



IPSOS VIEWS

# TESTE DE CONCEITO COM GÊMEOS DIGITAIS

Série "Humanizando a IA", parte 4  
Como dados sintéticos aceleram  
a inovação

Colin Ho, Ph.D  
Jiongming Mu  
Yuding Duan



**Na Ipsos, defendemos a combinação única de Inteligência Humana (HI) e Inteligência Artificial (AI) para impulsionar a inovação e oferecer insights impactantes e centrados nas pessoas para nossos clientes.**

**Nossa Inteligência Humana nasce da nossa expertise em prompt engineering, ciência de dados e de nossos conjuntos de dados próprios, de alta qualidade – que incorporam criatividade, curiosidade, ética e rigor às nossas soluções de IA, impulsionadas pela nossa plataforma de Gen AI, Ipsos Facto. Assim, nossos clientes têm acesso a insights mais seguros, mais rápidos e ancorados no contexto humano.**

**#IpsosHiAi**



**82% dos líderes esperam que a inovação impulse o crescimento nos próximos três anos, mas 60% acreditam que seus ciclos de inovação ficam aquém do ideal.**



## Dados sintéticos estão transformando a pesquisa de mercado – e isso é só o começo.

Na Ipsos, temos explorado as fronteiras do uso de dados sintéticos para boosting, imputação, fusão, persona bots e gêmeos digitais – buscando continuamente novas maneiras de acelerar o desenvolvimento e ampliar a compreensão do ser humano.

Na nossa publicação anterior, [\*The Power of Product Testing with Synthetic Data\*](#)<sup>1</sup>, mostramos como o uso de dados sintéticos para boosting pode acelerar os ciclos de desenvolvimento de produto, ao mesmo tempo em que oferece insights mais ricos sobre subgrupos. Partindo daí, “Concept Testing with Digital Twins” aprofunda o olhar sobre uma etapa mais inicial do funil de inovação, focando em como gêmeos digitais podem aprimorar a triagem/priorização de ideias e conceitos.

Profissionais de marketing e de insights em todo o mundo estão sentindo a pressão: **82% dos líderes esperam que a inovação impulse o crescimento nos próximos três anos, mas 60% acreditam que seus ciclos de inovação ficam aquém do ideal**<sup>2</sup>. Com milhões de dólares em jogo a cada lançamento de produto, as marcas precisam equilibrar, com dificuldade, velocidade e precisão.

Este paper oferece orientações claras sobre quando usar soluções de dados sintéticos para alcançar metas de inovação, quais são os diferentes tipos disponíveis e como garantir que essas soluções sejam confiáveis e adequadas ao objetivo. Nele, apresentamos:

- O papel dos gêmeos digitais na triagem/priorização de inovações
- A efetividade de diferentes abordagens
- Ferramentas adequadas para diferentes níveis de risco

Há mais de cinquenta anos, a Ipsos está na linha de frente da inovação centrada no consumidor, ajudando milhares de marcas, em diversos setores e no mundo todo, a lançar novos produtos e serviços que gerem conexão e tenham sucesso no mercado.

Mostramos como profissionais de marketing e de insights podem acelerar seus ciclos de inovação e confiar nas previsões — **tudo depende de como essas previsões são geradas.**

## Mas antes de tudo, o que é um gêmeo digital?

A expressão “gêmeos digitais” significa coisas diferentes para indústrias diferentes. Se você buscar esse termo, vai encontrar muitas definições – a maioria ligada a objetos físicos ou processos de manufatura. Já na pesquisa de mercado, gêmeos digitais são representações virtuais, feitas por IA, de pessoas reais, geradas a partir de dados reais, para simular atitudes, tomada de decisão e comportamentos. Gêmeos digitais permitem respostas instantâneas a perguntas e projeções para a população, desde que sejam construídos com amostras humanas suficientemente grandes e representativas (por exemplo, 1.000 gêmeos digitais para representar os consumidores de um país). O principal benefício dos gêmeos digitais é acelerar **a velocidade com que os insights podem ser obtidos e as hipóteses podem ser exploradas, antes de partir para a validação com pessoas reais.**

Gêmeos digitais são modelos orientados por dados voltados principalmente para simulação, entre outros usos. O processo para estabelecer a confiabilidade de gêmeos digitais é, portanto, semelhante ao dos modelos preditivos de IA que existiam antes dos Large Language Models (LLMs). Gêmeos digitais são confiáveis se predizem com precisão aquilo que queremos prever. Ao contrário dos modelos de IA anteriores aos LLMs, porém, gêmeos digitais podem ser usados de forma qualitativa ou quantitativa. Por isso, os requisitos de qualidade das respostas variam de acordo com a aplicação.

Este paper foca em casos de uso quantitativos para gêmeos digitais. Para decisões de baixo risco, em que a rapidez na obtenção de insights é o fator principal, fabricantes podem aceitar previsões que sejam apenas “boas o suficiente”. Já se os gêmeos digitais forem

usados em decisões de negócio de alto risco, a precisão precisa ser maior. Em resumo, gêmeos digitais não são, por natureza, bons ou ruins. A questão é se eles são bons o suficiente para o uso que se pretende dar a eles. Nas próximas páginas, vamos nos concentrar no desenvolvimento e na validação de gêmeos digitais com o propósito específico de triagem/priorização em massa de novas ideias ou conceitos de produto em determinados domínios (por exemplo, alimentos) e mercados (por exemplo, EUA). Com o avanço da IA, profissionais de marketing e de insights conseguem gerar grandes volumes de ideias de produto em questão de minutos ou horas. Nesses casos, eles precisam de uma forma eficiente de filtrar centenas de ideias, de modo que um conjunto menor possa então ser mais bem desenvolvido e avaliado com maior rigor por consumidores reais. Embora foquemos nesse caso específico, os princípios apresentados neste paper se aplicam ao desenvolvimento e à validação de gêmeos digitais em geral.



