This document was translated from English by machine and may contain mistakes. The English original is at https://www.novaracg.com/2025/05/23/four-numbers-and-no-counterfactual-reading-the-artificial-intelligence-case-study-in-human-resources/.

Dört Sayı ve Sıfır Karşı Olgusal Analiz: İnsan Kaynaklarında Yapay Zeka Vaka Çalışmasını Okumak

Heather Grizzle · May 23, 2025

Dört yüzde iyileştirmesi ve karşı olgusal hiçbir şey. Standart bir İK yapay zeka vaka çalışması metriğinin her birinin doğru olması için neyin gerekli olduğu ve on dakika süren sorular.

İçindekiler

1. Tür
2. Birinci iddia: işe alım süresinde azalma
3. İkinci iddia: manuel işlemsel faaliyette azalma
4. Üçüncü iddia: yüksek potansiyelli çalışanların elde tutulmasında artış
5. Dördüncü iddia: önyargının azaltılması
6. Dördünde de eksik olan şey
7. İddia sahibi bunu genellikle neden inanılır bulur
8. Alıcının sorması gereken sorular
9. İnandırıcı bir vaka çalışması neler içerirdi
10. Sonuç
11. Kaynakça

İnsan kaynaklarındaki yapay zeka vaka çalışması istikrarlı bir biçime oturmuş durumda.

Tür

İnsan kaynaklarındaki yapay zeka vaka çalışması istikrarlı bir biçime oturmuş durumda. Bir kuruluş bir uygulamayı tanımlar, üç veya dört puanlık iyileştirmeler

bildirir, önyargının azaltılmasına dair bir iddia ekler ve yetenek yönetiminin rekabet avantajı kaynağı haline geldiği sonucuna varır. Rakamlar kesindir. Metodoloji yoktur. Sonuç bir görüşmeye davet niteliğindedir.

Bu biçim varlığını sürdürüyor çünkü ikna aracı olarak işe yarıyor ve kimse bunu denetleyecek konumda değil. Satıcılar bunu yayınlıyor çünkü satış sağlıyor. Uygulayıcılar bunu yayınlıyor çünkü itibar kazandırıyor. Alıcılar bunu kabul ediyor çünkü aksi halde vermek üzere oldukları kararın hiçbir dayanağı olmadığını itiraf etmiş olurlar. Zincirdeki herkesin sayıyla ilgisi var, ama kimsenin paydayla ilgisi yok.

Novara Consulting Group's position is that these documents should be read as advertising claims rather than as evidence, that the specific figures most often cited are the ones most easily produced without any underlying improvement, and that a buyer capable of asking four questions can distinguish a real result from a reported one in about ten minutes. What follows is that method, applied to the four claims that appear most frequently.

Zincirdeki herkesin sayıyla ilgisi var, ama kimsenin paydayla ilgisi yok.

Birinci iddia: işe alım süresinde azalma

İşe alım süresi en yaygın bildirilen ölçüttür ve işe alım süreciyle ilgili hiçbir şeyi değiştirmeden en kolay oynanabilen ölçüttür.

Bu ölçüm, karşılaştırma dönemleri boyunca sabit tutulan tanımlı bir başlangıç noktası ve tanımlı bir bitiş noktası gerektirir. Uygulamada başlangıç noktası kuruluşa ve çoğu zaman talep formuna (requisition) göre değişir: bu, talebin onaylandığı tarih, ilanın yayınlandığı tarih, kaynak bulma sürecinin başladığı tarih veya işe alım yöneticisinin ihtiyacı ilk kez dile getirdiği tarih olabilir. Bu başlangıç noktasını daha ileri bir tarihe çekmek, kimsenin yaşadığı hiçbir şeyi kısaltmadan ölçülen aralığı kısaltır. İptal edilen veya geri çekilen talepleri hariç tutmak, en uzun süren rolleri hariç tutmak veya paydadaki rol karışımını değiştirmek de aynı etkiyi yaratır.

Karıştırıcı etken, etkiden daha büyüktür. İşe alım süresi işgücü piyasası koşullarına güçlü biçimde tepki verir ve günümüzdeki çoğu vaka çalışması için ilgili karşılaştırma dönemleri, akut işgücü kıtlığından oldukça gevşemiş koşullara geçişi kapsar. Hiçbir

şeyi değiştirmemiş bir kuruluş bile bu pencere boyunca iyileşme gösterirdi. Bu hareketi o dönemde uygulanan bir sisteme atfetmek bir karşı olgusal analiz gerektirir ve bu türden hiçbir vaka çalışması böyle bir analiz içermez.

Niteleyici ifadeler de dikkat gerektirir. Tüm roller yerine kilit roller için bildirilen iyileşme, bir alt kümeyi işaret eder; ve bu alt küme, sonuçlar bilindikten sonra seçildiyse, bu rakam sistemi değil seçimi tarif eder.

