artificial, porque as pessoas em geral se deixam levar pela propaganda e desconhecem que alguns dos

a partir de grandes quantidades de dados, mediante o uso de técnicas de aprendizagem de máquina (*machine learning*): “The advances and successes of machine learning methods on big data and then promise of more have prompted researchers and practitioners in the industry to name this endeavor as a new branch of science and engineering. There are still discussions about what this new field of data science covers, but it seems as if the major topics are machine learning, high-performance computing, and data privacy/security” (ALPAYDIN, 2016, p. 159). Sobre o tema, ver também DHAR, 2012.

⁸ Segundo RÊGO (2013, p. 861-868), gestão de dados é a um só tempo uma disciplina de estudos e uma atividade. Como disciplina, é responsável por “definir, planejar, implantar e executar estratégias, procedimentos e práticas necessárias para gerenciar de forma efetiva os recursos de dados e informações das organizações, incluindo planos para sua definição, padronização, organização, proteção e utilização”. Como atividade, designa o “planejamento, controle e entrega dos ativos de dados e de informação”.

algoritmos de IA mais em voga são quase tão velhos quanto os primeiros computadores. É claro que estamos vivendo nesta última década um salto tecnológico sem precedentes, mas isso se deve muito mais ao aumento da capacidade de processamento e da quantidade de dados disponíveis que a uma mudança na natureza das máquinas. Há dez anos os computadores não pensavam. Hoje continuam não pensando.

Mas há algo de verdadeiro na intuição dos que procuram um “sistema” para resolver seus problemas de trabalho. A solução para todos os problemas pode não estar nos computadores, mas certamente está em algum algoritmo.

Essa afirmação pode soar paradoxal: como é possível um algoritmo sem um computador? Explico: algoritmo é todo conjunto finito de passos adequado e suficiente à resolução de um certo tipo de problema. Assim, o conceito de algoritmo é bem mais amplo que o de código de programação. Algoritmos não são apenas os polêmicos programas de computador criados pelo Google ou pelo Facebook para colher dados ou impedir a disseminação de notícias falsas. Também são algoritmos as receitas de bolo, o ato de escovar os dentes, o ato de amarrar os sapatos e *os fluxos de trabalho*. Quando alguém pede um “sistema” para resolver um problema, intuitivamente sabe que deve existir algum procedimento mecanizável capaz de fazê-lo. Todavia, para identificar esse procedimento é primeiro necessário compreender o problema e saber descrever, ainda que de modo aproximado e provisório, os passos necessários para resolvê-lo. Somente depois disso é que se torna possível escolher ou criar um algoritmo adequado, o qual pode ou não ser implementado por meio de um código de computador. Esse trabalho envolve principalmente o uso da inteligência natural e a compreensão de fluxos de trabalho pelos próprios seres humanos envolvidos.

Por isso costumamos iniciar nossos projetos dizendo aos gestores que a solução para o problema apresentado pode não ter nada a ver com tecnologia e que alguns problemas se resolvem tão somente por modificações nos fluxos de trabalho. Também temos instruído os membros da equipe de gestão de dados a juntar à documentação de cada projeto o mapeamento dos fluxos de trabalho relacionados ao problema abordado⁹, tanto na fase inicial quanto na fase da entrega, de modo a registrar as alterações eventualmente realizadas nesses fluxos.

⁹ Esse trabalho foi bastante facilitado pelo fato de que todos os fluxos de trabalho das áreas administrativas da Justiça Federal de São Paulo se encontram mapeados desde 2017.

Fonte única e múltiplas versões da verdade

Um dos princípios que costumamos utilizar bastante para transmitir a ideia de gestão de dados é o de “fonte única da verdade”. Significa que, para assegurar a integridade das informações produzidas pela organização, deve haver um único repositório para cada tipo de dado¹⁰.

Esse princípio algumas vezes é invocado para restringir a liberdade de oferecer interpretações distintas para o mesmo conjunto de dados. Essa é, contudo, uma extrapolação indevida, pois a exigência de um fonte única da verdade não elimina a possibilidade de que haja “múltiplas versões da verdade”, ou seja, que se extraiam significados diversos e até conflitantes de um mesmo conjunto de dados¹¹. Isso é possível porque os dados precisam ser interpretados e a interpretação pode variar segundo o contexto e os propósitos do destinatário da informação. Isso é perfeitamente aceitável, desde que se tenha consciência da possibilidade de múltiplas interpretações e haja transparência quanto aos pressupostos da interpretação adotada.

Por exemplo, para medir a quantidade de trabalho de um órgão judicial, pode-se adotar diversas variáveis, como volume de distribuição, tempo de tramitação, dificuldade dos processos, números de decisões e sentenças etc., dando-se ênfase a uma ou outra segundo o destinatário da informação e o contexto. Para

¹⁰ DALLEMULE e DAVENPORT (2017) definem a “fonte única da verdade” ou “single source of truth – SSOT” da seguinte forma: “The SSOT is a logical, often virtual and cloud-based repository that contains one authoritative copy of all crucial data, such as customer, supplier, and product details. It must have robust data provenance and governance controls to ensure that the data can be relied on in defensive and offensive activities, and it must use a common language—not one that is specific to a particular business unit or function”.

¹¹ “In our experience, a more flexible and realistic approach to data and information architectures involves both a single source of truth (SSOT) and multiple versions of the truth (MVOTs). The SSOT works at the data level; MVOTs support the management of information. (...) Consider how a supplier might classify its clients Bayer and Apple according to industry. At the SSOT level these companies belong, respectively, to chemicals/pharmaceuticals and consumer electronics, and all data about the supplier’s relationship with them, such as commercial transactions and market information, would be mapped accordingly. In the absence of MVOTs, the same would be true for all organizational purposes. But such broad industry classifications may be of little use to sales, for example, where a more practical version of the truth would classify Apple as a mobile phone or a laptop company, depending on which division sales was interacting with. Similarly, Bayer might be more usefully classified as a drug or a pesticide company for the purposes of competitive analysis. In short, multiple versions of the truth, derived from a common SSOT, support superior decision making”. Ibidem.

fins de medir a produtividade dos juízes, por exemplo, provavelmente será dada maior ênfase ao número de decisões e sentenças. Para medir a carga de trabalho dos servidores, o peso maior recairá sobre a quantidade de processos e o tempo de tramitação. Vista por esses dois pontos de vista, uma mesma unidade judicial pode ser considerada sobrecarregada de trabalho ou não, o que pode colocá-la em posição mais ou menos elevada em um *ranking* de prioridades, por exemplo. O que é importante observar, contudo, é que nenhuma das duas abordagens é errada ou falsa. O mesmo conjunto de dados permite esses dois tipos de abordagem. Cabe ao gestor utilizar a abordagem adequada ao contexto e ao problema que pretende solucionar.

