MEMPEKERJAKAN AGEN AI DI PERUSAHAAN ANDA

“Great vision without great people is irrelevant.”

“Visi yang hebat tanpa orang yang hebat tidak ada artinya.”

— Jim Collins, *Good to Great* (2001)[278,279]

> **Pertanyaan bab ini:** Bagaimana memasang agen AI di perusahaan tanpa kehilangan kendali atas pekerjaan, uang, dan pelanggan?
>
> **Jawaban singkat:** Perlakukan agen seperti karyawan baru: periksa dulu apakah pekerjaannya memang perlu agen, tulis uraian jabatannya, tetapkan satu atasan bernama, naikkan wewenangnya hanya sesudah ada bukti, nilai kinerjanya dengan kelompok pembanding, dan tulis syarat pemberhentiannya sejak hari pertama.
>
> **Alat yang Anda bawa pulang:** uraian jabatan agen dalam tujuh baris, lima ukuran rapor agen, dan lima alasan memberhentikan agen.

Identitas Dokumen

BUKU	**Stay Human. Stay in Charge.** Transformasi AI di Korporasi: Maksimalkan Hasil Tanpa Halusinasi dan Kebocoran Data
PENULIS	Dian Martin, Ketua Umum Asosiasi AI Indonesia
BAGIAN	Bagian Keenam: Tata Kelola AI
BAB	Bab 16: Mempekerjakan Agen AI di Perusahaan Anda
HALAMAN	425–454 (nomor halaman cetak buku)
ISI BERKAS	Isi bab, Panduan Daftar Pustaka, dan Daftar Pustaka bab ini

KAMIS, 15 AGUSTUS 2024. Sebanyak 302 pekerja layanan pelanggan di Taobao, platform belanja daring milik Alibaba, mulai bekerja dengan rekan baru. Rekan itu agen AI yang dikembangkan Alibaba pada Mei 2024, dengan model bahasa besar sebagai mesin keputusannya. Sementara itu, 345 pekerja lain bekerja seperti biasa, tanpa agen, sebagai kelompok pembanding. Kedua kelompok itu berasal dari 647 pekerja yang dipilih secara acak untuk sebuah uji lapangan, dan pekerjaan mereka selama dua minggu sebelumnya, dari 1 sampai 14 Agustus, sudah dicatat sebagai pembanding awal.[146]

Agen itu tidak diberi semua pekerjaan. Ia hanya menangani percakapan tentang masalah standar, yang bisa diselesaikan dengan mengikuti prosedur baku tanpa campur tangan langsung manusia. Para peneliti menyebutnya percakapan yang memenuhi syarat ditangani agen, dan jumlahnya kurang dari 10 persen seluruh percakapan. Sisanya tetap dikerjakan manusia. Setiap percakapan yang memenuhi syarat itu ditugaskan kepada satu pekerja yang mengawasinya. Begitu pekerja itu turun tangan, ia memegang tanggung jawab penuh sampai percakapan selesai.

Agen juga tidak bekerja tanpa alarm. Satu algoritma memantau tanda pelanggan mulai kesal atau ragu, dan percakapan seperti itu dialihkan ke manusia; para peneliti menyebutnya pengalihan karena pelanggan kesal. Algoritma lain memantau pertanyaan yang bergeser ke luar kemampuan agen, dan percakapan itu pun dialihkan: pengalihan karena di luar kemampuan agen. Di luar kedua alarm itu, pengawas boleh mengambil alih atas penilaiannya sendiri.[146]

Agen itu bekerja selama 17 hari, sampai Sabtu, 31 Agustus 2024. Selama masa studi, tercatat 680.676 percakapan.

Seandainya Anda yang harus memutuskan pada hari terakhir uji coba itu: sistem ini diteruskan ke semua pekerja, atau dihentikan? Angka apa yang Anda minta diletakkan di meja?

Angka pertama yang biasanya diminta adalah kecepatan. Jawabannya menyenangkan: secara keseluruhan, waktu layanan per percakapan turun 3,2 persen. Pertanyaan kedua semestinya menyusul: apakah pelanggan dirugikan? Secara keseluruhan, tidak. Nilai dari pelanggan, yaitu rating pada skala satu sampai lima, tidak berubah secara bermakna. Begitu pula kontak ulang, yaitu pelanggan yang menghubungi lagi untuk masalah yang sama dalam tujuh hari.

Kalau rapat berhenti di dua angka itu, keputusannya mudah. Namun ada pertanyaan ketiga yang jarang diajukan: bagaimana hasilnya bila angka itu dipecah menurut jenis percakapan? Di sinilah ceritanya bercabang. Pada percakapan yang memenuhi syarat ditangani agen, waktu layanannya turun 16,8 persen, tetapi nilai dari pelanggan turun 0,412 poin dibanding kelompok pembanding, hampir setengah bintang. Pada percakapan yang tetap dikerjakan manusia, waktu layanannya turun 1,8 persen dan nilainya justru naik 0,091 poin. Menurut para peneliti, pekerja yang bekerja bersama agen mengalihkan perhatiannya ke percakapan yang memang membutuhkan penilaian manusia.[146]

Jadi agen mempercepat bagian pekerjaan yang diserahkan kepadanya, tetapi kepuasan pelanggan pada bagian itu menurun. Sementara itu, manusia yang terlepas dari sebagian pekerjaan rutin justru melayani lebih baik. Rata-rata keseluruhan menyembunyikan keduanya.

Pertanyaan ketiga itulah yang membedakan rapat yang memutuskan dari rapat yang hanya mengesahkan. Angka rata-rata menjawab apakah sebuah proyek layak diteruskan. Angka yang dipecah menjawab

bagaimana proyek itu harus dijalankan: pekerjaan mana yang diserahkan ke agen, siapa yang mengawasinya, dan kapan manusia harus masuk.

Makalah itu tidak menceritakan rapat keputusannya. Yang tercatat hanya ini: uji coba berakhir Sabtu, 31 Agustus 2024, dan mulai keesokan harinya, Minggu, 1 September, sistem itu dipakai semua pekerja.[146]

Bagi direksi di Indonesia, kisah ini berguna karena cara Alibaba memperlakukan agennya. Pekerjaannya dibatasi, setiap percakapan yang dipegangnya punya satu pengawas manusia, dan kinerjanya diukur dengan kelompok pembanding, termasuk hasil yang tidak manis. Begitulah perusahaan yang baik memperlakukan karyawan baru. Para konsultan yang menelaah lebih dari 50 proyek agen yang mereka pimpin sampai pada kesimpulan yang sama. Seorang pemimpin bisnis yang mereka kutip menyebut bahwa memasang agen lebih mirip merekrut karyawan baru daripada memasang perangkat lunak, dan para konsultan itu menyarankan agar setiap agen diberi uraian jabatan yang jelas, dikenalkan ke pekerjaannya, dan terus diberi umpan balik.[280]

Bagi perusahaan Indonesia yang baru mulai, cara pandang ini punya keuntungan praktis: hampir semua alatnya sudah dikenal di kantor. Lowongan, uraian jabatan, masa percobaan, rapor, dan surat pemberhentian adalah dokumen yang sudah biasa diurus bagian SDM. Yang baru hanya pegawainya.

Pembahasan sebelumnya ditutup dengan kalimat bahwa setiap agen perlu satu nama manusia sebagai atasannya. Di Taobao, nama itu ada di setiap percakapan. Bab ini meminjam cara pandang yang sama dan mengikutinya dari awal sampai akhir: seleksi, uraian jabatan, masa percobaan, atasan langsung, buku induk, rapor, sampai pemberhentian.

