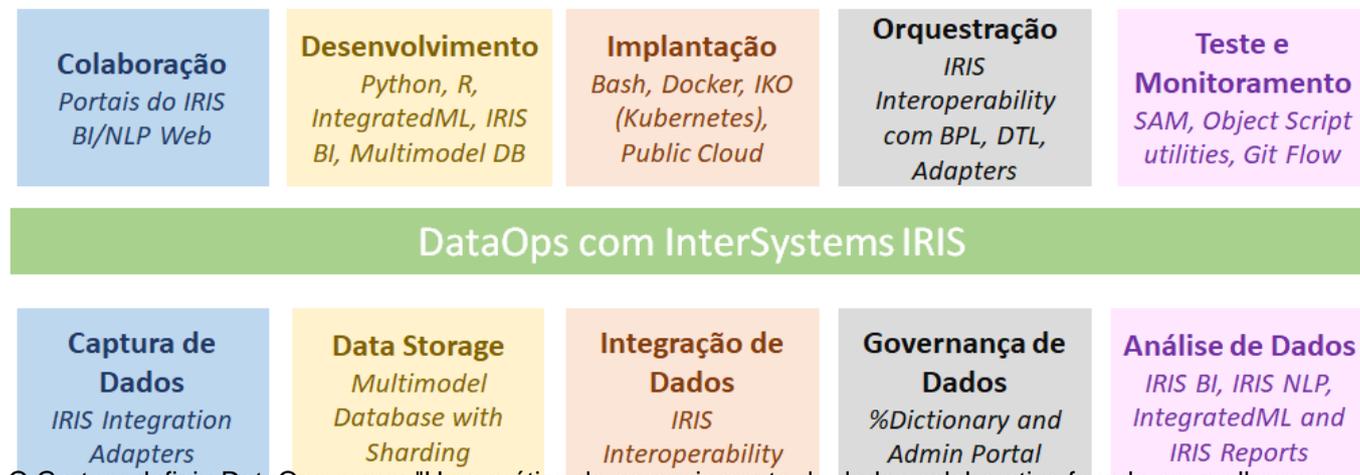


Artigo

[Yuri Marx Perei...](#) · Fev. 4, 2021 7min de leitura

## DataOps com InterSystems IRIS



O Gartner definiu DataOps como: "Uma prática de gerenciamento de dados colaborativo focada em melhorar a comunicação, integração e automação dos fluxos de dados entre os gerentes de dados e consumidores de dados em uma organização. O objetivo do DataOps é entregar valor mais rápido criando entrega previsível e gerenciamento de mudanças dados, modelos de dados e artefatos relacionados. DataOps usa tecnologia para automatizar o design, implantação e gerenciamento de entrega de dados com níveis apropriados de governança e usa metadados para melhorar a usabilidade e o valor dos dados em um ambiente dinâmico. "

O DataOps foi apresentado pela primeira vez por Lenny Liebmann, editor colaborador, InformationWeek, em uma postagem de blog no IBM Big Data & Analytics Hub intitulada "3 razões pelas quais DataOps é essencial para o sucesso de big data" em 19 de junho de 2014. O termo DataOps foi popularizado posteriormente por Andy Palmer em Tamr. DataOps é um apelido para "Operações de Dados". 2017 foi um ano significativo para DataOps com significativo desenvolvimento de ecossistema, cobertura de analistas, aumento de pesquisas de palavras-chave, pesquisas, publicações e projetos de código aberto. O Gartner nomeou DataOps no Hype Cycle for Data Management em 2018. (fonte: <https://en.wikipedia.org/wiki/DataOps>)

O manifesto DataOps estabeleceu os seguintes princípios DataOps: (<https://www.dataopsmanifesto.org/dataops-manifesto.html>)