Qualidade dos dados

Outro princípio fundamental da gestão de dados é zelar pela qualidade dos dados coletados. Isso envolve cuidados em todo o ciclo de vida dos elementos de dados, desde sua coleta até o descarte.

Segundo o *Data Management Body of Knowledge* editado pela DAMA – Data Management Association (DAMA-DMBOK)¹², a qualidade dos dados engloba os seguintes aspectos¹³:

- a. Acurácia: quão próximos estão de representar as entidades reais.
- b. Completude: se são ou não suficientes para fornecer a informação de que se necessita num dado contexto.
- c. Consistência: se apresentam integridade e coerência quando confrontados com outras fontes.
- d. Atualidade: se correspondem ao estado atual de coisas.
- e. Precisão: se refletem ou não o grau de precisão exigido em cada contexto (em casas decimais, por exemplo).
- f. Privacidade: se atendem as normas de privacidade e sigilo.
- g. Razoabilidade: se são produzidos de modo consistente com as expectativas gerenciais.

¹² <<http://www.dama.org.br/>>. Acesso em: 28 mar. 2019.

¹³ *Apud* RÊGO, 2013, p. 3775-3793. A lista foi adaptada para o livro.

•• Inovação no Judiciário

- h. Integridade referencial: se atendem aos parâmetros de integridade necessários para que sejam considerados confiáveis.
- i. Unicidade: se são gerados por uma única “fonte da verdade”.
- j. Validade: correspondem ao tipo ou formato adequado.

Segurança da Informação

Por fim, a gestão de dados deve nortear-se também pelas leis e pelas normas internas de segurança da informação¹⁴. A Política de Segurança da Informação da Justiça Federal está fixada, em seus traços gerais, pela Resolução n. 6, de 7abr. 2008¹⁵, do Conselho da Justiça Federal, que abrange “todos os Tribunais Regionais Federais, suas seções e subseções” e tem por público-alvo “agentes públicos dos órgãos participantes e ainda a estagiários, aprendizes, clientes e parceiros” (itens 2 e 3 do Anexo I).

Aplicam-se a todos os participantes os seguintes preceitos gerais, definidos no item 6 do Anexo I:

- a) “Toda e qualquer informação criada, armazenada, mantida ou descartada pelos participantes é considerada seu patrimônio e deve ser protegida conforme estabelecido na política de segurança da informação”.
- b) “Para acesso às informações dos órgãos participantes, que não sejam de domínio público, é necessário o aceite de um termo de responsabilidade, por parte dos Agentes Públicos”.

A implementação de regras específicas se dá por meio de “documentos acessórios”, classificados em três níveis distintos de abrangência:

¹⁴ No âmbito das legislações, merece destaque especial a chamada Lei Geral de Proteção de Dados Pessoais – LGPD (Lei n. 13.709/2018), que dispõe sobre “o tratamento de dados pessoais, inclusive nos meios digitais, por pessoa natural ou por pessoa jurídica de direito público ou privado, com o objetivo de proteger os direitos fundamentais de liberdade e de privacidade e o livre desenvolvimento da personalidade da pessoa natural” e se aplica, dentre outras hipóteses, à “atividade de tratamento tenha por objetivo a oferta ou o fornecimento de bens ou serviços ou o tratamento de dados de indivíduos localizados no território nacional”.

¹⁵ Links para os normativos em vigor podem ser encontrados em: <<https://www.cjf.jus.br/cjf/unidades/tecnologia-da-informacao/politica-de-seguranca-da-informacao>>. Acesso em: 27 mar. 2019.

- a) São documentos acessórios comuns, que se aplicam a todos os participantes, os que tratam do padrão para criação de documentos, da política de auditoria de segurança da informação, da política de gestão de risco, da política de segurança para aquisição, desenvolvimento e manutenção de sistemas e da metodologia de avaliação de efetividade da implementação da política de segurança.
- b) São documentos acessórios diferenciados até o nível da Região, os que definem a política de segurança de acesso físico, a política permanente de conscientização e treinamento e as penalidades.
- c) São documentos acessórios diferenciados até o nível de Seção Judiciária os que definem a política de controle de acesso lógico, a política de utilização de recursos de TI, a política de classificação de informações e o plano de continuidade de negócios.

Inserem-se no âmbito da política de utilização de recursos de TI os “meios de armazenamento de dados”. Por conseguinte, cabe à Seção Judiciária estabelecer as regras referentes aos meios de armazenamento de dados e à política de classificação de informações. Na Justiça Federal de São Paulo, até o momento em que escrevo este texto (abril de 2019), não existe ainda um normativo editado para disciplinar esses temas de forma ampla e sistemática. Existem apenas regras esparsas, como as que autorizam a abertura e utilização de repositórios em portais como GitHub, GitLab e BitBucket, para armazenamento de arquivos de projetos, respeitados os casos de sigilo (art. 55-D, parágrafo único, da Portaria n. 36, de 13 de agosto de 2018, e art. 4º da Portaria n. 49, de 20 de dezembro de 2018, ambas da Diretoria do Foro da Seção Judiciária de São Paulo)¹⁶.