**Gêmeos digitais são representações virtuais, feitas por IA, de pessoas reais, geradas a partir de dados reais para simular atitudes, tomada de decisão e comportamentos.**



**Pessoas reais, individuais**



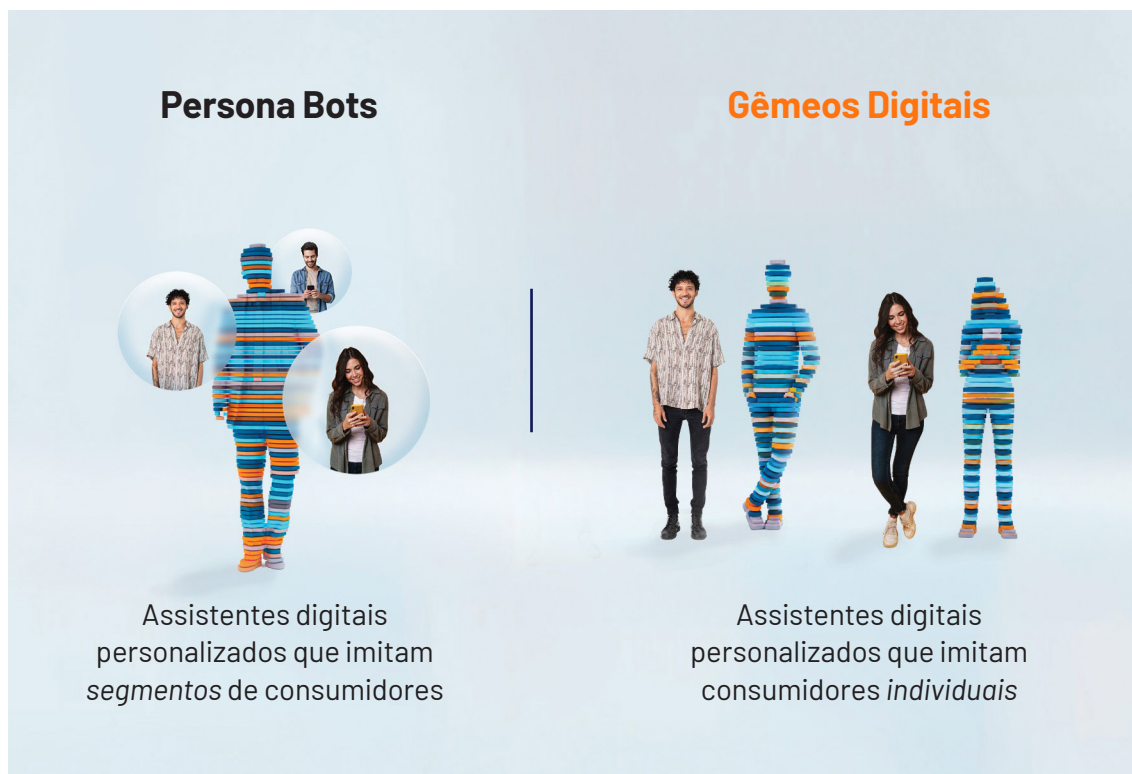
**Representações em IA (gêmeos digitais)**

## O que diferencia gêmeos digitais de persona bots?

Gêmeos digitais não devem ser confundidos com persona bots. Persona bots são representações virtuais em IA de *grupos de indivíduos*. Para representar cinco segmentos de consumidores, por exemplo, poderíamos desenvolver cinco personas em IA, uma para cada segmento. Persona bots são construídos usando dados agregados (por exemplo, resumos em texto de cada segmento). Já os gêmeos digitais são construídos com dados em nível individual. As respostas de persona bots são apenas qualitativas

e não devem ser usadas para projeções quantitativas. Gêmeos digitais são apenas uma forma de dado sintético. Data boosting, imputação e fusão de dados representam outras formas. Essas outras formas, porém, não permitem que o pesquisador faça novas perguntas – elas servem apenas para ampliar ou aprimorar conjuntos de dados *numéricos* já existentes. Não discutimos essas outras formas aqui e encaminhamos o leitor interessado para outras fontes<sup>3</sup>.

Figura 1: A diferença entre persona bot e gêmeos digitais



Fonte:  
Ipsos

## Saiba o que você está tentando prever

O mantra da inteligência humana e inteligência artificial (HI+AI) tem sido repetido tantas vezes que corre o risco de se esvaziar e perder sua força. Mas a interpretação do que HI+AI realmente significa nem sempre é bem compreendida. HI+AI não é apenas ter pessoas revisando o que a IA produz. Na verdade, HI+AI *significa entrelaçar a inteligência humana na própria essência de uma solução de IA*. Isso quer dizer incorporar aprendizados de pesquisas anteriores ou da experiência sobre como as pessoas pensam, sentem, acreditam ou se comportam.

Dois grandes princípios de pesquisas anteriores são fundamentais ao criar gêmeos digitais. Primeiro, **o princípio da especificidade** vem da psicologia social. Esse princípio afirma que o poder preditivo das atitudes sobre os comportamentos é maior quando as atitudes são mensuradas em um nível de *especificidade que corresponda de perto ao comportamento de interesse*.