İkinci iddia: manuel işlemsel faaliyette azalma

Bu iddia, manuel faaliyete ilişkin bir temel ölçüm gerektirir ve neredeyse hiçbir insan kaynakları fonksiyonu buna sahip değildir.

Where a figure exists it has usually been derived one of three ways: from a vendor's implementation model, from a survey in which staff estimated their own time allocation, or from a count of tickets or transactions that changed definition when the new system was introduced. None of these supports a percentage comparison across time, because the instrument changed at the same moment as the intervention.

There is also a scope question the metric conceals. Transactional work that disappears from the human resources function frequently reappears elsewhere, most often with employees and line managers through self-service. A sixty percent reduction in transactions handled by the function may represent a genuine efficiency gain, a transfer of the same work to people whose time is not being counted, or some combination. Distinguishing these requires measuring total organizational effort, which nobody does, because the metric exists to justify the function's investment rather than to describe the organization's cost.

Üçüncü iddia: yüksek potansiyelli çalışanların elde tutulmasında artış

Bu dördü arasında en ilgi çekici olanıdır, çünkü rakamı üreten mekanizma genellikle uygulamanın kendisidir.

Yüksek potansiyel gözlemlenen bir özellik değildir. Kuruluşun atadığı bir tanımlamadır ve bir yapay zeka yetenek sistemi, tipik olarak bu atamanın nasıl yapıldığını değiştirir. Belirtilen amaç da zaten budur. Sistem hangi çalışanların yüksek potansiyelli olarak tanımlandığını değiştiriyorsa, uygulamadan sonra ölçülen kohort,

öncekiyle aynı nüfus değildir ve ikisi arasındaki elde tutma karşılaştırması aynı şeyin karşılaştırması değildir.

Ortaya çıkan önyargının yönü tahmin edilebilirdir. Performans geçmişinden, bağlılık sinyallerinden ve kariyer ivmesinden yüksek potansiyeli belirleyen bir model, bağlı, yakın zamanda terfi etmiş ve ilerleme kaydeden çalışanları tercihen seçecektir; bunların hepsi kalıcılığın mevcut en güçlü göstergeleri arasındadır. Sistem zaten kalacak olan kişileri seçer ve kuruluş onların kalmasını sistemin bir etkisi olarak kaydeder.

Kohort büyüklüğü bunu daha da katmerlendirir. Çoğu kuruluşta yüksek potansiyelli nüfus, yüzde on sekizlik bir değişimin bir avuç kişiyi temsil ettiği, herhangi bir küçük kohortun yıldan yıla üreteceği varyasyonun tamamen içinde kaldığı kadar küçüktür. Bu türden bir rakamın yanında kohort büyüklüğü ve tanımın değişip değişmediği belirtilmelidir. Neredeyse hiçbir zaman belirtilmez.

Dördüncü iddia: önyargının azaltılması

Dördüncü iddia türü bakımından diğerlerinden farklıdır, çünkü sadece dayanaksız değildir. Bu bir hukuki beyandır.

Bunu doğrulamak, tanımlı gruplar üzerinde, karşılaştırmayı destekleyecek yeterli hücre büyüklükleriyle, öncesi ve sonrası önyargının ölçülmesini gerektirir. Ayrıca grupların bilinmesini de gerektirir ki iddia genellikle burada çöker. Engellilik bu şekilde hiç ölçülemez: işverenler, koşullu bir teklif öncesinde hangi başvuru sahiplerinin engelli olduğunu bilmez ve yasal olarak soramaz, dolayısıyla hiçbir etki-oranı aracı engellilik ayrımcılığını tespit edemez. Ardıllık planlamasında önyargının azaltıldığını iddia eden bir kuruluş, genellikle ırk ve cinsiyeti ölçmüştür (eğer bir şey ölçmüşse) ve performans verisi güdümlü modellerden en çok etkilenen nüfus hakkında hiçbir şey söylememiştir.

Ardıllık planlamasının ayrıca bir zorluğu daha vardır. Kıdemden, kariyer aşamasından, görevde geçirilen süreden, yeterlilik güncelliğinden ve önceki ilerleme hızından hazırlık veya potansiyeli belirleyen bir model, yaşla ağır biçimde ilişkili sinyallere dayanmaktadır. Böyle bir sistem için yaş dağılımını incelemeden önyargının azaltıldığını bildirmek, girdilerin öngördüğünün tam tersini iddia etmek demektir.