¹⁶ Entendemos que sistematização desse tema é algo bastante desejável e útil para orientar as ações de uma equipe de gestão de dados. É importante saber em que meios os dados coletados podem ser armazenados, quem pode ter acesso a eles e de que forma. Como modelo podem ser seguidas as regras do item 3 do Anexo I da Portaria n. CJF-POR-2013/00437, de 28 de novembro de 2013, atualizada pela Portaria n. CJF-POR-2018/00303, de 20 de agosto de 2018, as quais prevêem inclusive a possibilidade de armazenamento e tratamento de dados em nuvem: “Disposições Gerais 3.1. Todas as informações corporativas deverão ser armazenadas nos servidores de dados corporativos do CJF, sendo que as de natureza pública poderão ser também armazenadas em nuvem, mídias e serviços tecnológicos como correio eletrônico, videoconferência e redes sociais corporativas. As informações tratadas em ambiente de computação em nuvem devem passar por um processo de Gestão de Riscos de

•• Inovação no Judiciário

Para não correr riscos desnecessários numa área complexa, repleta de regras e armadilhas, a melhor estratégia para uma equipe de gestão de dados é alinhar-se com o departamento de TI, dando-lhe ciência de suas atividades e abertura para receber *feedbacks* sobre eventuais vulnerabilidades e riscos associados a seus projetos. Importante, ainda, que a equipe de gestão de dados utilize a infraestrutura disponibilizada pelo departamento de TI, pois a criação de infraestrutura própria envolve diversos tipos de risco, especialmente: (i) riscos de segurança e de agilidade, pois, se a equipe deixar de valer-se dos mecanismos de proteção oferecidos por seu departamento de TI precisará criar e manter seus próprios mecanismos de proteção, o que consumirá tempo que poderia ser empregado no desenvolvimento de seus projetos principais; e (ii) riscos de *compliance*, visto que a equipe passará a ter responsabilidade direta por eventual descumprimento das regras de segurança e poderá não receber suporte do departamento de TI.

Na Justiça Federal de São Paulo, o alinhamento entre a Equipe de Gestão de Dados do iJuspLab e o departamento de TI se dá pelo compartilhamento do arma-

Segurança da Informação. Informação sem restrição de acesso pode ser tratada, a critério do órgão, em ambiente de computação em nuvem, considerando os riscos de Segurança da Informação. Informação classificada como sigilosa não pode ser tratada em ambiente de computação em nuvem. É vedada a transmissão, processamento ou guarda de informações corporativas do CJF em contas pessoais de serviços de computação ou armazenamento em nuvem. 3.2. Os arquivos pessoais deverão ser armazenados nas estações de trabalho ou em nuvem. 3.3. A Secretaria de Tecnologia da Informação – STI proverá os mecanismos necessários para a proteção das informações armazenadas somente nos servidores de rede corporativos do CJF, visando garantir a integridade, a disponibilidade e a confidencialidade das informações e conceder os privilégios de acesso adequados aos perfis dos usuários. 3.4. A STI deverá efetuar backup periódico dos sistemas e das informações armazenadas nos servidores de rede corporativos do CJF, com ciclos periódicos de backup e tempos de retenção em conformidade com a política de temporalidade definida pelo órgão. 3.5. A STI deverá divulgar às unidades do CJF os tempos de retenção de backup e a granularidade de restauração da informação configurados em seus sistemas de informação. 3.6. A salvaguarda das informações armazenadas nas estações de trabalho ficará a cargo do usuário do recurso, que deverá responsabilizar-se por esses dados e pela realização de backup antes das manutenções do equipamento. 3.7. A STI deverá implementar mecanismos que reduzam a alocação de recursos devido à duplicação de dados nos sistemas de armazenamento. 3.8. É vedado o compartilhamento permanente de pastas nas estações de trabalho dos usuários. Caso haja a necessidade de nova área de trabalho para armazenamento de dados corporativos, esta deverá ser requerida ao SAC pelo titular da unidade. 3.9. A STI deverá prover mecanismos para o descarte seguro de informação armazenada em meio digital, de forma a preservar a confidencialidade dos dados do CJF”.

zém de dados (*data warehouse*) e pela adoção de todas as boas práticas da instituição no tocante aos bancos de dados utilizados, à modelagem das tabelas (especialmente nomenclatura e documentação) e aos requisitos de aprovação dos modelos criados. A equipe tem interlocução constante com o departamento de TI e há inclusive projetos em comum.

É importante ressaltar, ainda, que, em casos excepcionais, pode vir a ser conveniente ou até mesmo necessário que a equipe mantenha repositórios próprios, separados da infraestrutura fornecida pelo departamento de TI. Um exemplo são os repositórios abertos em portais na Internet para armazenamento de documentos de projetos. Essa é uma medida que constitui padrão de indústria em projetos colaborativos e que facilita o desenvolvimento dos projetos numa época em que arquivos são criados, editados, visualizados e compartilhados em nuvem. Outro exemplo é o da criação de um *data lake*, ou seja, um repositório que tem por finalidade armazenar dados estruturados e não estruturados em enormes quantidades. Como é difícil dimensionar de antemão o volume que será armazenado, a contratação de serviço escalável em nuvem pode resultar em solução muito mais econômica do que adquirir ou provisionar o espaço necessário na infraestrutura interna.

O uso desse tipo de infraestrutura traz riscos específicos, que devem ser medidos e avaliados em conjunto com o departamento de TI, mas não implicam, por si só, violação da segurança de dados. Ao contrário, em razão da alta qualificação das equipes que integram os grandes provedores de serviços em nuvem e dos ganhos de escala, frequentemente os serviços em nuvem oferecem proteção mais robusta do que aquela existente em infraestrutura própria¹⁷.

Por fim, é importante ao menos mencionar alguns riscos para a segurança da informação que surgirão com o emprego de técnicas de inteligência artificial ao tratamento de dados: (i) o risco para a transparência, visto que alguns algoritmos de IA são opacos em relação a seu funcionamento interno; e (ii) o risco para a

¹⁷ Um excelente estudo sobre o tema do uso de serviços em nuvem pelo poder público pode ser encontrado no Acórdão n. 1.739/2015 do Tribunal de Contas da União, que pode ser obtido no endereço <https://pesquisa.apps.tcu.gov.br/#/documento/acordao-completo/*/NUMACORDAO%253A1739%2520ANOACORDAO%253A2015/DTRELEVANCIA%20desc,%20NUMACORDAOINT%20desc/0/%20?uuid=7f77e370-5ad8-11e9-ae50-7b8f165af2f8>. Acesso em: 28 mar. 2019.

•• Inovação no Judiciário

isonomia e imparcialidade, tendo em vista o viés (*bias*) associado à seleção dos dados, ao treinamento do algoritmo e à seleção da “recompensa”, no caso dos algoritmos de aprendizagem por reforço (*reinforcement learning*)¹⁸⁻¹⁹.