Por exemplo, uma atitude especificamente em relação a “comer saladas vegetarianas” vai prever com mais precisão o comportamento de “escolher uma salada vegetariana no almoço” do que uma atitude ampla em relação a “alimentação saudável”. Esse princípio destaca a importância de alinhar o nível de especificidade das perguntas sobre atitudes ao comportamento ou resultado específico que está sendo estudado, aumentando assim a precisão das previsões de comportamento.

Aplicado a gêmeos digitais, isso significa que precisa haver uma relação direta entre as informações que usamos para construir os gêmeos e aquilo que estamos tentando prever. As informações usadas para construir gêmeos digitais não precisam se limitar a atitudes, mas devem ser diretamente relevantes ao que queremos prever. Isso significa que você precisa saber, antes de tudo, o que está tentando prever! Seguindo esse princípio, definimos um objetivo específico para o nosso primeiro modelo de gêmeos digitais: prever como consumidores *americanos* reagiriam a novas ideias de produtos alimentícios. Limitamos o escopo do modelo apenas a alimentos porque atitudes e comportamentos variam enormemente entre diferentes categorias de bens de consumo embalados. Também restringimos o escopo a um único país, já que atitudes e escolhas relacionadas a *comida* diferem de um país para outro. Em essência, estamos treinando LLMs, que são os blocos de construção dos gêmeos digitais, para responder a perguntas de pesquisas como se fossem um consumidor representativo dos EUA diante do cenário de compra em uma determinada categoria.



HI+AI não é apenas ter pessoas revisando o que a IA produz. Na verdade, HI+AI significa entrelaçar a inteligência humana na própria essência de uma solução de IA.



## Afinando nosso foco

Se seguíssemos o princípio da especificidade à risca, poderíamos ser ainda mais específicos ao prever o comportamento do consumidor em uma categoria. Poderíamos, por exemplo, construir um modelo que previsse como consumidores dos EUA reagiriam a novas ideias de produtos de *iogurte*, em vez de ideias de alimentos em geral. Para construir esse modelo, precisaríamos criar nossos gêmeos digitais com base nas atitudes de consumidores americanos em relação ao iogurte. But following the principle of specificity Mas seguir o princípio da especificidade à letra pode ser impraticável. Existe um trade-off entre especificidade e generalização. Um painel de gêmeos digitais criado para prever a aceitação de novas ideias de iogurte só pode ser usado para prever novas ideias de iogurte, e não alimentos congelados, snacks, sobremesas, biscoitos ou qualquer outra categoria de alimentos. **Por isso, é preciso haver um equilíbrio entre especificidade e pragmatismo.**

Pesquisas anteriores da Ipsos indicaram que existem muitos pontos em comum nos drivers de compra em diversas subcategorias de alimentos. Para ampliar as aplicações do nosso modelo de gêmeos, decidimos construir gêmeos digitais usando atitudes gerais em relação a alimentos como um todo, de modo que nossos gêmeos pudessem fazer previsões em diferentes categorias de alimentos. Pode haver alguma perda de precisão, mas, como o objetivo é a triagem/priorização em massa em uma etapa inicial, o alcance de aplicação se torna mais amplo. Um modelo geral de alimentos também permitiria que os profissionais de marketing avaliassem novas ideias de produto “nas bordas” ou que rompem com as categorias atuais (ou seja, novas ideias de produto que não se encaixam perfeitamente em subcategorias de alimentos pré-definidas). Ter um modelo geral de alimentos, no entanto, não nos impede de, no futuro, aprimorar os gêmeos digitais com dados de treinamento específicos de subcategorias.

## Ampliando nosso alcance

O princípio da especificidade vai contra a eficiência. O “santo graal” seriam gêmeos digitais que pudessem ser usados para prever *qualquer* atitude, reação ou comportamento do consumidor para qualquer produto ou serviço. Pesquisadores da Universidade de Stanford tinham esse objetivo em mente, ainda que no domínio da formulação de políticas públicas e das ciências sociais. Os pesquisadores construíram gêmeos digitais usando dados de uma entrevista em áudio de duas horas, que abrangia uma ampla variedade de tópicos, desde as histórias de vida dos participantes (por exemplo: *“Conte a história da sua vida – da infância, passando pela educação, família e relacionamentos, até quaisquer grandes eventos de vida que você tenha tido”*) até suas visões sobre questões sociais atuais (por exemplo: *“Como você reagiu ao foco crescente em raça e/ou racismo e policiamento?”*).

Com esses gêmeos digitais, os pesquisadores conseguiram replicar as respostas dos participantes no General Social Survey em uma ampla variedade de temas sociais, como liberdades civis, crime, relações entre grupos, saúde, moralidade e questões nacionais. A precisão foi de 85%, definida como a capacidade de reproduzir as próprias respostas dos participantes duas semanas depois. Essa pesquisa gerou grande entusiasmo, pois forneceu evidências de que gêmeos digitais de uso geral poderiam replicar o comportamento humano em muitos domínios relacionados à formulação de políticas públicas e às ciências sociais.

Seria difícil para nós alcançar o mesmo nível de sucesso no mundo de consumer insights. Em primeiro lugar, não é nada prático conduzir entrevistas de duas horas. A maioria dos consumidores não aceitaria isso. Com remuneração, alguns até poderiam aceitar, mas a gama de temas de consumo que precisaríamos cobrir iria muito além daqueles relacionados à formulação de políticas públicas. Além disso, como os mercados mudam com frequência, os gêmeos digitais precisam ser atualizados. Entrevistas longas e cansativas tornariam essas atualizações proibitivamente caras ou praticamente impossíveis. Mais importante ainda, os insights buscados em pesquisa de mercado precisam ser granulares e específicos. Por exemplo: uma formulação com alto teor de proteína será mais valorizada do que uma com baixo teor de açúcar em uma nova barra de cerea?