Seleksi: Apakah Pekerjaan Ini Perlu Agen?

Perusahaan yang sehat tidak merekrut sarjana untuk menekan tombol mesin fotokopi. Sarjana itu tentu bisa melakukannya, tetapi gajinya terlalu mahal untuk pekerjaan itu, dan kemampuannya terbuang. Agen AI diperlakukan sama. Sebelum lowongan dibuka, ada satu pertanyaan yang sering dilompati: apakah pekerjaan ini memang perlu dikerjakan agen?

Pertanyaan itu penting karena label "agen" sedang laris. Pada Rabu, 25 Juni 2025, firma riset Gartner memperkirakan lebih dari 40 persen proyek AI agentik akan dibatalkan sebelum akhir 2027, karena biaya yang membengkak, nilai bisnis yang tidak jelas, atau kendali risiko yang tidak memadai. Dalam rilis yang sama, Gartner menyoroti gejala yang disebutnya *agent washing*: produk lama, seperti asisten AI, otomasi proses robotik, dan *chatbot*, diberi label baru sebagai agen. Dari ribuan vendor yang menjual AI agentik, menurut perkiraan yang sama, hanya sekitar 130 yang sungguhan.[281]

Label seperti itu bisa menjadi *PHP*, singkatan pemberi harapan palsu yang biasa dipakai untuk janji yang tidak ditepati: brosurnya menjanjikan agen yang bekerja sendiri, isinya alat lama dengan nama baru. Firma riset yang sama juga mencatat bahwa banyak pemakaian yang ditawarkan sebagai agen sebenarnya tidak memerlukan agen. Sarannya untuk tahap awal lugas: pakai agen ketika diperlukan keputusan, otomasi untuk alur kerja rutin, dan asisten untuk sekadar mencari informasi; kejar AI agentik hanya bila nilainya atau imbal hasilnya jelas.[281]

Cara memeriksanya tidak rumit. Minta vendor memperlihatkan, pada data perusahaan Anda sendiri, tindakan apa yang dikerjakan agennya tanpa diarahkan langkah demi langkah, dan apa yang terjadi bila agen itu

menemui kasus yang tidak ada di prosedur. Bila jawabannya hanya berupa demo dengan contoh yang sudah disiapkan, yang ditawarkan kemungkinan besar alat lama dengan nama baru.

Para konsultan yang menelaah lebih dari 50 proyek agen tadi memberi patokan untuk pekerjaan rutin: bila tugasnya berbasis aturan dan berulang, dengan input yang terstruktur seperti entri data, pakai otomasi berbasis aturan.[280] Agen otonom untuk pekerjaan satu langkah itu salah kamar: ia ditempatkan di pekerjaan yang tidak cocok untuknya, lebih mahal dan lebih sulit diawasi daripada aturan otomatis biasa yang sebenarnya sudah cukup.

Lima Level Teknologi di Bagian Pembahasan "Desain Ulang: Benahi Dapurnya Dulu" sudah mengingatkan bahwa banyak pekerjaan cukup dibereskan dengan merapikan langkah prosesnya, atau dengan aturan otomatis biasa. Bila pekerjaannya hanya satu langkah, atau manusia tetap harus mengerjakannya di sistemnya sendiri, agen tidak diperlukan.

Untuk pekerjaan selebihnya, saya memetakan pilihannya seperti ini, dari yang paling sederhana.

Empat Pilihan Sebelum Merekrut Agen

Pilihan	Cocok bila	Jangan dipakai bila
Otomasi berbasis aturan	Langkah dan aturannya sudah diketahui, input terstruktur, kesalahan tidak bisa ditoleransi	Input tidak terstruktur dan sangat beragam
Asisten AI yang memberi saran	Menyusun draf atau analisis, manusia menjadi penjaga mutu	Hasilnya langsung dikirim ke pelanggan tanpa diperiksa
Satu agen	Banyak langkah lintas sistem, volume tinggi, tindakan bisa dibatalkan, ada jalur ke manusia	Tindakannya berisiko tinggi dan otonominya tinggi
Beberapa agen bekerja sama	Satu agen sudah stabil, tugasnya benar-benar bisa dipecah, ada pemeriksa hasil	Supaya terlihat canggih, atau tanpa pemeriksa

Perhatikan baris terakhir. Beberapa agen yang bekerja sama baru layak sesudah satu agen terbukti stabil. Bagian Pembahasan "Kala AI Mulai Bertindak" sudah memperlihatkan bagaimana tim agen sering gagal karena tidak jelas siapa memegang langkah yang mana.

Contohnya di bagian administrasi penjualan sebuah distributor. Mencocokkan nomor pesanan dengan nomor faktur cukup dikerjakan aturan otomatis, karena langkahnya pasti dan inputnya terstruktur. Menyusun draf jawaban untuk toko yang menanyakan status kiriman cocok untuk asisten yang memberi saran, karena staf masih membaca setiap jawaban sebelum dikirim. Menerima pesanan lewat pesan singkat yang ditulis dengan gaya santai, memeriksa stok di gudang, lalu menyiapkan surat jalan adalah pekerjaan banyak langkah lintas sistem. Di situlah satu agen mulai masuk akal, asalkan setiap surat jalan masih

bisa dibatalkan sebelum barang dimuat. Ketiga pekerjaan itu ada di satu meja, tetapi jawabannya berbeda.

Pekerjaan yang Cocok untuk Agen

Bila jawabannya memang perlu agen, lowongan baru boleh ditulis. Namun sebelum menulis uraian jabatan, pastikan dulu pekerjaannya cocok. Saya mencari enam ciri.

Ciri pertama dan kedua menyangkut bentuk pekerjaannya. Pekerjaan itu bervolume besar dan berulang: muncul setiap hari atau setiap ada transaksi. Agen yang hanya dipakai sebulan sekali tidak memberi cukup banyak kasus untuk diuji dan diperbaiki, dan tidak sempat menghasilkan manfaat yang menutup biayanya. Pekerjaan itu juga punya input yang beragam dan butuh sedikit penilaian; di situlah bedanya dengan otomasi biasa. Taobao memberi contohnya. Masalah pelanggan yang diserahkan ke agen adalah masalah standar yang penyelesaiannya mengikuti prosedur baku, tetapi pelanggan menyampaikannya lewat percakapan bebas, bukan lewat formulir.

Ciri ketiga dan keempat menyangkut risiko dan uang. Jalurnya sempit dan tindakannya bisa dibatalkan. Agen menghasilkan nilai pada pekerjaan yang cakupannya terbatas, volumenya tinggi, dan tindakannya bisa dibatalkan. Di luar itu, keandalannya pada rangkaian langkah yang panjang masih rendah, dan biayanya sulit ditebak. Dampaknya pun harus terukur dalam uang, bukan sekadar jam yang dihemat. Jam yang dihemat agen dihitung nol sampai ada cara tertulis untuk mengubahnya menjadi uang, aturan yang kita pakai sejak Bagian Pembahasan "Penghematan di Atas Kertas".

Ciri kelima dan keenam menyangkut kesiapan. Tuntutan ketepatannya wajar. Makin tinggi ketepatan yang diwajibkan, misalnya untuk teks

perjanjian, laporan kepada regulator, atau saran kepada nasabah, makin mahal biaya membuat hasil agen layak dipakai, dan makin kecil imbal hasilnya. Terakhir, data dan sistemnya siap. Pilih proses yang datanya sudah digital dan sudah berjalan di sistem kerja perusahaan, misalnya sistem pencatatan keuangan dan operasi (ERP) atau sistem pengelolaan pelanggan (CRM). Proses yang masih bergantung pada kertas dan ingatan orang sebaiknya dirapikan dulu.

