



RAPORT • SIERPIEŃ 2026

Dane podstawowe w polskich przedsiębiorstwach

*Rynek, regulacje i praktyka porządkowania danych
o kontrahentach, materiałach i pracownikach*

SNOK

SNOK Sp. z o.o. • Warszawa • snok.ai

Spis treści

1.	Streszczenie dla zarządu	3
2.	Rok konsolidacji rynku zarządzania danymi	5
3.	Polskie przedsiębiorstwa wobec danych: stan wyjściowy	7
4.	KSeF: powszechny sprawdzian jakości danych kontrahentów	11
5.	Wyzwania CIO: migracje, które czekają w kalendarzu	14
6.	Sztuczna inteligencja zaczyna się od kartoteki	18
7.	Dlaczego programy porządkowania danych nie przynoszą efektów	20
8.	Podejście przyrostowe: metoda i mierniki	23
9.	Podsumowanie	26
10.	Nota metodyczna i źródła	27
	O SNOK	28

1

Streszczenie dla zarządu

Niniejszy raport porządkuje wiedzę o danych podstawowych, czyli o rekordach opisujących kontrahentów, dostawców, materiały, produkty, pracowników i konta finansowe. Zebrane w nim liczby pochodzą z badań publikowanych między 2020 a 2026 rokiem; każda występuje z nazwą źródła i datą.

1 Rynek przeszedł konsolidację, jakiej nie widział od lat.

Salesforce zamknął przejęcie Informaticy (listopad 2025), SAP przejął Reltio (maj 2026), a Gartner po ponad czteroletniej przerwie przywrócił raport Magic Quadrant dla tej kategorii (kwiecień 2026). Obie transakcje ogłoszono z tym samym uzasadnieniem: dane gotowe do pracy ze sztuczną inteligencją.

2 Polskie przedsiębiorstwa wchodzą w ten okres z niską dojrzałością danych.

Według EY zaledwie 9 procent średnich i dużych firm w Polsce dysponuje kompletną infrastrukturą danych gotową do zasilenia modeli AI (badanie z przełomu 2025 i 2026 roku). Analizę danych wykonuje 24,5 procent polskich firm (Eurostat, 2025), a technologie AI stosuje 8,7 procent, przy średniej unijnej 20 procent (Eurostat, 2025).

3 KSeF stał się pierwszym powszechnym sprawdzianem jakości kartotek kontrahentów.

Od 1 lutego do 26 maja 2026 roku przez system przeszło ponad 289 milionów faktur od ponad 2 milionów wystawców (dane Ministerstwa Finansów). System nie weryfikuje, czy NIP nabywcy należy do właściwego podmiotu; błąd w kartotece oznacza fakturę doręczoną niewłaściwej firmie i korektę do zera, ponieważ noty korygujące przestały istnieć.

4 Koniec wsparcia SAP ECC ustawia kolejkę konwersji, a konwersja nie poprawia danych.

Podstawowe wsparcie kończy się 31 grudnia 2027 roku. Według DSAG 54 procent ankietowanych organizacji nadal korzysta z ECC lub starszych wersji Business Suite (2026), a według ISG 58 procent programów migracji przekracza budżet i harmonogram (2026). Podobne cezury ogłaszają inni dostawcy: Microsoft zakończy standardowe wsparcie Dynamics GP z końcem 2029 roku (komunikat z września 2024).

5 Programy porządkowania danych zawiodą częściej z powodu zakresu niż technologii.

Gartner prognozował, że ponad 75 procent programów zarządzania danymi podstawowymi nie osiągnie celów biznesowych w horyzoncie 2025 roku (Magic Quadrant, grudzień 2021). Typowe błędy się powtarzają: zbyt wiele domen naraz, brak właściciela danych z mandatem, scalanie rekordów bez nadzoru człowieka.

6 Alternatywą dla programu na lata jest podejście przyrostowe.

Jedna domena danych, jedno kryterium sukcesu ustalone przed startem i pomiar na rzeczywistych danych organizacji. Skalę problemu można oszacować własnymi siłami w kilka dni; cztery pytania kontrolne zawiera rozdział 8.

*Przewagę zbuduje nie ten, kto pierwszy kupi narzędzia,
lecz ten, kto wcześniej uporządkuje warstwę,
na której te narzędzia pracują*

2

Rok konsolidacji rynku zarządzania danymi

Między listopadem 2025 a majem 2026 roku najwięksi dostawcy oprogramowania wydali miliardy dolarów na firmy zajmujące się jakością i spójnością danych. Powód obu transakcji był ten sam i został zakomunikowany wprost: bez uporządkowanych danych nie ma wiarygodnej sztucznej inteligencji.

Salesforce zamknął 18 listopada 2025 roku przejęcie Informaticy, wycenione na około 8 miliardów dolarów (komunikat Salesforce). SAP ogłosił 27 marca 2026 roku zamiar przejęcia Reltio, dostawcy platformy zarządzania danymi podstawowymi, i zamknął transakcję 7 maja 2026 roku; w komunikacie wskazał, że celem jest przygotowanie danych, zarówno z systemów SAP, jak i spoza nich, do pracy z AI (news.sap.com).