1. Satisfaça continuamente seu cliente: Nossa maior prioridade é satisfazer o cliente por meio da entrega antecipada e contínua de valiosos insights analíticos de alguns minutos a semanas.
2. Análise de trabalho de valor: Acreditamos que a principal medida do desempenho da análise de dados é o grau em que análises criteriosas são fornecidas, incorporando dados precisos, sobre estruturas e sistemas robustos.
3. Abraça a mudança: acolhemos as necessidades em evolução dos clientes e, na verdade, as abraçamos para gerar vantagem competitiva. Acreditamos que o método mais eficiente, eficaz e ágil de comunicação com os clientes é a conversa cara a cara.
4. É um esporte de equipe: as equipes analíticas sempre terão uma variedade de funções, habilidades, ferramentas favoritas e títulos. Uma diversidade de experiências e opiniões aumenta a inovação e a produtividade.
5. Interações diárias: clientes, equipes analíticas e operações devem trabalhar juntos diariamente ao longo do projeto.
6. Auto-organizar: acreditamos que os melhores insights analíticos, algoritmos, arquiteturas, requisitos e projetos surgem de equipes auto-organizadas.
7. Reduza o heroísmo: à medida que o ritmo e a amplitude da necessidade de insights analíticos aumentam,

acreditamos que as equipes analíticas devem se esforçar para reduzir o heroísmo e criar equipes e processos analíticos de dados sustentáveis e escaláveis.

8. Refletir: As equipes analíticas devem ajustar seu desempenho operacional refletindo-se, em intervalos regulares, no feedback fornecido por seus clientes, por si mesmas e pelas estatísticas operacionais.
9. Análise é código: as equipes analíticas usam uma variedade de ferramentas individuais para acessar, integrar, modelar e visualizar dados. Fundamentalmente, cada uma dessas ferramentas gera código e configuração que descreve as ações realizadas nos dados para fornecer insights.
10. Orquestrar: A orquestração do início ao fim de dados, ferramentas, código, ambientes e o trabalho das equipes analíticas é um fator chave para o sucesso analítico.
11. Torne-o reproduzível: resultados reproduzíveis são necessários e, portanto, temos a versão de tudo: dados, configurações de hardware e software de baixo nível e o código e configuração específicos para cada ferramenta no conjunto de ferramentas.
12. Ambientes descartáveis: Acreditamos ser importante minimizar o custo para os membros da equipe analítica experimentarem, proporcionando-lhes ambientes técnicos fáceis de criar, isolados, seguros e descartáveis que reflitam seu ambiente de produção.
13. Simplicidade: acreditamos que a atenção contínua à excelência técnica e ao bom design aumentam a agilidade; da mesma forma, a simplicidade - a arte de maximizar a quantidade de trabalho não realizado - é essencial.
14. Análise é fabricação: os pipelines analíticos são análogos às linhas de manufatura enxuta. Acreditamos que um conceito fundamental de DataOps é um foco no pensamento de processo visando alcançar eficiências contínuas na fabricação de insights analíticos.
15. A qualidade é fundamental: os pipelines analíticos devem ser construídos com uma base capaz de detecção automática de anormalidades (jidoka) e problemas de segurança no código, configuração e dados, e devem fornecer feedback contínuo aos operadores para evitar erros (poka yoke).
16. Monitorar a qualidade e o desempenho: Nosso objetivo é ter medidas de desempenho, segurança e qualidade que sejam monitoradas continuamente para detectar variações inesperadas e gerar estatísticas operacionais.
17. Reutilizar: Acreditamos que um aspecto fundamental da eficiência de fabricação de visão analítica é evitar a repetição do trabalho anterior por parte do indivíduo ou da equipe.
18. Melhorar os tempos de ciclo: devemos nos esforçar para minimizar o tempo e o esforço para transformar uma necessidade do cliente em uma ideia analítica, criá-la no desenvolvimento, liberá-la como um processo de produção repetível e, finalmente, refatorar e reutilizar esse produto.