FERRAMENTAS, TÉCNICAS E METODOLOGIAS

A Equipe de Gestão de Dados do iJusLab adota diversos elementos extraídos das metodologias e ferramentas de desenvolvimento ágil (*scrum*, *kanban*, XP etc.)²⁰ e da metodologia *lean*, baseada no TPS (*Toyota Production System*)²¹.

Para evitar desperdício de recursos materiais e humanos e acelerar as entregas, tem-se sempre como objetivo entregar o mais rapidamente possível aos gestores um produto de viabilidade mínima *Minimum Viable Product* (MVP), que apresente as funcionalidades básicas necessárias para oferecer aos gestores

¹⁸ Há no YouTube um vídeo bastante didático sobre esse tema produzido pelo Google sob o título “Machine Learning and Human Bias”. O vídeo pode ser encontrado em: <<https://www.youtube.com/watch?v=59bMh59JQDo>>. Acesso em: 28 de março de 2019. Vale a pena ler também o artigo “Bringing Precision to the AI Safety Discussion”, publicado no Google Research Blog em 21/06/2016. Link para o resumo do texto, onde se pode encontrar um outro link para o artigo científico: <<https://ai.googleblog.com/2016/06/bringing-precision-to-ai-safety.html>>. Acesso em: 28 mar. 2019.

¹⁹ Segundo explica O’NEIL (2017), “To create a model, then, we make choices about what’s important enough to include, simplifying the world into a toy version that can be easily understood and from which we can infer important facts and actions. We expect it to handle only one job and accept that it will occasionally act like a clueless machine, one with enormous blind spots. Sometimes these blind spots don’t matter. A model’s blind spots reflect the judgments and priorities of its creators. Here we see that models, despite their reputation for impartiality, reflect goals and ideology. (...) It’s something we do without a second thought. Our own values and desires influence our choices, from the data we choose to collect to the questions we ask. Models are opinions embedded in mathematics. Whether or not a model works is also a matter of opinion. After all, a key component of every model, whether formal or informal, is its definition of success. (...) In each case, we must ask not only who designed the model but also what that person or company is trying to accomplish” (p. 20-21). Ver também sobre o tema NUNES e MARQUES, 2018.

²⁰ Sobre desenvolvimento ágil, ver <<http://agilemanifesto.org/>>. Acesso em: 5 mar. 2019; e <<http://www.desenvolvimentoagil.com.br/>>. Acesso em: 5 mar. 2019.

²¹ Sobre a metodologia *lean* aplicada à inovação e ao desenvolvimento de *software*, ver RIES, 2011. Conforme explica esse autor, “Estudei a manufatura enxuta, processo que nasceu no Japão com o Sistema de Produção Toyota: um modo totalmente novo de pensar acerca da manufatura de bens físicos. Constatei que, ao aplicar ideias da manufatura enxuta aos meus desafios empresariais – com alguns ajustes e mudanças –, eu tinha o começo de um arcabouço que fazia sentido para eles. Essa linha de pensamento evoluiu até a startup enxuta: a aplicação do pensamento enxuto ao processo de inovação”. Idem, p. 209-214.

algum benefício efetivo, logo no estágio inicial do projeto. Essa estratégia é o que torna possível iniciar ciclos contínuos de *feedback*, nos quais os gestores solicitam novas funcionalidades e apontam problemas. Aos ciclos de *feedback* correspondem ciclos de entregas pequenas e constantes (*continuous delivery*)²², pelas quais o projeto ganha maturidade muito rapidamente. A principal vantagem desse modelo é evitar o desperdício, algo muito importante num contexto de equipes pequenas, com tempo limitado. A estratégia do MVP evita, por exemplo, que a equipe perca tempo no desenvolvimento de funcionalidades que à primeira vista poderiam parecer necessárias, mas que não são efetivamente adotadas pelos usuários. Permite, ainda, que erros e falhas de concepção sejam identificados e retificados mais cedo a um custo menor. Até mesmo mudanças radicais de rumo (pivotagens) tornam-se menos custosas e, portanto, mais fáceis de implementar²³.

²² A ideia de entrega contínua está ligada ao “fluxo de peça única” introduzido pelo Sistema Toyota de Produção. Apesar de ser uma ideia contra-intuitiva, está comprovado empiricamente que a entrega contínua de pequenos lotes gera mais eficiência que a execução de grandes entregas em intervalos de tempo mais espaçados. Conforme explica RIES (2011, p. 3592-3713), “O fluxo de peça única é chamado assim porque tem o tamanho de lote de uma unidade. Por que encher um envelope por vez permite que o trabalho seja feito de modo mais rápido, ainda que pareça ser mais lento? Porque nossa intuição não leva em consideração o tempo extra requerido para classificar, empilhar e deslocar as grandes pilhas de envelopes semicompletos quando fazemos de outra maneira. (...). A maior vantagem do trabalho em pequenos lotes é que os problemas de qualidade podem ser identificados muito antes. A ideia principal da implantação contínua (continuous deploy) é que, em vez de transformar a implantação em um evento especial e único dentro da vida do projeto, ela deve acontecer automaticamente e com frequência, talvez várias vezes por dia, sendo automatizada e exercitada de forma a se tornar trivial”. Sobre o conceito de entrega contínua, ver também HUMBLE e FARLEY, 2014.

²³ “Na era digital, as empresas precisam inovar de maneira radicalmente diferente, por meio da experimentação rápida e do aprendizado contínuo. Em vez de se concentrar basicamente no produto acabado, essa abordagem enfoca a identificação do problema certo e, então, no desenvolvimento, teste e aprendizado, envolvendo múltiplas soluções possíveis. Como as *startups* enxutas do Vale do Silício, ela se concentra no desenvolvimento e na repetição reiterada de protótipos de viabilidade mínima – antes, durante e até depois do lançamento”. (ROGERS, 2017, p. 3676-3682). “O ponto mais importante é que o MVP não deve ser, de modo algum, um produto completo ou acabado. A maneira mais comum de estourar os orçamentos para inovação é desenvolver demais os protótipos (por meio de processos de desenvolvimento técnico longos e dispendiosos), antes de validá-los com clientes reais. Scott Cook afirma que um MVP deve ter ‘atributos apenas suficientes para permitir *feedback* útil dos primeiros adeptos’. Idem, p. 4370-4375. “Esse é o exemplo perfeito dos objetivos de um MVP: custo mínimo + aprendizado máximo. Se o MVP for bem-sucedido, a ele se seguirão

•• Inovação no Judiciário

Não há uma fase inicial de planejamento exaustivo ou de especificação de requisitos, porque a adoção de tal fase seria incompatível com a ideia de criar MVPs e de rever fluxos de trabalho²⁴. No lugar disso, dá-se maior ênfase a histórias de usuários e à necessidade de compreender a fundo o problema apresentado, inclusive lançando mão das técnicas de *design thinking* do laboratório, quando necessário.