Bila diringkas, agen cocok untuk pekerjaan yang sering terjadi, jalurnya sempit, tindakannya bisa dibatalkan, dan hasilnya bisa dihitung dalam rupiah.

Ada satu pertanyaan lagi yang jarang diajukan saat lowongan ditulis: otak sebesar apa yang dibutuhkan pekerjaan itu? Agen untuk pekerjaan rutin bervolume besar tidak perlu otak terbesar. Agen Taobao hanya menangani masalah standar, kurang dari sepersepuluh percakapan. Sekelompok peneliti industri berpendapat bahwa sebagian besar pekerjaan di dalam sistem agen memang berupa tugas sempit yang diulang-ulang dengan sedikit variasi, sehingga model bahasa kecil sudah cukup kuat, lebih cocok, dan lebih hemat untuk banyak panggilan.[282] Untuk tugas mengelompokkan teks, sebuah studi menemukan bahwa model kecil yang dilatih ulang dengan contoh dari tugas itu sendiri secara konsisten mengungguli model besar yang hanya diberi perintah.[283] Di kantor, tugas seperti itu misalnya memilah *email* masuk, menggolongkan keluhan, atau menandai mutasi rekening.

Agen juga berbeda dari asisten yang hanya menjawab pertanyaan: ia membaca data pelanggan, menulis ke sistem perusahaan, dan menyimpan catatan kerjanya. Makin banyak data yang lewat di tangannya, makin penting pertanyaan di mana ia dijalankan. Saran saya, untuk agen pekerjaan rutin seperti itu, pertimbangkan model kecil yang

bisa dijalankan di server sendiri, di server bersama beberapa perusahaan, atau di mitra pusat data dalam negeri. Pilihan ini sekaligus menjaga data pelanggan tetap dalam kendali perusahaan. Cara memilihnya memakai lembar harga kedaulatan di Bagian Pembahasan "Benarkah Server Sendiri Lebih Murah?", dan kapan server kecil seperti itu menguntungkan dibahas di Bagian Pembahasan "Aturan Main AI Kita".

Uraian Jabatan Agen dalam Tujuh Baris

Sesudah pekerjaannya cocok, tulis uraian jabatannya di selembar kertas, seperti untuk karyawan baru. Uraian itu tidak perlu panjang. Tujuh baris cukup, asalkan setiap baris terisi.

1. Nama jabatan dan tugas utamanya. Satu kalimat: apa yang dikerjakan agen, untuk siapa, dan apa hasilnya.

2. Hal-hal yang tidak boleh ia lakukan. Termasuk data apa yang tidak boleh dibawanya keluar, dan layanan mana saja yang boleh disambungkan kepadanya. Model, data, dan catatan kerja agen tercatat sebagai milik perusahaan, siapa pun yang menjalankannya.

3. Siapa atasan langsungnya. Satu nama manusia, bukan nama tim atau nama bagian.

4. Ukuran kinerjanya, beserta angka dasarnya. Angka dasar dicatat sebelum agen dipasang, supaya ada yang bisa dibandingkan.

5. Batas biaya per bulan. Plafon dalam rupiah, dan siapa yang dihubungi bila plafon hampir habis. Gunanya agar tagihan yang tiba-tiba melonjak langsung dibicarakan pada hari itu juga, tidak menunggu laporan akhir bulan.

6. Kapan ia harus menyerahkan pekerjaan ke manusia. Jenis kasus yang langsung dialihkan, dan seberapa cepat pengalihannya.

7. Syarat kapan ia diberhentikan. Ditulis sekarang, sebelum ada masalah.

Siapa yang menulisnya? Saran saya, tiga orang duduk bersama: atasan langsung agen, yang paling tahu pekerjaannya; pemegang hasil dari bagian yang memakai agen, yang memegang angkanya; dan orang TI yang menjalankan agen, yang tahu batas teknisnya. Vendor boleh dimintai pendapat, tetapi tidak menulisnya, karena uraian jabatan dari vendor cenderung menggambarkan apa yang bisa dilakukan produknya, bukan apa yang dibutuhkan perusahaan. Uraian itu ditinjau setiap kali agen naik atau turun level, setiap kali modelnya diganti, dan sedikitnya sekali setahun.

Agen yang dipasang dengan uraian jabatan setengah terisi ibarat masakan setengah matang yang sudah disajikan ke tamu: kelihatannya siap, tetapi yang menanggung akibatnya adalah pelanggan. Bila salah satu dari tujuh baris ini tidak bisa diisi, lowongannya belum siap dibuka.

Masa Percobaan: Naik Level dengan Bukti

Karyawan baru tidak langsung diberi kunci brankas. Ia menjalani masa percobaan, dan wewenangnya dinaikkan sesudah ia menunjukkan hasil. Agen diperlakukan sama.

Di Bagian Pembahasan "Kala AI Mulai Bertindak" kita sudah mengenal enam level otonomi, A0 sampai A5, dari sekadar memberi informasi sampai menentukan sendiri langkah dan alatnya. Bagian Pembahasan "Kala AI Mulai Bertindak" juga menetapkan prinsipnya: otonomi ditetapkan per jenis tindakan, bukan per agen. Di masa percobaan,

prinsip itu menjadi jadwal kerja. Ambil contoh agen di bagian pengadaan. Membandingkan tiga penawaran pemasok boleh ia kerjakan sendiri sejak minggu pertama, karena tidak mengubah apa pun. Menyiapkan pesanan pembelian boleh, tetapi pesanan itu baru terkirim sesudah disetujui staf. Menandatangani kontrak tidak boleh sama sekali. Ketiganya dikerjakan agen yang sama, dengan level otonomi yang berbeda.

Kapan sebuah tindakan boleh naik level? Aturannya sudah kita bahas di Bagian Pembahasan "Keamanan Dimulai dari Pembatasan Akses": lamanya agen diamati dan banyaknya tindakan yang tercatat, dua-duanya harus terpenuhi, ditambah uji serangan yang lulus dan cara pemulihan yang sudah diuji. Level tertinggi, A5, tidak saya sarankan untuk perusahaan pada 2026. Pegangan saya: otonomi secukupnya. Tujuan boleh tinggi, tetapi setiap kenaikan level ditebus dengan bukti.

Masa percobaan juga punya tanggal akhir, seperti masa percobaan karyawan. Pada tanggal itu ada tiga kemungkinan keputusan. Agen lulus, dan tindakan tertentu naik satu level. Masa percobaannya diperpanjang dengan perbaikan yang tertulis. Atau agen diberhentikan, dengan alasan yang sudah ditulis sejak awal. Yang tidak boleh terjadi adalah masa percobaan yang berakhir tanpa keputusan, sehingga agen terus bekerja di level yang sama hanya karena tidak ada yang sempat menilainya. Tanggal itu ditulis di lembar syarat berhenti yang kita bahas di pembahasan sebelumnya.

Mengapa harus berhati-hati? Hitungan peluang di Bagian Pembahasan "Kala AI Mulai Bertindak" sudah memperlihatkan bagaimana kesalahan kecil di setiap langkah menumpuk pada rangkaian yang panjang. Dari situ saya menarik pelajaran untuk masa percobaan: pecah alur kerja yang

panjang, pasang pemeriksaan di antara langkah, dan batasi otonomi pada rangkaian yang pendek.