Trzeci sygnał przyszedł od analityków. Gartner przywrócił w kwietniu 2026 roku raport Magic Quadrant for Master Data Management Solutions, którego poprzednia edycja ukazała się w grudniu 2021 roku; w nowej edycji ocenił dwudziestu dostawców. Mordor Intelligence szacuje wartość światowego rynku rozwiązań tej kategorii na około 18,2 miliarda dolarów w 2025 roku, a Grand View Research wycenił go na 19,9 miliarda dolarów już w 2023 roku; obie firmy prognozują wzrost rzędu kilkunastu procent rocznie.

Dla polskich zarządów znaczenie tych wydarzeń jest praktyczne. Kategoria, którą przez lata traktowano jako techniczne zaplecze, została wyceniona przez rynek jako warunek wstępny automatyzacji. Kto planuje wdrożenia AI, agentów lub migrację systemu ERP, ten prędzej czy później stanie przed pytaniem o stan własnych kartotek.

~8 mld USD

wartość przejęcia Informatycy przez Salesforce

Salesforce, komunikat o zamknięciu transakcji,
18.11.2025

3

Polskie przedsiębiorstwa wobec danych: stan wyjściowy

Zanim padnie pytanie, jak porządkować dane, warto ustalić, od czego zaczynamy. Statystyka publiczna i badania rynkowe z lat 2025-2026 rysują obraz spójny: polskie firmy z informatyzowały procesy, ale nie uporządkowały danych, na których te procesy pracują.

3.1. Systemy są, porządku w danych brak

Według GUS z systemów klasy ERP korzysta 40,5 procent polskich przedsiębiorstw, o 4,5 punktu procentowego więcej niż dwa lata wcześniej; z systemów CRM 25,1 procent („Społeczeństwo informacyjne w Polsce w 2025 r.", GUS, październik 2025). Płatne usługi chmurowe kupuje 55,3 procent firm. Infrastruktura zatem istnieje i rośnie.

Inaczej wygląda praca na danych. Analizę danych wykonuje 24,5 procent polskich firm, podczas gdy w Danii 60 procent (Eurostat, 2025). Technologie AI stosuje w Polsce 8,7 procent przedsiębiorstw według GUS i 8,4 procent według Eurostatu; średnia unijna to 20 procent, co daje Polsce drugie miejsce od końca, przed Rumunią. Różnica między GUS a Eurostatem wynika z odmiennych metodyk i obie wartości podajemy z tego powodu obok siebie.

3.2. Dane jako bariera, nazwana po imieniu

Za średnią kryje się silne rozwarstwienie: AI stosuje 42 procent dużych firm, 15,6 procent średnich i 6,1 procent małych (GUS, 2025). Wskaźniki cyfryzacji opisują przy tym dojrzałość otoczenia, a nie mierzą wprost jakości kartotek; tę pokazują badania omówione niżej.

Badanie EY przeprowadzone na próbie około 500 polskich firm (CubeResearch, badanie z IV kwartału 2025, wyniki publikowane w 2026 roku) dostarcza liczby, którą warto zapamiętać: zaledwie 9 procent średnich i dużych przedsiębiorstw w Polsce ma kompletną infrastrukturę danych gotową do zasilenia modeli AI, a kolejne 30 procent dysponuje wyłącznie danymi tradycyjnymi. Jakość danych została w tym badaniu wskazana wśród najważniejszych barier wdrożeń, obok trudności z operacjonalizacją strategii i deficytu kompetencji.

Podobny obraz przynosi badanie Algolytics i SW Research na próbie ponad 700 firm (2025): 73 procent polskich przedsiębiorstw deklaruje brak kompetencji do wykorzystania potencjału danych, a 66 procent nie wykorzystuje danych w procesach decyzyjnych.

Z naszej praktyki projektowej wynika prosty mechanizm stojący za tymi deklaracjami. Ten sam kontrahent istnieje w ERP z pełną nazwą, w CRM ze skrótem wpisanym ręcznie, w systemie zakupowym bez numeru NIP. Każdy system ma rację lokalnie i żaden nie ma pełnej. Koszt tego stanu nie występuje w budżecie jako osobna pozycja, ponieważ rozkłada się na ręczne uzgadnianie sald, wyjaśnianie warunków płatności i raporty składane z kilku wersji prawdy.