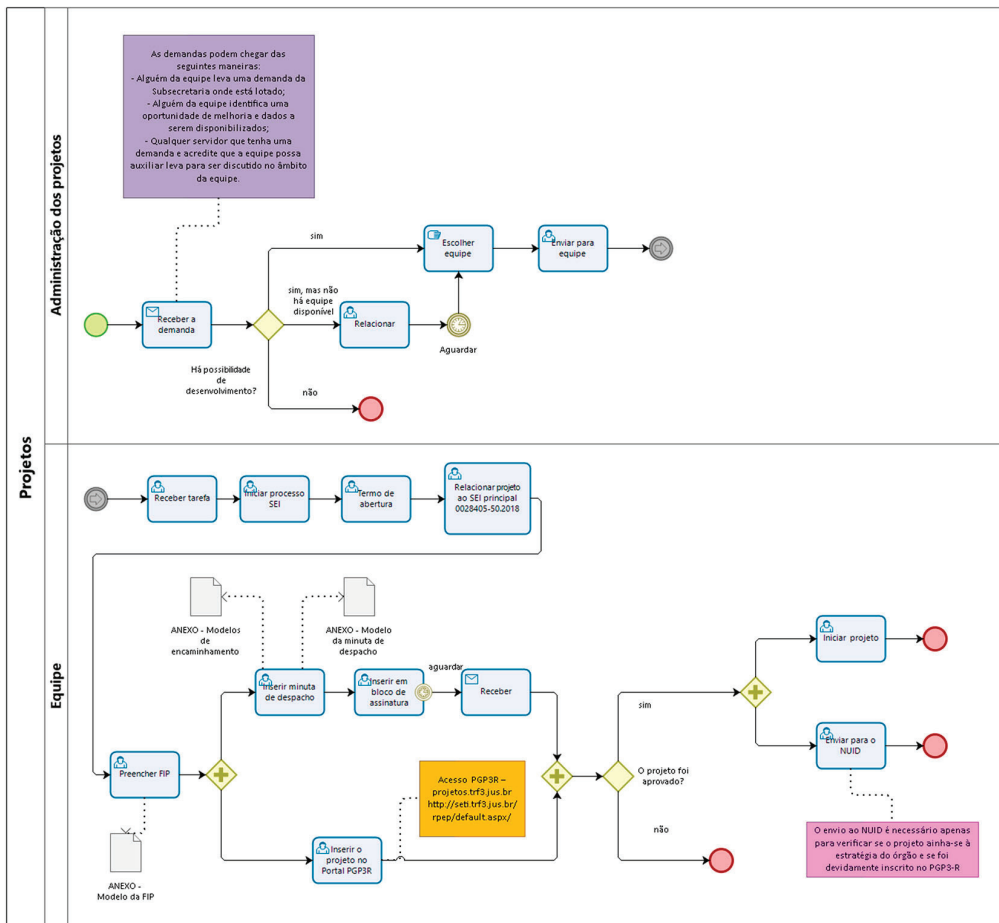
Para dar agilidade ao processo de documentação, as formalidades de criação de um novo projeto limitam-se à abertura de um expediente eletrônico, à elaboração de uma ficha inicial de projeto e à aprovação pelo órgão diretivo, de modo a assegurar que todo projeto tenha um escopo bem definido e esteja alinhado com os objetivos estratégicos da instituição (a figura 3 descreve o fluxo de abertura de projetos).

A partir daí qualquer tipo de material apto a registrar as atividades e decisões de desenvolvimento é aceito como documentação válida. Isso inclui, por exemplo, anotações manuscritas, fotos e gravações de áudio e vídeos. Também são muito utilizadas ferramentas de *kanban* que permitem converter o andamento inteiro dos trabalhos para o formato de um único documento. Os *backlogs* criados para cada etapa do desenvolvimento ficam registrados nessas ferramentas, tornando muito fácil incorporá-los depois à documentação.

outras iterações. À medida que você progride, seus sucessivos projetos devem evoluir em fidelidade (por exemplo, um esboço, um modelo, um produto funcional) e em funcionalidade, de parcial para total (por exemplo, do teste de um atributo importante para o teste da oferta completa, com todos os atributos)". Idem, p. 4379-4384. "Quando levamos em conta a construção do produto mínimo viável, essa regra simples é suficiente: elimine todo recurso, processo ou esforço que não contribui diretamente com a aprendizagem que você procura". RIES, 2011, p. 2006-2009.

²⁴ "Esse sistema é denominado desenvolvimento ágil por um bom motivo: as equipes que o empregam são capazes de mudar de direção com rapidez, permanecer ágeis, e ser altamente responsivas às mudanças nos requisitos de negócios do dono do produto". RIES, 2011, p. 2446-2450. "Essa modelagem com fases distintas de design e implementação não é ótima ao ser aplicada em software. O cliente raramente sabe de antemão tudo o que precisa antes que possa experimentar o software em uma situação real e tende a se basear em suposições; além do que as especificações, necessidades de negócio e mesmo a tecnologia disponível mudam com o tempo". Idem, p. 2512-2514. "Em vez de um desenvolvimento tradicional dividido em 'planejamento, implementação, teste, implantação' (conhecido como 'desenvolvimento em cascata' ou 'waterfall'), as metodologias ágeis trabalham com ciclos curtos de desenvolvimento, geralmente de poucas semanas, de forma que o produto é constantemente apresentado para o cliente à medida que é construído". Idem, p. 2530-2533.

Figura 3. Fluxo de abertura de novos projetos.



Fonte: Justiça Federal de Primeiro Grau em São Paulo.

Outro aspecto importante do trabalho da equipe é a definição de boas práticas uniformes para todos os projetos, a adoção de técnicas de desenvolvimento de *software* e a automatização do versionamento e dos fluxos de entrega contínua.