Ao analisar esses princípios, é possível ver alguns pontos onde a InterSystems IRIS pode ajudar:

- Satisfaça continuamente seu cliente: você pode criar novas produções curtas de integração, orquestrações, cubos IRIS, relatórios, visualizações de BI e modelos de ML por sprints ou iterações.
- Análise de trabalho de valor: o IRIS ajuda você a fornecer dados com qualidade (usando produções, adaptadores e métodos de classe nas classes persistentes) e permite que você faça a exploração de dados em tabelas dinâmicas IRIS BI (designer de análise) e em IRIS NLP (análise de texto).
- Auto-organização: o IRIS simplifica a auto-organização, porque com uma plataforma de dados unificada, você coleta, processa, analisa e publica insights, com uma ferramenta.
- Refletir: no Portal do usuário você pode interagir com os usuários e coletar feedback para melhorar os produtos entregues.
- Análise é código: tudo é código, incluindo, modelo de dados IRIS, cubos e painéis; com controle de versão e governança.
- Orquestrar: o IRIS é uma plataforma de dados que orquestra a ingestão de dados, enriquecimento, trabalho analítico, visualização de dados e ML sobre dados, em uma única ferramenta, IRIS.
- Torne-o reproduzível: o IRIS utiliza docker, kubernetes (IKO) e devops para reproduzir os resultados.
- Ambientes descartáveis: o IRIS suporta a criação de ambientes docker descartáveis para integração, modelos de dados, cubos de BI e visualizações.
- Simplicidade: a criação do cubo de dados IRIS é muito simples e elimina a criação de scripts ETL, a criação de análises, cubos, dashboards são visuais, web e passíveis de serem feitos pelos usuários, não apenas pela equipe de desenvolvedores. E IntegratedML permite criar ML para cenários comuns sem desenvolvimento de código-fonte.
- Monitore a qualidade e o desempenho: a IRIS usa o SAM para monitorar o desempenho e tem um portal de gerenciamento da web.
- Reutilizar: no IRIS, os artefatos DataOps são classes e as classes são extensíveis e reutilizáveis por

padrão.

- Melhore os tempos de ciclo: os usuários podem criar dashboards, análises, relatórios, publicar e compartilhar seu trabalho em autoatendimento.

O ODSC

(<https://opendatascience.com/maximize-upstream-dataops-efficiency-through-ai-and-ml-to-accelerate-analytics/>)

indica a seguinte estratégia de DataOps:

O InterSystems IRIS auxilia nos pontos acima, veja:

- Provisionamento de autoatendimento: os usuários podem criar e publicar cubos e painéis.
- Compartilhe, marque, anote: o portal do usuário pode ser usado para compartilhar painéis, o IRIS Analytical Web Portal permite que o usuário crie, documente, organize em pastas e marque seu trabalho.
- Enrichement: BPL pode ser usado para enriquecer dados.
- Preparação: BPL, DTL, adaptadores e lógica ObjectScript podem ajudar na preparação de dados.
- Mercado de dados: os ativos de dados podem ser publicados na API REST e monetizados com o IRIS API Manager.
- Catálogo de dados: os dados no IRIS são organizados em classes, essas classes são armazenadas no sistema de catálogo de classes (% Dicionário)
- Perfil e Classificação: podem ser criados grupos, pastas para artefatos analíticos no Portal do Usuário e Portal do Administrador.
- Qualidade: IRIS tem classes utilitárias para gerar dados de amostra e fazer testes de unidade.
- Linhagem: no IRIS todos os ativos de dados são conectados, a partir do modelo de dados você constrói cubos, dos cubos você constrói painéis e todos os ativos de dados podem ser controlados por curadores de dados (sistema de permissão IRIS)
- Mastering: Admin Portal permite que você domine todos os aspectos em projetos analíticos.
- Dados de banco de dados, dados de arquivos, API SaaS, fluxos: IRIS é multimodelo e oferece suporte à persistência e análise de dados e texto (PNL). Suporta SaaS API usando IRIS API Manager e funciona com Streams usando Adaptadores de Integração e PEX (com kafka).
- Aplicativos, ferramentas de BI, sandboxes analíticos: com IRIS, você pode criar aplicativos DataOps com sua linguagem preferida (Java, Python, .NET, Node.js, ObjectScript). IRIS é uma ferramenta de BI, mas você pode usar conectores com Power BI ou ponte MDX e IRIS é uma área restrita para análises, em uma única ferramenta.

Veja meu resumo de mapeamento de IRIS e DataOps:



**DataOps com InterSystems IRIS**



[#Analytics](#) [#InterSystems IRIS](#)

---

URL de origem: <https://pt.community.intersystems.com/post/dataops-com-intersystems-iris>