Treinar gêmeos com temas sociais gerais não chega a esse nível de granularidade e não é específico o suficiente para prever como uma pessoa pode reagir a ideias de novos produtos específicos. Desenvolver gêmeos digitais de uso geral para avaliar QUALQUER comportamento de mercado não é uma solução viável. Dito isso, o método de construção de modelo usado pelos pesquisadores de Stanford para transformar dados de treinamento em nível individual em seus gêmeos digitais serve como um bom modelo e é aplicado na nossa prática. A diferença, porém, é que os dados de treinamento que usamos para construir nossos gêmeos são focados em uma aplicação muito específica.

## Humanizando gêmeos digitais

Para espelhar o processo de decisão de compra no mundo real, aproveitamos o conhecimento de duas unidades de negócio da Ipsos para construir nossos gêmeos digitais:

- 01 Unidade de Market Strategy and Understanding da Ipsos**, que revela o que as pessoas pensam, sentem, dizem ou fazem
- 02 Unidade de Inovação da Ipsos**, que se concentra em prever o sucesso de inovações

Os aprendizados da nossa unidade de Market Strategy and Understanding orientaram as informações que usamos para construir os gêmeos digitais (ver Figura 2). Capturamos tudo o que descobrimos sobre o que impulsiona as escolhas alimentares. Isso inclui desde os ingredientes específicos que as pessoas procuram nas embalagens (por exemplo, baixo teor de açúcar, quantidade de proteína) até atitudes relacionadas à alimentação, como se a pessoa prioriza conveniência, preço, prazer, saúde ou preocupações ambientais na hora de escolher o que comer. Também coletamos informações sobre a probabilidade de as pessoas trocarem de marca de alimentos. A disposição para experimentar novos produtos alimentícios está fortemente ligada ao grau de apego ao que elas consomem hoje. Não coletamos nenhuma informação sobre preço, já que o objetivo era fazer a triagem de conceitos iniciais ainda sem preço. Construir gêmeos digitais que reflitam com precisão a sensibilidade de cada indivíduo a preço seria uma tarefa muito mais complexa.

Em vez de usar respostas a estímulos passados do banco de dados, usamos atitudes, preferências, comportamentos e hábitos para treinar os gêmeos digitais – permitindo que eles gerem respostas como consumidores reais diante de novos estímulos, sem depender de dados antigos. O princípio é simples: os insights de ontem não preveem as ações de hoje. Esse método de treinamento é fundamental para garantir que os gêmeos estejam “prontos para o futuro”, em vez de apenas repetir respostas a estímulos antigos.

As informações foram coletadas por meio de perguntas fechadas em pesquisa, com o apoio de um bot de moderação em IA. Essa combinação única de dados estruturados e não estruturados forneceu uma fonte rica de informação para construir o perfil de cada gêmeo digital e exemplifica como a inteligência humana pode ser diretamente entrelaçada em uma ferramenta de IA para consumer insights.



**Em vez de usar respostas a estímulos passados do banco de dados, usamos atitudes, preferências, comportamentos e hábitos para treinar os gêmeos digitais – permitindo que eles gerem respostas como consumidores reais.**

Figura 2: Input dos gêmeos digitais dos EUA para previsão de novas ideias de alimentos

## Quais influências de categoria moldam a modelagem de gêmeos digitais?

**Exemplo:** cinco dimensões que impactam na escolha de alimentos

### Preferências

- Açúcar, gordura, proteína
- Preferências alimentares
- Rotina de lanches

Entre outros

### Objetivos

Necessidades não atendidas e aspirações em relação à alimentação

### E... dimensões pessoais

- Composição do domicílio
- Variáveis demográficas
- Ocasões e estilo de vida relacionados à alimentação

### Atitudes

- Facilidade/conveniência
- Valor/sensibilidade a preço
- Prazer/relaxamento
- Atitude em relação à saúde
- Preocupações alimentares

Entre outros

### Atitudes / comportamento em relação à marca

Atitude em relação às marcas atuais, comportamento de troca, busca por variedade etc.

### Hábitos de compra

Leitura de rótulos, planejamento de compra, busca de informações etc.



O conhecimento da nossa unidade de Inovação orientou o que pedimos para os gêmeos preverem. Quando os consumidores se deparam com um novo produto no mundo real, eles não avaliam esse produto no vácuo. Na verdade, a escolha é sempre entre um produto já estabelecido e a nova oferta. Por exemplo, imagine que você veja um novo produto de iogurte. Se você já compra uma determinada marca de iogurte, então a escolha é entre continuar com sua marca atual ou experimentar o novo produto. Se você não compra iogurte hoje, a escolha é entre manter sua rotina alimentar atual ou experimentar um novo iogurte. Portanto, os gêmeos digitais foram configurados para prever se um consumidor escolheria um produto já estabelecido ou o novo produto. Essas escolhas foram feitas com base em métricas-chave que nosso banco de dados demonstrou estar correlacionadas com o sucesso no mercado (por exemplo, relevância). Seguir os protocolos validados sobre o que pedimos para os gêmeos digitais preverem também permite

que os dados gerados sejam usados em modelos quantitativos já existentes que podem ser comparados ao nosso banco de dados; isso torna a análise e a apresentação dos resultados das previsões dos gêmeos digitais fáceis de interpretar, por serem comparáveis a pesquisas de teste de conceito feitas com humanos.