Dari lebih dari 50 proyek itu, para konsultan yang sama mencatat satu kebiasaan yang membuat pemeriksaan seperti itu sulit: banyak perusahaan hanya melacak hasil akhir. Saran mereka, kinerja agen diperiksa di setiap langkah alur kerja. Dalam salah satu proyek yang mereka ceritakan, akurasi agen turun tiba-tiba ketika sistem bertemu sekelompok kasus baru, dan penelusuran per langkah menunjukkan bahwa sebagian kelompok pengguna mengirim data bermutu lebih rendah, sehingga agen salah menafsirkannya.[280] Tanpa catatan per langkah, tim hanya tahu bahwa hasilnya memburuk, tanpa tahu di mana harus memperbaikinya.

Mengalihkan ke Manusia Sebelum Pelanggan Kesal

Taobao memberi satu pelajaran lagi untuk masa percobaan, dan pelajaran ini soal waktu. Ketika percakapan dialihkan karena pertanyaan pelanggan bergeser ke luar kemampuan agen, campur tangan manusia berhasil menjaga mutu layanan. Ketika percakapan dialihkan karena algoritma mendeteksi pelanggan kesal, ceritanya lain. Pada percakapan seperti itu, waktu layanan menjadi 40,8 persen lebih lama, kontak ulang naik 6 poin persen, dan nilai dari pelanggan turun 0,928 poin.

Para peneliti juga menemukan sesuatu yang sangat manusiawi. Petugas yang mewarisi percakapan dengan pelanggan yang sudah kesal mengirim lebih sedikit pesan dan mengambil porsi percakapan yang lebih kecil. Campur tangan yang lebih awal menjaga upaya petugas tetap tinggi dan kekesalan pelanggan tetap pada tingkat sedang.[146]

Pelajarannya berlaku di kantor mana pun. Perusahaan yang menunggu pelanggan meledak sebelum memanggil manusia akan membayar

harganya dua kali: sekali di nilai dari pelanggan, dan sekali lagi di usaha petugas yang mewarisi percakapan itu.

Saran saya, alihkan ke manusia pada keluhan pertama, bukan sesudah keluhan berulang. Tanda keluhan pertama bisa sederhana: kata-kata kesal, permintaan bicara dengan orang, atau pertanyaan yang sama diulang. Dalam percakapan pelanggan di Indonesia, tanda itu sering datang dalam bahasa santai, misalnya "kak, ini kok dari kemarin belum beres juga?". Kalimat seperti itu perlu masuk ke kumpulan uji agen, supaya terlihat apakah agen mengenalinya sebagai keluhan. Aturan ini mengurangi sedikit penghematan agen, karena lebih banyak percakapan berpindah ke manusia. Namun yang berpindah hanya sebagian kecil percakapan, sedangkan yang dijaga adalah percakapan yang paling mahal bila gagal.

Satu Nama sebagai Atasan Agen

Setiap karyawan baru punya atasan. Agen juga harus punya, dan atasannya seorang manusia yang namanya bisa ditulis di buku induk.

Di Taobao, setiap percakapan yang memenuhi syarat ditangani agen ditugaskan kepada satu pekerja yang mengawasinya, dan begitu ia turun tangan, tanggung jawab penuh berpindah kepadanya. Gagasan satu nama seperti ini sudah lama dipakai di dunia usaha. Di Apple, menurut sebuah liputan majalah bisnis pada 2011, istilah internalnya DRI, singkatan *directly responsible individual*: satu orang yang namanya tertulis di samping setiap butir tindakan dalam rapat, supaya semua orang tahu siapa yang memegangnya.[284]

Untuk agen, atasan seperti itu memegang empat tugas. Ia menetapkan tujuan dan batas kerja agen, memberi konteks yang tidak tertulis di sistem, memeriksa hasil kerjanya secara berkala, dan memutuskan

kapan agen naik level, turun level, atau diberhentikan. Satu orang bisa menjadi atasan beberapa agen sekaligus, selama ia sanggup memeriksa hasil kerja semuanya.

Memeriksa hasil kerja tidak berarti membaca semua yang dikerjakan agen. Atasan cukup mengambil sampel kecil setiap minggu, ditambah semua kasus yang dialihkan ke manusia, lalu mencatat apa yang keliru dan mengapa. Catatan mingguan itu menjadi bahan rapor di akhir masa percobaan. Atasan yang tidak pernah membuka hasil kerja agennya sebenarnya hanya meminjamkan namanya. Atasan itu pula yang pertama kali menjawab bila pelanggan, auditor, atau regulator bertanya mengapa agen mengambil tindakan tertentu. Kisah chatbot Air Canada dalam cerita sebelumnya sudah memperlihatkan bahwa perusahaan tidak bisa berdalih chatbot-nya bertanggung jawab atas tindakannya sendiri.

Di pembahasan sebelumnya, satu nama memegang hasil sebuah pemakaian AI di tingkat perusahaan. Atasan agen bekerja lebih dekat ke lapangan. Di perusahaan kecil, keduanya bisa orang yang sama. Di perusahaan yang lebih besar, pemegang hasil biasanya seorang direktur atau kepala unit, sedangkan atasan agen adalah orang yang setiap hari melihat pekerjaan agen itu.

Yang berbahaya adalah agen yang tetap bekerja sesudah atasannya pindah bagian, naik jabatan, atau keluar dari perusahaan. Agen seperti itu bom waktu: hari ini tidak terjadi apa-apa, tetapi ketika ia keliru, tidak ada orang yang tahu apa tugasnya dan siapa yang berwenang menghentikannya. Saran saya, setiap pergantian atasan agen dicatat seperti serah terima pemegang use case di pembahasan sebelumnya, bersama uraian jabatan dan rapor terakhir agen itu.

Buku Induk Agen

Berapa agen yang bekerja di perusahaan Anda hari ini, dan siapa atasan masing-masing? Bila jawabannya butuh lebih dari sehari untuk dikumpulkan, perusahaan Anda belum punya buku induk agen.

Perusahaan yang rapi punya buku induk karyawan: siapa saja yang bekerja, di bagian apa, siapa atasannya, dan apa wewenangnya. Untuk agen, buku induk itu adalah registri agen yang kita bahas di Bagian Pembahasan "Keamanan Dimulai dari Pembatasan Akses" dari sisi keamanan. Setiap agen tercatat dengan pemiliknya, sponsornya, serta versi perintah, alat, dan model yang dipakainya, ditambah lisensi modelnya seperti yang diusulkan dalam cerita sebelumnya. Dari sisi kepegawaian, saya menambah tiga kolom: nama atasan langsung, tanggal uraian jabatan terakhir ditinjau, dan rapor terakhirnya. Agen yang tidak tercatat di buku induk tidak boleh bekerja.

Bagi perusahaan menengah yang belum punya registri, buku induk bisa dimulai dari satu lembar kerja dengan satu baris per agen. Mintalah setiap kepala bagian mendaftar agen yang dipakai di bagiannya, termasuk agen yang datang bersama aplikasi yang sudah dibeli, seperti agen bawaan aplikasi penjualan yang ditemukan PT Contoh di pembahasan itu. Agen bawaan seperti itu paling sering lolos dari daftar, karena tidak ada yang merasa merekrutnya.

Buku induk juga memperlihatkan pembagian tugas antaragen. Saran saya, agen yang merencanakan tidak sama dengan agen yang mengeksekusi tindakan berisiko tinggi. Bank Contoh sudah melakukannya dalam cerita sebelumnya, ketika agen pengembalian biaya dipecah menjadi agen pembaca yang hanya menggolongkan pesan nasabah dan agen pelaksana yang hanya menerima data kelayakan dari

sistem bank. Bila keduanya tercatat terpisah di buku induk, masing-masing punya atasan, batas, dan rapornya sendiri.

Pengawas Khusus atau Peran Gabungan?