Ve iddianın kendisi bir hukuki risk yaratır. 1991'de değiştirildiği haliyle Title VII, istihdam veya terfi için aday seçimi veya yönlendirmesiyle bağlantılı olarak, ırk, renk,

din, cinsiyet veya ulusal köken temelinde puanları ayarlamayı, farklı kesme puanları kullanmayı veya istihdamla ilgili testlerin sonuçlarını başka şekilde değiştirmeyi yasa dışı kılar. Burada hiçbir şey bu sistemlerin bunu yaptığını göstermiyor. Önemli olan nokta şudur: otomatik bir terfi aracını demografik sonuçlarla ilişkilendiren kamuya açık bir beyan, kuruluşun kendisinin yazdığı, geri çekemeyeceği ve açıklaması istenebilecek bir belgedir.

Dördünde de eksik olan şey

Yukarıdaki başarısızlıkların her biri aynı eksikliğin bir versiyonudur. Bu vaka çalışmalarının hiçbiri bir karşı olgusal analiz içermez; yani müdahale olmasaydı ne olacağının bir tarifini içermez.

Kuruluş, işgücü piyasasının değiştiği, kadro sayısının değiştiği, yöneticilerin devir yaşadığı, ücretlerin ayarlandığı ve başka girişimlerin eşzamanlı yürütüldüğü bir dönemde bir sistem devreye almıştır. Ardından dört ölçütteki hareketi bu değişikliklerden birine atfetmiştir. Bu küçük bir metodolojik itiraz değildir. Bu meselenin bütünüdür ve bu eksiklik, bu türü çürütülemez kılan şeydir.

Bitişik iddialara ilişkin yayınlanmış kanıtlar, herhangi bir alıcıyı düşündürmelidir. Harvard Business School ve Burning Glass Institute, beceri temelli işe almayı kamuya açık biçimde benimseyen firmalarda gerçekte ne olduğunu incelediğinde, diploma şartının kaldırılmasının diplomasız adayların işe alımını yaklaşık üç buçuk yüzde puanı hareket ettirdiğini, bunun her yedi yüz işe almadan birinden azına denk geldiğini ve firmaların neredeyse yarısının açıklamalarına rağmen hiçbir davranış değişikliği göstermediğini tespit etti. Açıklamalar samimiydi. Uygulama tamamen farklılaştı ve bunu yapan kuruluşlar bunu bilmiyordu.

Buradaki ilgili çıkarım insanların yalan söylediği değildir. Kuruluşların kendi girişimlerinin işe yarayıp yaramadığını güvenilir biçimde bilmediği ve bir vaka çalışmasının bunu öğrenmeye en az konumlu tarafça yazıldığıdır.

İddia sahibi bunu genellikle neden inanılır bulur

Bunu doğrudan ifade etmekte fayda var, çünkü aksi okuma iyi niyetten uzak ve genellikle yanlıştır.

Bu rakamlar nadiren uydurulur. Bunlar, hiçbir temel ölçümün oluşturulmadığı, tanımların değiştiği, analitik çerçeveyi satıcının sağladığı ve sonuçları derleyen kişinin

satın almayı savunan kişi olduğu bir ölçüm ortamında üretilir. Bu koşulların her biri çıktıyı aynı yöne doğru eğer ve hiçbiri kimsenin kötü niyetle hareket etmesini gerektirmez. Sayı, birinin onu hesapladığı anlamında gerçektir. Sadece etrafındaki cümlenin söylediği şeyi ifade etmez.

Düzenleyici ortamın harekete geçmeye başlamasının nedeni de budur. Federal Ticaret Komisyonu, desteklenmeyen yapay zeka yetenek iddialarına karşı aktif olmuştur; bunlar arasında, Ocak 2025'te bir erişilebilirlik yazılımı sağlayıcısının, otomatik ürününün herhangi bir web sitesini erişilebilirlik standartlarına uygun hale getirebileceği yönündeki beyanları nedeniyle bir milyon dolar ödemesini gerektiren bir karar da bulunmaktadır. İlgili ilke o ürün kategorisiyle sınırlı değildir. Yapay zeka hakkında performans iddiaları doğrulama gerektirir ve bir pazarlama belgesinde sonuç yüzdeleri yayınlayan bir kuruluş tam da bu türden bir iddiada bulunmaktadır.

Alıcının sorması gereken sorular

Yaklaşık on dakika süren ve vaka çalışmalarının büyük çoğunluğunun karşısında ayakta kalamayacağı dört soru.

Temel ölçüm neydi, ne zaman ölçüldü ve kim tarafından ölçüldü. Temel ölçüm uygulamadan sonra oluşturulduysa veya satıcı tarafından sağlandıysa, karşılaştırma bir karşılaştırma değildir.