Isso ilustra o segundo princípio que seguimos ao construir nossos gêmeos: **o princípio do que nos torna humanos**. Esse não é um princípio acadêmico formal. É simplesmente a ideia de que existem traços que nos tornam exclusivamente humanos. Além de fazermos escolhas relativas, há outros fatores que tornam o raciocínio humano único: limitações de memória, distrações, fadiga, emoções, valores e contexto. Gêmeos digitais não são afetados por nenhum desses. Ao construir gêmeos digitais, devemos levar esses fatores em conta. O objetivo, porém, é torná-los parecidos com humanos, não humanos. Só conseguimos ir até certo ponto.

## Avaliando precisão e estabilidade

Para serem úteis, as previsões feitas por gêmeos digitais precisam ser precisas. Embora muitos pensem em precisão como a capacidade de reproduzir as respostas que os humanos dariam, também precisamos considerar o conceito de confiabilidade: obter previsões que não apresentem grandes flutuações também é importante. Confiabilidade significa que os resultados são relativamente estáveis entre amostras ou ao longo do tempo. Precisamos dos dois. Gêmeos digitais podem ser confiáveis, mas não precisos: você obtém a mesma resposta errada todas as vezes! Um modelo de gêmeos também pode ser preciso, mas não confiável: de vez em quando você até consegue um resultado correto, mas seus resultados podem variar de um projeto para outro. O ideal é ter ambos: precisão e confiabilidade.

Também precisamos falar sobre acurácia em nível de pessoa versus acurácia agregada.

No mundo de ciência de dados, a acurácia em nível de pessoa é a forma de acurácia mais aceita. Mas, como estamos construindo uma ferramenta de triagem/priorização em estágio inicial para ajudar marcas a reduzir um conjunto inicial de novas ideias de produto, o nível de pessoa não é o critério mais crítico. A validação da acurácia em nível individual é particularmente relevante quando o objetivo é personalizar produtos ou serviços às preferências de cada pessoa. Embora isso seja importante, não é a meta principal na priorização inicial de ideias de produto. Já a validação da acurácia em nível agregado se torna mais relevante quando o objetivo é entender tendências mais amplas ou fazer previsões em nível de população ou de grupo. Essa abordagem é benéfica para empresas que desenvolvem estratégias ou campanhas voltadas a grandes segmentos de consumidores e é exatamente o que se faz em testes de conceito e de produto.

Dessa forma, avaliamos nosso modelo de gêmeos digitais de alimentos dos EUA usando acurácias agregadas (ver Figura 3). Compartilhamos as previsões para 12 conceitos provenientes de um conjunto diverso de categorias de alimentos. As previsões dos gêmeos digitais foram comparadas aos resultados humanos no índice de experimentação validado da Ipsos e ao quintil em que esse índice se enquadra. O quintil é uma classificação de desempenho de conceito com base no banco de dados da Ipsos. Um conceito pode ser classificado como Alto (H), Médio Alto (MH), Médio (M), Médio Baixo (ML) e Baixo (L). Não exigimos uma correspondência perfeita de quintil, já que os gêmeos são destinados à triagem/priorização de ideias em estágio inicial, para reduzir grandes volumes de ideias. As ideias que passarem por essa triagem inicial seriam, posteriormente, reavaliadas de forma mais rigorosa. Se um conceito ficar dentro de +/-1 quintil em relação à avaliação humana, consideramos isso uma acurácia suficientemente boa (isto é, um "aprovado").

Entre os 12 conceitos, 11 foram aprovados usando esse critério. Nossas descobertas fornecem evidências iniciais de que gêmeos digitais podem ser usados para prever o sucesso de novos produtos de maneira razoável quando treinados com os dados adequados.

Em estudos subsequentes de P&D e pilotos, constatamos de forma consistente que os gêmeos digitais conseguem igualar as respostas humanas em mais de 85% dos casos quando o critério é definido como corresponder a +/-1 quintil da avaliação humana. Esse critério vem sendo usado desde o início dos anos 2020 com outra ferramenta de previsão de conceitos utilizada na Ipsos – o InnoPredict AI. Esse nível de precisão é adequado para a triagem em estágio inicial, que será seguida por testes mais rigorosos com respondentes humanos, e é amplamente aceito pelos fabricantes como "bom o suficiente" para iniciativas de baixo risco em situações práticas.

**Figura 3:** Exemplos de acurácia agregada: modelo de gêmeos digitais dos EUA (categoria de alimentos)

Índice de Potencial de Experimentação			
	Humano	Gêmeos digitais	Padrão de acurácia para aprovação
Conceito 1	M	H	Sem aprovação
Conceito 2	MH	M	Aprovado
Conceito 3	MH	MH	Aprovado
Conceito 4	MH	MH	Aprovado
Conceito 5	H	H	Aprovado
Conceito 6	MH	MH	Aprovado
Conceito 7	MH	H	Aprovado
Conceito 8	H	MH	Aprovado
Conceito 9	H	MH	Aprovado
Conceito 10	M	M	Aprovado
Conceito 11	M	MH	Aprovado
Conceito 12	M	M	Aprovado

Fonte:  
Ipsos

## E quanto à variabilidade?

Para colocar isso em perspectiva, testamos repetidamente o mesmo conjunto de conceitos em ondas separadas, porém equivalentes, de amostras humanas (por exemplo, balanceadas por idade e gênero), junto com seus gêmeos digitais. O que descobrimos foi que a variação entre humanos e seus gêmeos digitais na mesma onda não é maior do que a variação que observamos entre diferentes ondas de amostras humanas. Visto por esse ângulo, gêmeos digitais oferecem um grau de estabilidade semelhante ao do erro amostral, o que traz evidências adicionais de que usar gêmeos digitais para triagem/priorização de estímulos em estágio inicial pode não aumentar os riscos de acurácia além dos riscos amostrais inerentes à pesquisa por survey.