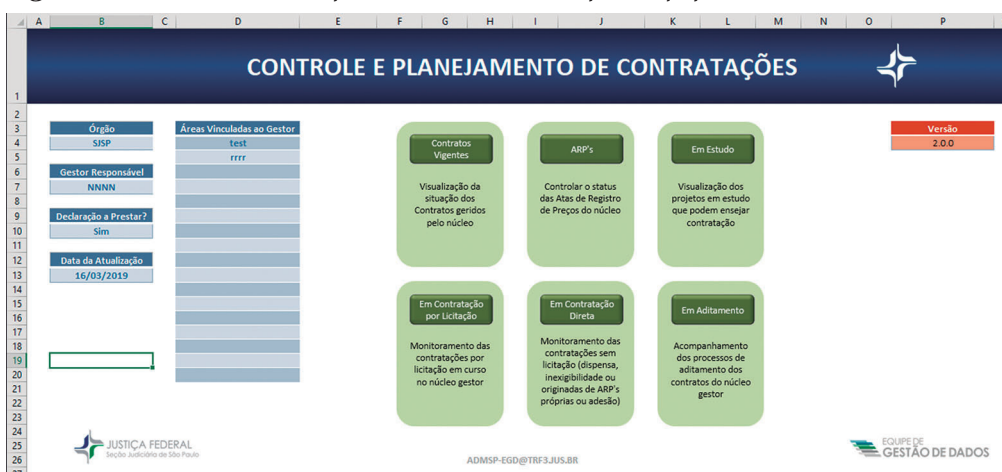
Na definição das boas práticas, são estabelecidos desde os critérios de nomenclatura até os códigos de cores para abas e intervalos das planilhas. Como exemplo, reproduzimos no Anexo II o guia de boas práticas da equipe para projetos em Excel.

Há também uma grande preocupação com a identidade visual e com o emprego de técnicas de interface gráfica (botões de navegação, alertas, instruções de preenchimento, validação não obstrutiva etc.), para assegurar uma boa experiência do usuário.

•• Inovação no Judiciário

No que se refere às técnicas de desenvolvimento, a equipe vem adotando de modo consistente alguns recursos de abstração e arquitetura modular, o que facilita a manutenção e possibilita o reaproveitamento de funcionalidades²⁵. Nas planilhas, por exemplo, há uma rígida separação entre funcionalidades de entrada de dados, de processamento de dados e de visualização de dados, as quais são até visualmente distinguidas pelo uso de um código de cores. Utiliza-se o sistema de versionamento semântico²⁶ e o controle de versões é efetuado preferencialmente com o Git. Os repositórios são mantidos na rede local ou em serviços online, como GitLab e GitHub.

Figura 4. Detalhe de uma planilha desenvolvida pela Equipe de Gestão de Dados.



Fonte: Acervo da autoria.

Para automatizar o fluxo de entrega, a equipe tem utilizado ferramentas *open source* e criado seus próprios *scripts* em Python e VBA. Algumas técnicas de automação utilizadas em desenvolvimento de *software* têm sido transpostas para o desenvolvimento de planilhas²⁷.

²⁵ Uma excelente referência para o desenvolvimento profissional em Excel, muito usada pela equipe, é o clássico *Professional Excel Development. The Definitive Guide to Developing Applications Using Microsoft Excel, VBA, and .NET.*, de BOVEY, WALLETIN, BULLEN e GREEN, citado na bibliografia. Apesar de publicado em 2009, o livro não perde a atualidade, porque aborda o desenvolvimento de planilhas Excel a partir dos conceitos fundamentais de ciência da computação.

²⁶ <<https://semver.org/lang/pt-BR/>>. Acesso em: 28 mar. 2019.

²⁷ Por exemplo, por meio de um *script* em Python e de uma extensão para Git, conseguimos automatizar o versionamento de qualquer planilha Excel. Por meio de *scripts* em VBA, encapsulados

CONCLUSÃO E VISÃO PARA O FUTURO

A Equipe de Gestão de Dados do iJuspLab realizou muito em pouco tempo. Alguns de seus integrantes, que pouco conheciam de Excel, em menos de seis meses estavam entregando seus primeiros projetos de BI. Esse resultado só foi possível em razão de cada um dos acontecimentos e personagens que compõem a pequena história que contei até aqui.

Se esses fatores pudessem ser reescritos na forma de uma receita, talvez ficassem parecidos com o seguinte:

- 1) Faça um amplo diagnóstico dos problemas e anote as queixas.
- 2) Separe os problemas que exigem desenvolvimento de *software* e encaminhe essas demandas para quem puder resolvê-las.
- 3) Enquanto isso, forme uma equipe interdepartamental de voluntários. Procure integrar nessa equipe pessoas com algum conhecimento em tecnologia e pessoas da área de negócios.
- 4) Defina objetivos ambiciosos, que possam nortear e inspirar a equipe por um longo tempo.
- 5) Organize o funcionamento da equipe. Dê especial atenção à participação dos gestores nos times de desenvolvimento. Veja também se esses times entendem dos problemas apresentados e se interessam pelo projeto.
- 6) Organize ciclos de capacitação e de aperfeiçoamento das ferramentas de trabalho.
- 7) Defina claramente os princípios de gestão de dados.
- 8) Estude e adote metodologias e ferramentas de desenvolvimento ágil.
- 9) Estude e adote os princípios da metodologia *lean*. Crie MVPs e trabalhe com ciclos contínuos de entrega e *feedback*.
- 10) Use documentação ágil. Não deixe que a equipe consuma muito tempo com trabalho burocrático.
- 11) Defina boas práticas e adote os métodos de trabalho consagrados, como, por exemplo, os utilizados no desenvolvimento de *software*.
- 12) Automatize seus fluxos de entrega.

e isolados num suplemento (*add-in*), conseguimos automatizar todas as tarefas usuais de publicação de novas versões, incluindo a limpeza e proteção de células e a ocultação e exclusão de abas.