Tanım değişti mi. Yüksek potansiyelli veya kilit roller gibi belirlenmiş bir nüfusu içeren herhangi bir ölçüt için, atama kriterlerinin öncesinde ve sonrasında aynı olup olmadığını sorun. Sistem bunları değiştirdiyse, nüfuslar farklıdır ve karşılaştırma geçersizdir.

Bu dönemde başka ne değişti. İşgücü piyasası, kadro sayısı, ücretlendirme, yönetim, eşzamanlı girişimler. Dürüst bir yanıtlayanın bir listesi olacaktır. Listesi olmayan bir yanıtlayan bakmamıştır demektir.

Payda nedir ve kaç kişi. Kohort büyüklükleri olmadan yüzdeler yorumlanamaz ve küçük kohortlar küçük hareketlerden büyük yüzdeler üretir.

BEŞİNCİ BİR SORU

Yanlılığa değinen her iddiayla ilgili sorulması gereken beşinci bir soru şudur: hangi gruplar, hangi hücre büyüklüklerinde ölçüldü ve ölçülemeyen gruplar için ne yapıldı. Son sorunun yanıtı neredeyse her zaman hiçbir şey yapılmadığıdır ve alıcının öğreneceği en yararlı bilgi de zaten budur.

İnandırıcı bir vaka çalışması neler içerirdi

Bu tür çalışmalar onarılamaz değildir ve aşağıdakileri yayımlamaya istekli bir kuruluş, kimsenin bunu yapmadığı bir alanda hemen fark edilirdi.

Başlangıç değerini ve tarihini belirtin. Tanımları belirtin ve değişenleri not edin. Kohort büyüklüklerini yüzdelerle birlikte raporlayın. Hareketi açıklayabilecek eşzamanlı değişiklikleri adlandırın ve hangilerini dışlayamayacağınızı söyleyin. Neyin işe yaramadığını anlatın, çünkü her şeyin iyileştiği bir uygulama, hiçbir şeyin ölçülmediği bir uygulamadır. Ve bir yanlılık iddiasında bulunulduğunda, grupları, yöntemi, hücre büyüklüklerini ve ölçülmemiş olanı belirtin.

Bu belge, mevcut halinden daha az etkileyicidir ama çok daha kullanışlıdır. Ayrıca doğru olup olmadığıyla ilgilenen biri tarafından okunduğunda ayakta kalabilen tek versiyondur.

Sonuç

Vaka çalışması, kurumsal insan kaynakları teknolojisi satın alımlarında başlıca kanıt belgesidir ve rutin profesyonel kullanımdaki en zayıf kanıt biçimidir. Tek kuruluşa dayalıdır, kontrolsüzdür, kendi kendine bildirilmiştir, geriye dönük olarak tanımlanmıştır ve bulguda çıkarı olan bir taraf tarafından yayımlanmıştır.

None of that makes it worthless. It makes it a hypothesis. The organizations that will actually realize returns from these systems are the ones that treat other people's numbers as a reason to investigate rather than a reason to buy, and that hold their own numbers to the standard they would want applied to a vendor's. The discipline is not difficult and it is not expensive. It is a baseline, a stable definition, a cohort size, and an honest account of what else was happening at the time.

Kaynakça

- Federal Ticaret Komisyonu, *accessiBe, Inc. Davası*, Dosya No. 2223156, şikayet ve önerilen karar 3 Ocak 2025'te açıklanmıştır.
- Fuller, Joseph, ve diğerleri. *Beceri Bazlı İşe Alım: Söylemlerden Uygulamaya Uzun Yol*. Harvard Business School, Çalışmanın Geleceğini Yönetme Projesi ve The Burning Glass Institute, Şubat 2024.
- 1964 Medeni Haklar Yasası'nın, değiştirilmiş haliyle, VII. Başlığı, 42 U.S.C. § 2000e-2, § 2000e-2(l) dahil.
- Çalışan Seçim Prosedürlerine İlişkin Tekdüzen Kılavuzlar, 29 C.F.R. Bölüm 1607, işe alım, terfi ve diğer istihdam fırsatlarını etkileyen seçim prosedürlerine uygulanır.

TİPS

Know something about a sign language AI system that the public should know? Send it to tips@novaracg.com. Every tip is read by me and kept confidential. If you work in the UK, writing to us is not by itself a protected disclosure under the Public Interest Disclosure Act 1998, so before you share employer information, check your options with Protect (protect-advice.org.uk), the free whistleblowing advice line.

Nasıl alıntılanır

Grizzle, H. M. (2025, May 23). Dört Sayı ve Sıfır Karşı Olgusal Analiz: İnsan Kaynaklarında Yapay Zeka Vaka Çalışmasını Okumak. Novara Consulting Group. https://www.novaracg.com/tr/2025/05/23/four-numbers-and-no-counterfactual-reading-the-artificial-intelligence-case-study-in-human-resources/