Uma diferença importante, porém, é que, nos casos em que os conceitos são apenas marginalmente diferentes entre si (como pequenas diferenças de redação na descrição de um ingrediente), os gêmeos digitais apresentam respostas que variam menos do que as dos respondentes humanos. Isso pode ser interpretado de forma positiva, como um sinal de que os gêmeos digitais são mais estáveis quando as diferenças entre conceitos são pequenas. Mas também pode ser interpretado como falta de capacidade dos gêmeos digitais de perceber nuances sutis entre conceitos, porque os dados de

treinamento nunca chegaram a esse nível de granularidade, enquanto os humanos conseguem. Em aplicações de negócios, isso significa que gêmeos digitais não são uma boa ferramenta para diferenciar ideias de novos produtos que só mudam em detalhes. Muitas vezes, pesquisadores testam conceitos com pequenas variações de texto ou conteúdo. Nossa recomendação é usar gêmeos digitais para testar a proposta central de novas ideias de produto que sejam claramente diferentes entre si. Gêmeos digitais não devem ser usados para testar pequenas diferenças de redação entre conceitos.



**Nossa recomendação é usar gêmeos digitais para testar a proposta central de novas ideias de produto que sejam claramente distintas entre si.**



## Como escolher a solução ágil certa de previsão de conceitos

Há diferentes maneiras de prever rapidamente o sucesso de novos produtos. Neste paper, apresentamos uma visão completa, para além de gêmeos digitais; a mais popular é construir modelos preditivos de IA usando bancos de dados de conceitos para prever o sucesso de um novo produto. Isso envolve analisar ideias de novos produtos que já foram testadas no passado. Se certos atributos tiveram bom desempenho antes, parte-se do pressuposto de que esses atributos também terão bom desempenho hoje. E, inversamente, se certos atributos tiveram desempenho ruim no passado, assume-se que eles também terão desempenho ruim agora. Esses modelos preditivos partem da premissa de que conhecer o passado permite prever o futuro. Em categorias em que os produtos mudam muito pouco (por exemplo, detergente líquido para lavar louça à mão), o passado pode de fato prever o futuro. Já em outras categorias em que os produtos mudam bastante (por exemplo, pense nos primeiros celulares em comparação com os que temos hoje), é improvável que o passado consiga prever o futuro. Essa é a principal fragilidade dos modelos de IA baseados no passado, quando a finalidade é prever inovação que, por definição, não existia anteriormente. Embora a previsão com base em dados históricos possa alcançar uma acurácia satisfatória em outros campos de insights,

o risco é alto quando utilizada para prever o sucesso de um novo produto.

Em vez de usar dados históricos, a Ipsos opera um modelo de previsão de inovação chamado InnoPredict AI, que utiliza machine learning baseado nas respostas em linguagem natural dos consumidores a um estímulo de inovação para prever o sucesso, detectando temas e emoções nessas respostas que indiquem potencial de adoção. Esse método, na prática, reduz o teste de conceito a uma única pergunta e pode ser escalado para qualquer país ou categoria. A desvantagem, porém, é que não se trata de uma previsão instantânea, já que ainda precisamos esperar que as amostras de consumidores respondam a esses estímulos.

O uso de persona bots para avaliar conceitos pode ser suficiente quando as marcas precisam de uma leitura qualitativa rápida ou de uma forma de otimizar conceitos para testes quantitativos no futuro. Insights qualitativos iniciais vindos de persona bots podem ser extremamente úteis. Mas, como persona bots representam grupos de pessoas, suas respostas não devem ser usadas para projeções quantitativas. Resumimos as diferentes abordagens na Figura 4.

Figura 4: Quando usar gêmeos digitais versus alternativas

*Risco	Solução	Aplicabilidade
Baixo	<b>PersonaBots</b> Modelos agregados de grupos que fornecem feedback qualitativo sobre conceitos	Melhor para Insights qualitativos  Limitação Sem projeções quantitativas
	<b>Modelos de banco de dados</b> Modelos preditivos agregados que fornecem feedback quantitativo sobre conceitos com base no passado	Categorias estáveis (por exemplo, detergentes para louça)  Pode falhar se os produtos inovarem substancialmente ao longo do tempo.
	<b>InnoTest com respondentes sintéticos</b> Simulações de consumidores individuais reagindo a conceitos de inovação	Triagem/priorização em massa de conceitos diversos  Não indicado para diferenças pequenas ou sutis entre conceitos
	<b>InnoPredict AI com humanos</b> Previsões de reações de consumidores a conceitos de inovação com inteligência artificial de machine learning	Países/categorias em que painéis de gêmeos digitais ainda não foram desenvolvidos  Trabalho de campo com pessoas reais significa que a previsão não é obtida instantaneamente.
Alto	<b>InnoTest com humanos</b> Previsões de reações de consumidores a conceitos de inovação usando benchmarks robustos e validados	Teste de validação, otimização de conceitos, teste de pequenas diferenças e coerência entre proposta/imagem/preço  <b>Sem limitação, o padrão-ouro</b>

Fonte: Ipsos

\*Nível de risco para inovação

## Direções futuras e ressalvas

Este paper detalhou o que é necessário para construir gêmeos digitais confiáveis que simulem decisões reais de consumidores – capacitando você a avaliar criticamente soluções concorrentes e a usar dados sintéticos para inovar com confiança. Não existem atalhos. Primeiro, precisamos de dados sob medida para a tarefa preditiva em questão e que reflitam o mercado e a categoria de produto para os quais queremos fazer previsões. LLMs prontos, pré-treinados “de prateleira”, por si só, não conseguem captar atitudes locais/culturais em relação a categorias de produto específicas. Depois de desenvolvidos, os gêmeos digitais precisam ser validados. Alguma forma de validação, mesmo que limitada, é necessária para entender se os gêmeos digitais conseguem prever com precisão em um mercado e categoria de produto específicos.