Siapa yang sebaiknya mengawasi agen sehari-hari? Para peneliti Taobao membahas dua pilihan, dan pembahasan mereka terbuka tentang kelebihan serta kekurangan masing-masing.

Pilihan pertama, pengawas khusus: sekelompok orang yang tugasnya hanya mengawasi agen dan mengambil alih bila perlu. Menurut para peneliti, orang seperti itu bisa lebih cepat mengenali percakapan yang mulai memburuk. Namun mereka terus-menerus menghadapi pelanggan yang sudah kesal, sehingga beban emosinya berat dan risiko kelelahannya tinggi. Lama-lama, karena tidak lagi melayani pelanggan secara langsung, mereka juga bisa kehilangan pemahaman tentang masalah pelanggan, dan keterampilannya terkikis.

Pilihan kedua, peran gabungan: orang yang sama mengawasi agen sambil tetap melayani kasus yang tidak bisa ditangani agen. Inilah yang diuji di Taobao, dan hasilnya sudah kita lihat di awal bab: pekerjaan manusia justru membaik. Kelemahannya, menurut pembahasan yang sama, pengambilalihan percakapan yang memburuk bisa terlambat. Kesimpulan para peneliti: tidak satu pun dari kedua cara itu unggul di semua sisi.[146]

Pilihan ini adalah keputusan tentang manusia, bukan tentang teknologi: siapa yang mengawasi agen, bagaimana orang itu tetap tajam, dan siapa yang menjaga agar ia tidak kelelahan. Vendor tidak bisa memutuskannya. Yang memutuskan adalah kepala unit yang memakai agen, bersama bagian SDM.

Di perusahaan menengah, pilihan ini sering diputuskan oleh jumlah orang. Bagian layanan yang hanya punya beberapa petugas hampir pasti memakai peran gabungan, dan itu tidak masalah, asalkan tiga hal dijaga. Giliran mengawasi dibagi rata, supaya kasus sulit tidak menumpuk pada satu orang. Pengalihan karena pelanggan kesal ditangani petugas yang paling berpengalaman. Petugas yang lebih muda tetap mendapat kasus biasa untuk dilayani sendiri, supaya keterampilan melayaninya terus tumbuh, seperti yang kita bahas di Bagian Pembahasan "Mesin Jawaban atau Guru Privat?".

Satu pelajaran lain dari para konsultan itu sejalan dengan pembahasan para peneliti Taobao: manusia tetap diperlukan, tetapi peran dan jumlah orangnya akan berubah. Peran manusia bergeser: mereka mengawasi ketepatan, memastikan kepatuhan, menggunakan pertimbangan, dan menangani kasus-kasus di luar kebiasaan.[280] Soal jumlah orang itu kita bahas di pembahasan berikutnya.

Sketsa: Siapa yang Mengawasi Agen Pengembalian Biaya?

Pilihan antara pengawas khusus dan peran gabungan pernah diperdebatkan di Bank Contoh. Bank, tokoh, dan angka dalam sketsa ini rekaan, dan kesamaan dengan pihak nyata tidak disengaja. Yang nyata adalah temuan uji acak Taobao tentang dua jenis pengalihan, dan pembahasan para penelitinya tentang dua cara mengawasi agen.[146]

Di Bagian Pembahasan "Keamanan Dimulai dari Pembatasan Akses", agen pelaksana pengembalian biaya keterlambatan turun otomatis ke A2 ketika vendor mengganti modelnya. Uji ulangnya kemudian lulus, dan agen itu kembali ke A3: pengembalian kecil ke rekening asal nasabah boleh dieksekusinya sendiri, dalam batas yang ditetapkan sejak awal

cerita. Agen pembaca tetap menggolongkan pesan nasabah dan memutuskan mana yang dialihkan ke petugas. Datanya tetap diproses di mitra pusat data dalam negeri, sesuai keputusan Bagian Pembahasan "Benarkah Server Sendiri Lebih Murah?".

Pada akhir bulan pertama sesudah itu, kepala layanan nasabah membawa catatan pengalihan ke rapat. Misalkan dari sekitar 3.000 permohonan bulan itu, 360 dialihkan ke petugas, atau 12 dari setiap 100. Sebanyak 240 dialihkan karena permintaannya di luar aturan, misalnya keterlambatan yang bukan pertama kali. Sebanyak 120 dialihkan karena nasabahnya kesal.

Angka kedua itulah yang membuat rapat berlangsung lama. Di Bagian Pembahasan "Kala AI Mulai Bertindak", aturannya jelas: nasabah yang menulis dengan nada marah, atau menyebut kata keluhan, langsung diteruskan ke petugas berpengalaman. Dalam praktik, agen pembaca baru menandai nasabah sebagai kesal sesudah ia mengeluh untuk kedua atau ketiga kalinya. Misalkan rata-rata 14 menit berlalu dari keluhan pertama sampai petugas masuk ke percakapan. Pada 18 persen percakapan yang dialihkan karena nasabah kesal, nasabah menghubungi bank lagi untuk masalah yang sama dalam tujuh hari.

Pak Yusuf, direktur teknologi, mengusulkan jalan yang terdengar paling tegas: tiga petugas khusus yang tugasnya hanya mengawasi agen, duduk di depan layar sepanjang hari, dan mengambil alih begitu ada tanda bahaya. “Mereka pasti lebih cepat, karena tidak sambil mengerjakan hal lain,” katanya.

Kepala layanan nasabah membawa temuan Taobao ke meja. Pengawas khusus memang bisa lebih cepat, katanya, tetapi para penelitinya mengingatkan soal beban emosi dan keterampilan yang terkikis. Bu Ratna, direktur keuangan, menambahkan pertanyaan yang biasa ia

ajukan. “Kalau tiga orang itu kelelahan, sakit, atau keluar bersamaan, siapa yang menanggung agen ini?”

Keputusannya peran gabungan. Lima belas petugas layanan bergiliran mengawasi agen sambil tetap melayani kasus yang bukan untuk agen. Dua aturan ditambahkan. Pertama, agen pembaca mengalihkan percakapan begitu nasabah menulis keluhan pertama, tanpa menunggu keluhan berikutnya. Kedua, setiap pengalihan karena nasabah kesal tetap masuk ke petugas berpengalaman di antara kelima belas orang itu. Dua ukuran dipantau setiap bulan: waktu dari keluhan pertama sampai petugas masuk, dan kontak ulang tujuh hari pada percakapan yang dialihkan karena nasabah kesal.

Sebulan sesudah aturan baru berjalan, misalkan waktu dari keluhan pertama sampai petugas masuk turun dari 14 menit menjadi 3 menit, dan kontak ulang turun dari 18 persen menjadi 11 persen. Percakapan yang dialihkan menjadi lebih banyak, karena aturan keluhan pertama menangkap nasabah yang dulu lolos. Bu Ratna meminta kenaikan itu dicatat di laporan bulanan bersama dua ukuran tadi, supaya penghematan agen tidak dihitung dengan angka lama.

Satu hal lagi ia minta sebelum rapat ditutup: nama atasan agen itu harus ditetapkan. Kepala layanan nasabah menjadi atasannya, dan namanya dicatat di registri agen bank yang disusun di Bagian Pembahasan "Keamanan Dimulai dari Pembatasan Akses". Usulan Pak Yusuf tidak dibuang. Ia diminta menyimpannya sebagai pilihan bila kelak jumlah pengalihan naik jauh melampaui kemampuan kelima belas petugas itu.