•• Inovação no Judiciário

Como, no entanto, não há receita infalível na arte de lidar com pessoas, eu preferiria deixar duas ideias apenas, extraídas de dois textos citados na bibliografia, porque elas sintetizam, a meu ver, o essencial:

“Os aspectos de ‘infraestrutura’ da gestão de dados podem não ser tão belos e atraentes quanto os modelos preditivos e os dashboards coloridos que eles produzem, mas são vitais para um alto desempenho.”²⁸

A transformação digital não tem a ver com tecnologia – tem a ver com estratégia e novas maneiras de pensar. Transformar-se para a era digital exige que o negócio atualize sua mentalidade estratégica, muito mais que sua infraestrutura de TI.”²⁹

Esses seriam os meus conselhos para uma equipe de gestão de dados em formação. Havendo uma base sólida, fundada em trabalho de infraestrutura e mudança cultural, os modelos preditivos e os *dashboards* coloridos virão naturalmente.

À Equipe de Gestão de Dados da Justiça Federal de São Paulo, que eu tive o privilégio de ajudar a formar e conduzir, nada tenho a transmitir senão minha gratidão. Sinto que estão no caminho certo.

BIBLIOGRAFIA

ALPAYDIN, ETHEM. *Machine learning*. Cambridge (MA): MIT Press, 2016.

APRENDA sobre desenvolvimento ágil de software. Disponível em: <<http://www.desenvolvimentoagil.com.br/>>. Acesso em: 5 mar. 2019.

BOVEY, Rob; WALLETIN, Dennis; BULLEN, Stephen; GREEN, John. *Professional Excel Development*. The Definitive Guide to Developing Applications Using Microsoft Excel, VBA, and .NET. Upper Saddle River: Addison-Wesley, 2009, second edition (e-book Kindle).

DHAR, Vasant. *Data science and prediction* (maio 2012). NYU working paper n. CeDER-12-01. Disponível em: <<http://hdl.handle.net/2451/31553>>.

DALLEMULE, Leandro; DAVENPORT, Thomas H. *What’s your data strategy?* Harvard Business Review, maio/jun. 2017 (p.112-121). Disponível em: <<https://hbr.org/2017/05/whats-your-data-strategy>>. Acesso em: 5 mar. 2019.

HUMBLE, Jez; FARLEY, David. *Entrega contínua*. Como entregar software de forma rápida e confiável. Trad. Marco Aurélio Valtas Cunha, Ronaldo Melo Ferraz. Porto Alegre: Bookman, 2014 (e-book Kindle).

²⁸ DALLEMULE e DAVENPORT, 2017: “The ‘plumbing’ aspects of data management may not be as sexy as the predictive models and colorful dashboards they produce, but they’re vital to high performance”.

²⁹ ROGERS, 2017, p. 101-104.

MANIFESTO for Agile Software Development. Disponível em: <<http://agilemanifesto.org/>>. Acesso em: 5 mar. 2019.

NUNES, Dierle; MARQUES, Ana Luiza Pinto Coelho. Inteligência artificial e direito processual: vieses algorítmicos e os riscos de atribuição de função decisória às máquinas. *Revista de Processo*, v. 285/2018, p. 421-447, nov. 2018.

O'NEIL, Cathy. *Weapons of Math destruction*. How big data increases inequality and threatens democracy. New York: Broadway Books, 2017.

RÊGO, Bergson Lopes. *Gestão e governança de dados*. Promovendo dados como ativo de valor nas empresas. Rio de Janeiro: Brasport, 2013 (e-book Kindle).

RIES, Eric. *A startup enxuta*. Como os empreendedores atuais utilizam a inovação contínua para criar empresas extremamente bem-sucedidas. Trad. Texto Editores. São Paulo: Leya, 2011.

ROGERS, David L. *Transformação digital*. Repensando o seu negócio para a era digital. São Paulo: Autêntica Business, 2017 (e-book Kindle).

ANEXO I

Guia de boas práticas para projetos em Excel, vigente em abril de 2019³⁰

I – Regras de Nomenclatura

I.1 – Variáveis e constantes:

I.1.1 – Forma geral: <escopo><tipo><parte descritiva>

I.1.2 – Prefixos:

a) escopo: «**g**» (público); «**m**» (módulo); não há prefixo específico para variáveis de escopo menor.

b) tipos: «**s**» (String); «**i**» (Integer); «**l**» (Long); «**dt**» (Date); «**b**» (Boolean); «**v**» (Variant); «**c**» (Currency); «**dbl**» (Double); «**sgl**» (Single); «**by**» (Byte); «**col**» (Collection); «**ctl**» (Control); «**obj**» (Object); «**e**» (Enum); «**cls**» (instância de uma Classe); «**u**» (“User Defined Type” ou “UDT”); «**xl**» (objetos específicos do Excel); «**wd**» (objetos específicos do Word).

c) para as Arrays, incluir o prefixo «**a**» logo após o tipo. Assim, por exemplo, uma Array de Strings deve ser prefixada com «**sa**».

d) controles em formulários: «**cbo**» (ComboBox); «**chk**» (CheckBox); «**cmd**» (CommandButton); «**frm**» (Frame); «**lbl**» (Label); «**lst**» (ListBox);

³⁰ As boas práticas aqui descritas foram em boa parte inspiradas ou simplesmente extraídas do clássico *Professional Excel Development. The Definitive Guide to Developing Applications Using Microsoft Excel, VBA, and .NET.*, de Rob Bovey, Dennis Wallatin, Stephen Bullen e John Green (Upper Saddle River: Addison-Wesley, 2009, 2. ed. e-book Kindle).

•• Inovação no Judiciário

«**mpg**» (Multipage); «**pg**» (Page); «**opt**» (OptionButton); «**spn**» (SpinButton); «**txt**» (TextBox).

I.1.2 – Completar os prefixos com uma parte descritiva, o mais intuitiva possível, para facilitar a compreensão do código (auto-documentação).

I.1.2.1 – Não deve haver espaço nem underline entre o prefixo e a parte descritiva.

I.1.2.2 – A parte descritiva não deve conter acentos, espaços ou caracteres especiais. Nas variáveis, cada palavra utilizada deve iniciar por letra maiúscula, sem uso de underline (exemplo: «**sNomeDaPlanilha**»). Nas constantes, todas as letras são maiúsculas e as palavras devem ser separadas por «_» (exemplo: «**sNOME_DA_PLANILHA**»).