O trabalho não termina após a validação do modelo. Pode ser necessário atualizar um painel de gêmeos digitais se houver mudanças no mercado que alterem de forma significativa as preferências e atitudes dos consumidores. Seguindo com nosso exemplo de alimentos, se as pessoas passarem a acreditar que gordura animal faz bem, em vez de ser algo a ser evitado, teremos que atualizar os gêmeos digitais para refletir isso. Embora todos desejemos que a IA torne nosso trabalho mais eficiente, a verdade é que o desenvolvimento desses modelos exige tempo e esforço.

Uma área que merece mais atenção – se quisermos usar gêmeos digitais para fazer triagem/priorização em massa de novas ideias de produto – é o uso de benchmarks:

- Os benchmarks que usamos para indicar diferenças entre ideias para humanos também valem para gêmeos digitais?

- Ou precisamos de benchmarks completamente diferentes?

Se a saída dos gêmeos digitais tiver menor variabilidade, as diferenças consideradas estatisticamente significativas para humanos podem precisar ser ajustadas para dados de gêmeos digitais (por exemplo, as faixas que definem cada quintil podem precisar ser mais estreitas). Embora tenhamos boas ferramentas e convenções para lidar com a variabilidade em respondentes humanos e separar o sinal do ruído (por exemplo, teoria amostral, intervalos de confiança), essas ferramentas não se aplicam quando lidamos com resultados de modelos de IA. Por fim, vale notar que há variabilidade e inconsistência inerentes nas respostas de LLMs, que já foi demonstrado que LLMs individuais podem se degradar ao longo do tempo, e que novos/atualizados LLMs chegam ao mercado o tempo todo. Por causa dessa mudança constante nos LLMs, o treinamento dos gêmeos digitais precisará ser atualizado e recalibrado periodicamente.

Gêmeos digitais, como método, oferecem um futuro promissor para aplicações em pesquisa de inovação. Enquanto ampliamos gradualmente o uso dessa abordagem para cobrir mais países e categorias, também estamos explorando o uso de gêmeos para responder a outros tipos de questões de inovação, como otimização de conceitos e de atributos de produto. O treinamento dos gêmeos digitais pode ser enriquecido com dados de consumidores da Ipsos e da indústria de forma mais ampla. Os dados de treinamento específicos de inovação também podem ser adaptados às necessidades específicas de nossos clientes, tornando os gêmeos versáteis para demandas de pesquisa customizadas. Estamos apenas arranhando a superfície.

## Conclusão

Buscamos oferecer uma visão simples, porém clara, da nossa abordagem para desenvolver gêmeos digitais na Ipsos para triagem/priorização de ideias e conceitos. Embora gêmeos digitais tenham o potencial de transformar a pesquisa de mercado, é essencial que sua adoção seja cautelosa. Ao integrar um LLM de base, como o Gemini, com dados em nível individual, buscamos ampliar a capacidade do LLM de prever comportamentos e atitudes específicos.

Para chegar lá, é fundamental seguir o princípio da especificidade e incorporar as nuances do contexto humano. É importante lembrar que gêmeos digitais não são réplicas de seres humanos; eles são aproximações. Quando as decisões de comportamento são guiadas por um conjunto limitado de fatores racionais, gêmeos digitais podem ser

altamente eficazes. No entanto, quando influências emocionais ou contextuais entram em cena, é necessário adicionar contextualização humana ao modelo.

Além disso, gêmeos digitais devem ser aplicados principalmente em cenários de baixo risco, como a triagem em estágios iniciais, onde sua capacidade de aproximação pode ser melhor aproveitada. Existem trade-offs inerentes a qualquer aplicação tecnológica, e com gêmeos digitais não é diferente. Eles se destacam em triagens de alto volume e baixo risco, mas não são indicados para decisões finais de seguir ou não em frente. Em última análise, o futuro da pesquisa de mercado não é substituir seres humanos – é usar IA para tornar a pesquisa com humanos mais estratégica e impactante.



**O futuro da pesquisa de mercado não é substituir os seres humanos – é usar IA para tornar a pesquisa com pessoas mais estratégica e impactante.**



## Principais aprendizados

### 01

**Você PODE fazer a priorização de ideias e conceitos instantaneamente e ainda assim obter os resultados certos.** Uma vez desenvolvidos, gêmeos digitais são consumidores virtuais movidos a IA que podem ser usados para prever a aceitação de produtos em questão de horas, não semanas. Eles permitem a triagem/priorização em massa e imediata de centenas de ideias de produto – liberando profissionais de marketing e de insights para focar a pesquisa centrada em pessoas nos conceitos mais promissores e acelerar as decisões de time-to-market.