Rapor Agen: Lima Ukuran

Setelah masa percobaan, setiap karyawan dinilai. Pertanyaannya, apa yang dinilai? Rapor yang hanya berisi satu mata pelajaran menyesatkan, sebagus apa pun nilainya. Taobao memperlihatkan bahayanya. Bila yang diukur hanya kecepatan, agen di sana tampak sangat berhasil. Bila angka yang sama dipecah menurut jenis percakapan dan dibandingkan dengan kelompok pembanding, ceritanya menjadi empat.[146]

Satu Uji Coba, Empat Cerita

Jenis percakapan	Waktu layanan	Nilai dan kontak ulang
Semua percakapan	3,2 persen lebih singkat	Tidak berubah secara bermakna
Yang memenuhi syarat ditangani agen	16,8 persen lebih singkat	Nilai turun 0,412 poin
Yang tetap dikerjakan manusia	1,8 persen lebih singkat	Nilai naik 0,091 poin
Dialihkan karena pelanggan kesal	40,8 persen lebih lama	Kontak ulang naik 6 poin persen; nilai turun 0,928 poin

Dari tabel itu saya menyusun lima ukuran yang layak masuk rapor agen.

1. Persentase tugas yang selesai dengan benar tanpa diserahkan ke manusia. Inilah ukuran pokok pekerjaan agen. Agen yang cepat tetapi sering menyerahkan pekerjaannya ke manusia sebenarnya hanya memindahkan beban. Hitung dari catatan sistem, bukan dari laporan vendor: berapa tugas yang masuk, berapa yang selesai tanpa disentuh manusia, dan berapa di antaranya yang ternyata keliru ketika diperiksa ulang.

2. Mutu di mata pelanggan dan di catatan operasional. Taobao memisahkan keduanya. Nilai dari pelanggan mewakili pengalaman pelanggan, sedangkan kontak ulang dalam tujuh hari menunjukkan apakah masalahnya benar-benar selesai. Keduanya bisa bergerak ke arah yang berbeda, jadi keduanya dilihat.

3. Biaya per tugas yang selesai, dengan plafon. Saran saya, biaya agen dihitung per tugas yang selesai, bukan per token atau per pengguna, dan setiap agen diberi plafon anggaran bulanan sejak hari pertama. Di Bagian Pembahasan "Kala AI Mulai Bertindak" kita sudah melihat perkiraan para konsultan Bain bahwa versi pertama alur kerja agen sering jauh lebih mahal daripada yang semestinya. Menurut tulisan yang sama, sebagian besar kelebihan biaya itu bisa dipangkas lewat rancangan yang lebih baik dan pemilihan model yang lebih cermat.[142] Pemilihan model itu pula yang membuat model kecil untuk pekerjaan rutin layak diuji.

4. Nilai dalam rupiah, di luar jam yang dihemat. Jam yang dihemat baru bernilai bila berubah menjadi uang: pesanan yang lebih banyak dengan orang yang sama, piutang yang tertagih lebih cepat, atau kerugian yang dicegah. Untuk pekerjaan berulang yang bersifat transaksional, para konsultan mencatat bahwa pemasangan awal agen menghasilkan siklus kerja 20 sampai 40 persen lebih cepat, atau biaya penanganan yang lebih rendah.[285] Angka seperti itu baru masuk rapor bila sudah muncul di laporan laba rugi perusahaan Anda sendiri. Firma riset yang memperkirakan pembatalan proyek agen tadi mengingatkan hal yang sama: nilai agen dilihat dari biaya, mutu, kecepatan, dan skala, dan yang dikejar adalah produktivitas perusahaan, bukan sekadar bantuan untuk tugas perorangan.[281]

5. Dampak ke manusia di sekitarnya. Di Taobao, pekerjaan manusia ikut membaik karena beban rutinnya berkurang. Di tempat lain, dampaknya bisa sebaliknya: petugas yang hanya menerima sisa kasus sulit bisa kelelahan. Dampak itu ikut diukur. Ukurannya bisa sederhana: jumlah kasus sulit per petugas, jam lembur, dan jumlah petugas yang keluar dari bagian itu, dibandingkan dengan sebelum agen dipasang.

Semua ukuran itu dibandingkan dengan kelompok pembanding, seperti di Taobao. Tanpa pembanding, perusahaan tidak tahu apakah perubahan datang dari agen, dari musim, atau dari hal lain.

Rapor agen dibaca sebulan sekali selama masa percobaan, lalu setiap kuartal sesudahnya. Bentuknya satu halaman: lima ukuran, angka dasar, angka kelompok pembanding, dan satu kalimat keputusan. Yang membacanya dua orang: atasan langsung agen, yang paling tahu sebab di balik angka, dan pemegang hasil, yang memutuskan langkah berikutnya.

Ukuran kelima membawa satu batas yang saya pegang. Agen boleh diawasi seketat apa pun. Manusia yang bekerja bersamanya berhak tahu data apa tentang dirinya yang ikut dipantau, untuk apa data itu dipakai, dan untuk apa data itu tidak akan pernah dipakai. Batas itu ditulis sebagai janji kepada karyawan, bukan sekadar pengaturan di sistem.

Sketsa: Rapor Masa Percobaan Agen Penagih

Seperti apa rapor itu ketika diisi? Kita kembali ke PT Contoh Distribusi. Perusahaan, tokoh, dan angka dalam sketsa ini rekaan, dan kesamaan dengan pihak nyata tidak disengaja. Yang nyata adalah cara mengukurnya, yang meniru uji acak Taobao dengan kelompok pembanding, dan beban per jam karyawan yang kita hitung di Bagian Pembahasan "Langganan AI Dibayar Dolar, tapi Gaji Rupiah" dari data upah BPS.[146]

Uji coba tiga bulan agen penagih, yang disetujui Pak Hendra di Bagian Pembahasan "Kala AI Mulai Bertindak", hampir selesai. Sejak awal, atas usul kepala bagian piutang, agen itu hanya menyusun pengingat untuk separuh toko yang tagihannya jatuh tempo; separuh lainnya tetap ditangani staf seperti biasa. Pembagiannya diundi. Misalkan setiap bulan ada 600 toko yang tagihannya jatuh tempo: 300 masuk kelompok agen dan 300 masuk kelompok staf. Di kelompok agen, pengingat disusun pada malam hari lewat layanan terkelola yang menjalankan model A, model yang di Bagian Pembahasan "Benarkah Server Sendiri Lebih Murah?" dipakai untuk meringkas catatan kunjungan, lalu disetujui staf pada pagi harinya. Agen tetap di A2: tidak ada satu pengingat pun yang terkirim tanpa persetujuan.

Menjelang rapat keputusan, Bu Sandra, direktur keuangan dan pemegang hasil agen piutang sejak peta di pembahasan sebelumnya, meminta satu hal yang belum pernah dibuat: uraian jabatan tertulis untuk agen penagih. Ternyata sebagian besar isinya sudah ada, tetapi tersebar di empat tempat. Kondisi berhenti ada di Uji Dompet Bagian Pembahasan "Kala AI Mulai Bertindak", daftar konektor yang diizinkan di Bagian Pembahasan "Ketika AI Lepas Kendali", pemilik dan sponsornya di registri Bagian Pembahasan "Keamanan Dimulai dari Pembatasan

Akses", dan nama pemegang hasilnya di peta Bagian Pembahasan "Siapa yang Bertanggung Jawab?". Dua baris belum pernah ditulis sama sekali: batas biaya per bulan, dan kapan agen harus menyerahkan pekerjaan ke manusia.