I.2 – Definição de UDTs: o nome da UDT segue as mesmas regras para os nomes descritivos das constantes e os itens da UDT seguem as mesmas regras para os nomes descritivos das variáveis, sendo, ainda, antecedidos por prefixos referentes ao tipo. Exemplo:

```
Type NOME_DESCRITIVO
sPrimeiroItem As String
bSegundoItem As Boolean
End Type
Enum gdCoresDasAbas
gdAzulEscuro = 0
gdVerdeClaro = 1
gdVerdeOliva = 2
End Enum
```

I.3 – Definição de Enums: os nomes da Enum e de seus itens seguem as mesmas regras para os nomes descritivos das variáveis e devem ser antecedidos de um prefixo do projeto. O nome da Enum deve expressar o sentido geral e os nomes dos itens as funções próprias de cada um. Um exemplo considerando o prefixo «**gd**» para o projeto “GestaoDados”:

I.4 – Módulos:

a) módulos comuns são precedidos de «**M**» e seguem, de resto, as mesmas regras para as partes descritivas das variáveis, com apenas uma exceção: pode ser utilizado o underline livremente sempre que necessário para organizar melhor a visualização dos módulos no IDE. Exemplo: «**MUtils**».

b) formulários são precedidos de «**F**» e seguem, de resto, as mesmas regras dos módulos comuns. Exemplo: «**FProgresso**».

c) módulos de classe são precedidos de «C» e seguem, de resto, as mesmas regras dos módulos comuns. Exemplo: «**CConfig**».

I.5 – Rotinas: formalmente, mesmas regras aplicáveis à parte descritiva das variáveis. Todavia, devem ser observadas duas particularidades: (i) as subrotinas executam ações e seu nome deve, por isso, procurar expressar a ação executada, o que pode ser obtido com a utilização de um ou mais verbos no infinitivo, seguidos ou não do objeto da ação. Exemplo: «**ProcessarDados**»; (ii) como as funções retornam valores, seu nome deve procurar refletir o resultado almejado, o que pode ser feito utilizando-se, por exemplo, o verbo no particípio. Exemplo: «**AmbienteChecado**».

I.6 – Intervalos nomeados:

I.6.1 – Prefixos: «**cht**» (*Chart Data Range*); «**con**» (intervalo utilizado para indicar uma constante); «**err**» (intervalo utilizado para checagem de erros); «**for**» (intervalo utilizado para indicar uma fórmula); «**inp**» (intervalo de input); «**out**» (intervalo de output); «**loc**» (intervalo utilizado para indicar uma localização específica); «**rgn**» (intervalo correspondente a uma região); «**ui**» (intervalo com dado de configuração da UI); «**sys**» (intervalo com dado de funcionamento do sistema); «**T**» (tabela).

I.6.2 – Intervalos criados como parte do “design” de um código em VBA devem ter seus nomes precedidos por underline. Exemplo: «conDesenvolvedores».

I.6.3 – Os prefixos utilizados como parâmetros de configuração livremente estabelecidos pelo usuário para processamento por um aplicativo em VBA devem ser precedidos e seguidos de underline. Exemplos: «limpar_», «proteger_», «ocultar_».

I.7 – Abas ou páginas: podem seguir tanto as regras definidas para a parte descritiva de variáveis quanto para a parte descritiva das constantes. Abas criadas como parte do “design” de um código em VBA devem ser precedidas por underline.

I.8 – Quando **não** aplicar as regras de nomenclatura definidas: (i) ao definir e utilizar as APIs do Windows; e (ii) ao utilizar códigos de terceiros.

II – Desenho de Planilhas

II.1 – Separar claramente as funcionalidades de (i) entrada de dados; (ii) armazenamento e recuperação de dados; (iii) processamento de dados; (iv) visualização de dados ou *UI*.

II.1.1 – Criar abas distintas para essas funcionalidades, com o seguinte código de cores:

- a) entrada de dados: azul.
- b) processamento de dados: laranja ou cinza.
- c) visualização: verde.
- d) armazenamento e recuperação de dados: preta.

II.2 – Definir e utilizar constantes e intervalos nomeados

•• Inovação no Judiciário

II.3 – Utilizar estilos e formatação condicional para tornar mais fácil a compreensão da entrada de dados e a visualização pelo usuário.

III – Desenho de Rotinas

III.1 – As rotinas no código VBA devem ser decompostas em suas funcionalidades mais elementares, de modo a que se obtenha (i) o encapsulamento das operações lógicas; (ii) a eliminação de código duplicado; (iii) o isolamento de operações complexas; (iv) a redução do tamanho das rotinas; (v) a redução do número de argumentos das rotinas.

III.2 – Deve haver sempre previsão para tratamento de erros adequado para as rotinas não triviais, inserindo-se números de linha quando necessário.

III.3 – O código deve ser explicado por meio de comentários em número e tamanho suficientes, que não sejam triviais, e que contenham o registro completo de todas as alterações posteriores.

III.4 – O código deve ser limpo e bem organizado, com uso de recuos e quebras de linha.

III.5 – Todas as regras devem ser aplicadas de modo consistente por todos os participantes da equipe.

III.6 – Usar sempre no topo de cada módulo as declarações Option Explicit e Option Private Module quando aplicável.

III.7 – Evitar o uso do tipo Variant.

III.8 – Evitar reuso de variáveis.

III.9 – Evitar a sintaxe As New para instanciar classes.

III.10 – Sempre qualificar de forma completa os objetos.

III.11 – Sempre utilizar LBound e UBound para manipular Arrays.

III.12 – Sempre especificar o contador do loop após o comando Next.

III.13 – Faça uso de constantes.

III.14 – Evite usar variáveis públicas. Procure seguir o princípio do encapsulamento.

III.15 – Use “late binding” e “early binding” de forma consciente.

III.16 – Use de modo explícito ByVal e ByRef.

III.17 – Sempre invoque de forma explícita a propriedade “default” de um objeto.

III.18 – Valide os argumentos antes de invocar uma rotina.

III.19 – Use contadores em loops While e Do While para evitar loop infinito.

III.20 – Sempre compilar o código.

III.21 – Procure não se referir às abas por seus nomes.

III.22 – Evite o uso de Select/Selection. Alterar e obter valores por meio de Range.Value.

III.23 – Procure abstrair funcionalidades e tornar declarativa a configuração elementar do código lançando mão de tabelas, intervalos nomeados, UDFs, constantes etc.