### 02

**O desenvolvimento dos gêmeos digitais faz toda a diferença.** Modelos de gêmeos digitais devem equilibrar especificidade (por exemplo, serem treinados com atitudes e comportamentos reais em relação a alimentos), flexibilidade (por exemplo, funcionar em todas as categorias de alimentos) e praticidade (por exemplo, questionário de 15 minutos). Além disso, é fundamental não simplesmente alimentar um LLM com qualquer tipo de dado. A ciência do comportamento deve ser incorporada sempre que necessário. No nosso exemplo, nossa inovação crucial foi garantir que os gêmeos avaliassem produtos como os humanos fazem – comparando as novas opções com seu status quo, e não de forma isolada.

### 03

**Use bem, ou é melhor nem usar.** Use gêmeos digitais para aquilo para que foram projetados. Uma vez desenvolvidos, é preciso tomar cuidado para garantir que os gêmeos digitais sejam usados apenas para seu propósito original. No nosso caso, projetamos nossos gêmeos digitais para se destacarem na triagem de conceitos fundamentalmente diferentes. Eles não foram feitos para testes A/B de pequenas mudanças de texto ou leves variações de redação. Garantir o uso adequado evita o mau uso da ferramenta.



**Para saber mais sobre como acelerar ciclos de inovação com dados sintéticos, entre em contato com a equipe local da Ipsos.**

## Notas metodológicas

### Construindo e validando gêmeos digitais para triagem/priorização de conceitos

Para construir perfis fundamentais abrangentes para nossos Gêmeos Digitais de Alimentos nos EUA, extraímos pontos de vista, preferências e estilos de linguagem únicos de 1.000 respondentes reais dos EUA, usando dados de pesquisa estruturados e não estruturados. Treinamos esses modelos usando 12 conceitos de alimentos, aumentando sistematicamente os dados, alterando o tom, o comprimento e a estrutura para garantir que os modelos fossem robustos diante de entradas variadas.

Para prever com precisão tanto KPIs quantitativos quanto feedback qualitativo, ajustamos finamente modelos avançados de Large Language Models usando técnicas de adaptação de última geração, o que garante que os gêmeos digitais forneçam respostas que se assemelham às de humanos reais em variabilidade, realismo da linguagem e positividade dos sentimentos.

Para respostas qualitativas, criamos uma estrutura com uma “camada de sentimento” preliminar. Ela funciona como uma espécie de “intuição” humana, garantindo que as opiniões geradas pelo gêmeo se alinhem naturalmente à sua persona específica antes que ele formule uma resposta completa.

Quando um gêmeo digital detecta que um conceito tem baixa relevância, ele aciona um processo de otimização passo

a passo. O gêmeo revisa suas preferências pessoais, identifica inconsistências no conceito e gera sugestões direcionadas de melhoria.

### Validação e otimização rigorosas

Para comprovar que nossa abordagem funciona, aplicamos uma validação científica rigorosa. Usando uma abordagem de validação cruzada leave-one-out, testamos repetidamente o desempenho do modelo em dados totalmente inéditos. No fim, validamos o modelo de alimentos dos EUA em 30 conceitos distintos de alimentos, alcançando uma taxa de acurácia de 97% em relação à nossa métrica proprietária de potencial de experimentação (definida como estar dentro de +/-1 quintil de potencial de experimentação).

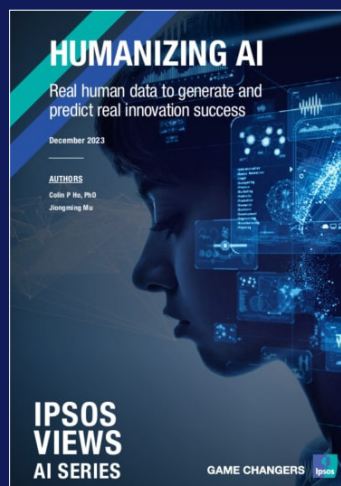
Para comprovar que essa abordagem é generalizável, recentemente desenvolvemos um modelo para ração pet nos EUA, que atingiu 88% de acurácia em 16 conceitos de alimentos para pets. Importante ressaltar que esse padrão de acurácia de +/-1 quintil não é um novo benchmark de IA; ele é um padrão já estabelecido na Ipsos, no qual nossos clientes confiam há muito tempo para a triagem inicial com humanos.



## Notas finais

- 1 (2025) Ho, C. Reynolds, N. Ipsos Views - The Power of Product Testing with Synthetic Data.
- 2 (2026) Ipsos and MarketLogic - Innovation, Reignited. Base: Executive-weighted sample (n=250) across CPG and adjacent sectors - 64% CEO, 24% CMO, 11% CGO, 2% Regional/Market/President/CEO - from organizations with \$100M to \$2B in annual revenue, fielded September 5-October 16, 2025.
- 3 (2024) Guidi, M. Hubert, B. Sava, C. Timpone, R. Ipsos Views - Synthetic Data, a Guide to Responsible Adoption.

# Leituras adicionais



JUNHO 2026

# TESTE DE CONCEITO COM GÊMEOS DIGITAIS

Série "Humanizando a IA", parte quatro

## AUTHORS

### **Colin Ho, Ph.D**

Chief Research Officer,  
Innovation and Market Strategy  
& Understanding, Ipsos

### **Yuding Duan**

Data Scientist, Innovation,  
Ipsos

### **Jiongming Mu**

Global Innovation Solutions  
Leader, Ipsos

Os white papers **Ipsos Views**  
são produzidos pelo **Ipsos**  
**Knowledge Centre.**

[www.ipsos.com](http://www.ipsos.com)

@Ipsos