Uraian jabatan yang ia tulis pagi itu muat di satu halaman. Tugasnya menyusun pengingat untuk faktur yang jatuh tempo. Larangannya menjanjikan potongan, menghapus denda, mengubah jadwal bayar, dan mengirim data pelanggan ke luar layanan yang disetujui. Atasan langsungnya kepala bagian piutang, yang setiap pagi melihat pengingat yang disetujui stafnya. Ukurannya dua: berapa toko yang membayar dalam 14 hari sesudah pengingat, dan berapa piutang yang lewat 60 hari, masing-masing dengan angka dasar tiga bulan sebelum uji coba. Plafonnya, misalkan, Rp1,2 juta sebulan. Syarat berhentinya disalin dari Uji Dompet. Baris tentang kapan pekerjaan diserahkan ke staf masih kosong; baris itu baru terisi sesudah rapor dibaca.

Rapor tiga bulan, misalkan, berisi angka-angka ini. Di kelompok agen, 64 persen toko membayar dalam 14 hari sesudah pengingat; di kelompok staf, 61 persen. Piutang yang lewat 60 hari di kelompok agen turun dari Rp210 juta menjadi Rp180 juta, sedangkan di kelompok staf turun dari Rp205 juta menjadi Rp195 juta. Selisih penurunannya Rp20 juta. Bu Sandra menuliskannya dengan hati-hati: Rp20 juta itu uang yang masuk lebih cepat, bukan laba. Sesuai periksa silang yang disepakati, angka rapor itu divalidasi Pak Dedi, yang tidak ikut membangun agen penagih.

Sisi lain rapor kurang manis. Toko di kelompok agen mengirim 11 keluhan, kelompok staf 4. Ketika ditelusuri, keluhan itu datang terutama dari toko yang sedang mengurus retur barang: mereka menerima pengingat untuk faktur yang nilainya masih diperdebatkan. Polanya mirip Taobao. Pekerjaannya lebih cepat, tetapi pelanggan yang

masalahnya tidak standar justru dirugikan. Pak Dedi menambahkan bahwa beberapa keluhan itu lebih dulu sampai ke tenaga penjualannya di lapangan, persis seperti kekhawatiran Bu Sandra.

Lalu biayanya. Misalkan biaya agen Rp900 ribu sebulan untuk 300 pengingat, atau Rp3.000 per pengingat. Staf masih menyetujui setiap pengingat, rata-rata satu menit, yang dengan beban Rp26.600 per jam bernilai sekitar Rp443. Satu pengingat lewat agen jadi berbiaya Rp3.443. Staf yang menyusun sendiri butuh enam menit, sekitar Rp2.660. Per pengingat, agen lebih mahal Rp783, atau Rp234.900 sebulan. Agen memang membebaskan sekitar 25 jam kerja staf setiap bulan, tetapi Bu Sandra menghitungnya nol sampai jam itu benar-benar dipakai untuk menelepon toko yang piutangnya lewat 60 hari, seperti rencana awal.

Rapat keputusan berlangsung pendek, karena semua angka sudah tersedia di satu halaman. Pak Hendra menolak usulan menaikkan agen ke A3. Agen tetap di A2. Baris yang tadinya kosong di uraian jabatan kini terisi: toko yang punya retur atau sengketa terbuka diserahkan ke staf sebelum pengingat disusun, tanpa menunggu toko itu mengeluh. Pelajarannya sama dengan pengalihan dini di Taobao. Plafon tetap Rp1,2 juta, dan rapor berikutnya dibaca tiga bulan lagi.

Satu pertanyaan Pak Hendra dicatat untuk tahun depan: apakah pekerjaan malam seperti menyusun pengingat ini kelak lebih hemat dijalankan dengan model kecil di server sendiri atau di server bersama. Bu Sandra tidak menjawabnya di rapat itu. Ia menuliskannya di lembar harga kedaulatan Bagian Pembahasan "Benarkah Server Sendiri Lebih Murah?", yang memang dijadwalkan dihitung ulang setahun lagi.

Memberhentikan Agen

Bagian ini paling mudah terlupa ketika perusahaan bersemangat memasang agen, padahal sama pentingnya dengan seleksi. Saran saya, ada lima keadaan yang menjadi alasan untuk memberhentikan agen.

1. Biaya per tugas yang selesai lebih besar daripada biaya manusia ditambah biaya pemeriksaannya, dan tidak ada nilai lain dalam rupiah yang menutup selisih itu.

2. Persentase tugas yang selesai mandek di bawah titik impas, bulan demi bulan, meski rancangannya sudah diperbaiki.

3. Terjadi satu insiden berat. Satu saja cukup, karena insiden berat biasanya menyentuh uang, data, atau nama baik perusahaan.

4. Nilai dari pelanggan atau ukuran mutu lain turun di bawah ambang yang disepakati.

5. Modelnya diganti tanpa uji ulang. Agen seperti itu sebaiknya langsung turun level, seperti agen Bank Contoh di Bagian Pembahasan "Keamanan Dimulai dari Pembatasan Akses" yang turun otomatis ke A2 sampai uji ulangnya lulus. Bila uji ulang tidak pernah dijalankan, agen itu diberhentikan.

Kelima syarat itu ditulis di baris ketujuh uraian jabatan, sejak hari pertama. Di pembahasan sebelumnya kita sudah melihat McDonald’s berani menghentikan uji coba pemesanan suara meski masih percaya pada arahnya. Keberanian yang sama dibutuhkan untuk agen. Pilihan di gerbang keputusan, dari memperluas sampai menghentikan, dibahas di Bagian Pembahasan "Sembilan Puluh Hari Pertama". Di sini cukup alasan-alasannya.

Tidak setiap masalah harus berakhir dengan pemberhentian. Agen yang mutunya turun bisa diturunkan dulu satu level: tindakan yang tadinya ia kerjakan sendiri kini menunggu persetujuan, sementara penyebabnya dicari. Bila rapornya pulih sesudah perbaikan, levelnya boleh dinaikkan lagi dengan aturan yang sama seperti masa percobaan. Pemberhentian menjadi pilihan untuk keadaan yang tidak bisa diperbaiki, atau yang biaya perbaikannya lebih mahal daripada nilai agen itu sendiri.

Agen yang diberhentikan juga perlu dipensiunkan dengan rapi: aksesnya dicabut, kuncinya ditarik, catatan kerjanya disimpan, dan namanya dicoret dari buku induk dengan tanggal yang jelas. Rincian pencabutan kunci sudah ada di Bagian Pembahasan "Keamanan Dimulai dari Pembatasan Akses". Rapor dan catatan kerja agen itu disimpan untuk penggantinya. Tanpa itu, agen baru yang dipasang untuk pekerjaan yang sama akan jatuh di lubang yang sama, mengulang kesalahan yang ongkosnya sudah pernah dibayar perusahaan.

Memberhentikan agen jauh lebih ringan daripada memberhentikan karyawan, karena agen tidak sakit hati. Yang sulit justru manusia di sekitarnya: tim yang sudah berbulan-bulan membangunnya, dan atasan yang telanjur melaporkan keberhasilannya ke direksi. Karena itu syarat berhenti ditulis sebelum siapa pun punya kepentingan agar agennya terus bekerja.

Lakukan dan Jangan

Kembali sebentar ke Taobao. Agen di sana tidak diberi semua pekerjaan, punya pengawas yang jelas, dialihkan ke manusia oleh dua jenis alarm, dan dinilai dengan kelompok pembanding. Hasilnya tidak sempurna, tetapi perusahaan tahu persis di mana agennya membantu dan di mana ia merugikan. Pegangan bab ini bisa diringkas menjadi lima hal yang dilakukan dan lima hal yang dihindari.

Lakukan

1. Tanyakan dulu apakah pekerjaan itu perlu agen. Otomasi berbasis aturan atau asisten yang memberi saran sering cukup, dan lebih murah.

2. Tulis uraian jabatan tujuh baris sebelum agen bekerja. Bila satu baris kosong, lowongannya belum siap dibuka.

3. Tetapkan satu nama manusia sebagai atasan setiap agen, dan catat agen itu di buku induk. Setiap pergantian atasan dicatat seperti serah terima jabatan.

4. Naikkan level otonomi per jenis tindakan, hanya sesudah ada bukti. Lama pengamatan dan jumlah tindakan, dua-duanya harus terpenuhi.

5. Nilai agen dengan rapor lima ukuran dan kelompok pembanding. Kecepatan hanya satu mata pelajaran.

Jangan

1. Membeli label "agen" tanpa memeriksa isinya. Minta vendor memperlihatkan tindakan yang benar-benar dikerjakan agen itu sendiri.

2. Menunggu pelanggan kesal sebelum memanggil manusia. Alihkan pada keluhan pertama.

3. Membiarkan alur kerja panjang berjalan tanpa pemeriksaan di antara langkah. Pecah alurnya dan periksa setiap langkah.

4. Menghitung jam yang dihemat sebagai uang sebelum jam itu dipakai. Nilai agen dicatat dalam rupiah, di laporan laba rugi.

5. Membiarkan pengawas manusia kelelahan, atau agen bekerja tanpa atasan. Keduanya menyimpan masalah yang baru terlihat ketika sudah mahal.

Jim Collins menulis bahwa visi yang hebat tanpa orang yang hebat tidak ada artinya, dan kalimat itu tetap berlaku di kantor yang mulai dipenuhi agen. Agen bisa direkrut dalam seminggu, tetapi orang yang mengawasinya, membaca rapornya, dan berani memberhentikannya tetap harus dipilih dengan cermat.

Pertanyaan berikutnya lebih sulit, karena menyangkut orang-orang itu sendiri: bila agen mengambil sebagian pekerjaan, apa yang terjadi dengan karyawan yang dulu mengerjakannya? Pembahasan berikutnya akan menjawab pertanyaan itu.

Panduan Daftar Pustaka

Halaman-halaman sebelum ini adalah isi bab *Mempekerjakan Agen AI di Perusahaan Anda* (Bab 16), dengan satu halaman Identitas Dokumen yang disisipkan setelah halaman pembuka bab. Halaman-halaman setelah ini adalah Daftar Pustaka bab tersebut, memuat 10 sumber. Panduan ini menerangkan cara keduanya saling menunjuk.

Penanda di dalam bab

Sumber sebuah pernyataan ditandai dengan angka kecil di ujung kalimat. Angka itu adalah nomor catatan, bukan bagian dari kalimat dan bukan data.

Contoh bentuknya: “Kalimat yang dikutip dari sebuah sumber.[12]” Angka 12 menunjuk ke entri nomor 12 di Daftar Pustaka.

- Kalimat dengan angka kecil memuat pernyataan atau temuan yang diambil dari sumber lain. Sumbernya adalah entri bernomor sama di Daftar Pustaka.
- Kalimat tanpa angka kecil adalah penjelasan, pertimbangan, atau kesimpulan penulis buku.
- Satu kalimat dapat memiliki lebih dari satu nomor. Setiap nomor menunjuk satu sumber.
- Satu nomor menunjuk satu sumber di seluruh buku. Sumber yang sama memakai nomor yang sama di bab mana pun, sehingga nomor di bab ini tidak harus berurutan dari 1.

Membaca Daftar Pustaka

Setiap entri di Daftar Pustaka dimulai dengan nomor catatan, diikuti nama penulis atau lembaga, judul, tahun terbit, dan alamat tautan bila ada. Untuk mengetahui asal sebuah pernyataan, cocokkan nomor kecil di bab dengan nomor di awal entri.

Daftar Pustaka di berkas ini hanya memuat sumber yang dirujuk bab ini, diurutkan menurut nomor: 142, 146, 278–285. Nomor yang tidak ditemukan di sana berarti sumbernya tidak dicantumkan di dokumen ini.

Menyebut tempat sebuah kutipan

Untuk menyebut asal sebuah pernyataan secara lengkap, ada dua keterangan yang saling melengkapi:

1. **Tempat di buku ini:** judul bab dan nomor halaman cetak.
2. **Sumber aslinya:** entri Daftar Pustaka dengan nomor catatan yang sama.

Nomor halaman cetak terletak di bagian atas setiap halaman isi bab, misalnya “Mempekerjakan Agen AI di Perusahaan Anda | 107”. Nomor ini berbeda dari urutan halaman pada berkas PDF, karena berkas ini memuat halaman tambahan di

luar isi bab. Identitas Dokumen, Panduan ini, dan Daftar Pustaka tidak bernomor halaman.

Daftar Pustaka

Bab 16: Mempekerjakan Agen AI di Perusahaan Anda. 10 sumber.

142. Wang, J., A. Hoecker, A. Koh, dkk. *Managing Token Spending without Choking Off Opportunity*. Bain Technology Report 2026, 29 September 2026. https://bain.com/insights/managing-token-spending-without-choking-off-opportunity-technology-report-2026/.

146. Wang, Yiwei, Chuan Zhu, Tianjun Feng, dkk. “Agentic AI and Human-in-the-Loop Interventions: Field Experimental Evidence from Alibaba’s Customer Service Operations.” Pracetak, arXiv 2605.14830, versi 1 terbit 14 Mei 2026, versi 2 terbit 1 Juni 2026. https://arxiv.org/abs/2605.14830.

278. Collins, Jim. “First Who – Get the Right People on the Bus.” jimcollins.com, t.t. https://jimcollins.com/article_topics/articles/first-who.html.

279. Collins, Jim. *Good to Great*. New York: HarperBusiness, 2001.

280. Yee, Lareina, Michael Chui, Roger Roberts, dkk. “One year of agentic AI: Six lessons from the people doing the work.” McKinsey & Company (QuantumBlack), 12 September 2025. https://mckinsey.com/capabilities/quantumblack/our-insights/one-year-of-agentic-ai-six-lessons-from-the-people-doing-the-work.

281. Gartner, Inc. “Gartner Predicts Over 40% of Agentic AI Projects Will Be Canceled by End of 2027.” Siaran pers, 25 Juni 2025. https://gartner.com/en/newsroom/press-releases/2025-06-25-gartner-predicts-over-40-percent-of-agentic-ai-projects-will-be-canceled-by-end-of-2027.

282. Belcak, P., G. Heinrich, S. Diao, dkk. “Small Language Models are the Future of Agentic AI.” Pracetak, arXiv 2506.02153, versi 1 terbit 2 Juni 2025, versi 3 terbit 22 September 2026. https://arxiv.org/abs/2506.02153.

283. Bucher, M. J. J., dan M. Martini. “Fine-Tuned ‘Small’ LLMs (Still) Significantly Outperform Zero-Shot Generative AI Models in Text Classification.” Pracetak, arXiv 2406.08660, versi 1 terbit 12 Juni 2024, versi 2 terbit 16 Agustus 2024. https://arxiv.org/abs/2406.08660.

284. Lashinsky, Adam. “How Apple works: Inside the world’s biggest startup.” *Fortune*, 9 Mei 2011. https://fortune.com/2011/05/09/inside-apple/.

285. Singla, Alex, Alexander Sukharevsky, dkk. “The change agent: Goals, decisions, and implications for CEOs in the agentic age.” McKinsey & Company, 1 Oktober 2025. https://mckinsey.com/capabilities/strategy-and-corporate-finance/our-insights/the-change-agent-goals-decisions-and-implications-for-ceos-in-the-agentic-age.