Pięć miejsc, które płacą za ten sam duplikat

- **Finanse**
ręczne uzgadnianie sald tego samego kontrahenta
- **Zakupy**
pytania do dostawcy, czy to ta sama firma
- **Magazyn**
druga dostawa części, która leży na półce
- **Kontroling**
raport składany z trzech wersji prawdy
- **IT**
skrypt scalający kartoteki, pisany „na chwilę”

*Żadne z pięciu miejsc nie widzi pełnego kosztu; do budżetu nie trafia ani jedna pozycja.
Opracowanie SNOK.*

9%

średnich i dużych firm w Polsce ma kompletną infrastrukturę danych gotową do zasilenia modeli AI

EY Polska, badanie CubeResearch (ok. 500 firm, IV kw. 2025), publikacja 2026

3.3. Fuzje i przejęcia: rosnąca ekspozycja

W 2025 roku na polskim rynku odnotowano 330 transakcji fuzji i przejęć, rok wcześniej 348 (M&A Index Poland, Fordata i Navigator Capital). Największą transakcją 2025 roku była sprzedaż 49 procent akcji Santander Bank Polska na rzecz Erste Group za około 29,5 miliarda złotych.

Każde przejęcie oznacza łączenie katalogów kontrahentów, indeksów materiałowych i planów kont prowadzonych według różnych zasad. Publicznych badań mierzących skalę tego zjawiska w Polsce brak, dlatego traktujemy wolumen transakcji jako miarę ekspozycji, a nie dowód. Z projektów, które prowadziliśmy w grupach kapitałowych po akwizycjach, w tym w koncernie chemicznym łączącym dane pięciu spółek z trzech krajów, wynika jednak powtarzalny wzorzec: pierwsze tygodnie integracji schodzą na ustalenie, która spółka używa którego katalogu i czym różnią się definicje pozornie tych samych pojęć.

4

KSeF: powszechny sprawdzian jakości danych kontrahentów

Obowiązkowy Krajowy System e-Faktur jest pierwszym mechanizmem, który sprawdza kartoteki kontrahentów wszystkich polskich podatników jednocześnie, codziennie i bez taryfy ulgowej dla przyzwyczajień.

4.1. Skala i harmonogram

Obowiązek wystawiania faktur w KSeF objął od 1 lutego 2026 roku przedsiębiorstwa o sprzedaży przekraczającej w 2024 roku 200 milionów złotych, a od 1 kwietnia 2026 roku pozostałych podatników; najmniejsze podmioty dołączą 1 stycznia 2027 roku (Ministerstwo Finansów, ksef.podatki.gov.pl). Od 1 lutego wszyscy podatnicy mają przy tym obowiązek odbierania faktur przez system.

Skala jest bezprecedensowa: od 1 lutego do 26 maja 2026 roku w KSeF wystawiono ponad 289,5 miliona e-faktur, a wystawców było ponad 2 miliony (dane MF podane za pośrednictwem PAP, maj 2026).

4.2. Co system sprawdza, a czego nie

Faktura ustrukturyzowana według wzoru FA(3) zawiera kilkaset pól; analizy branżowe mówią o ponad trzystu, przy czym tylko część jest obowiązkowa (Ministerstwo Finansów nie publikuje oficjalnej liczby). Dane nabywcy, w tym NIP, nazwa i adres, wypełnia wystawca na podstawie własnej kartoteki.

Najważniejsza właściwość systemu bywa zaskoczeniem: KSeF waliduje zgodność pliku ze schematem, natomiast nie weryfikuje merytorycznie, czy NIP nabywcy należy do podmiotu, którego dotyczy transakcja. Faktura z NIP-em pomylnym, nieaktualnym lub należącym do innej firmy otrzyma numer KSeF i zostanie udostępniona podmiotowi o tym numerze NIP, a jeżeli podany NIP nie istnieje, nie zostanie udostępniona żadnemu podmiotowi. Takie stanowisko przedstawił Dyrektor KIS w interpretacji indywidualnej z 24 marca 2026 roku (sygn. 0114-KDIP1-3.4012.44.2026.1.AMA), wydanej dla opisanego w niej stanu faktycznego.

Konsekwencje błędu również się zmieniły. Od lutego 2026 roku nabywca nie może naprawić cudzej pomyłki notą korygującą; błędny NIP wymaga od wystawcy faktury korygującej do zera oraz wystawienia nowej faktury. Za prawidłowość danych na fakturze odpowiada wystawca również wtedy, gdy błędne dane przekazał kontrahent; skutki konkretnego błędu zależą od okoliczności i właściwych przepisów.

4.3. Pierwsze półrocze w praktyce

Doradcy podatkowi opisywali w kwietniu 2026 roku powtarzalne problemy pierwszych tygodni: numery VAT UE wpisywane w miejsce NIP, te same dane umieszczane w różnych polach przez różnych kontrahentów oraz podwójny obieg dokumentów, gdy równolegle z fakturą w KSeF krąży PDF wysłany pocztą elektroniczną (Prawo.pl, 2 kwietnia 2026). Według danych jednego z dostawców oprogramowania, firmy Exorigo-Upos, 18 procent faktur korygujących pobieranych z KSeF zawiera błędy, a około 10 procent faktur trafia do systemu bez części kluczowych danych (Bankier.pl, 29 lipca 2026); są to dane z próby klientów jednego dostawcy, nie statystyka publiczna, i tak należy je czytać.

Ministerstwo Finansów zapowiedziało odroczenie do końca 2026 roku sankcji związanych z obowiązkami KSeF (za Prawo.pl, kwiecień 2026). Okres ochronny ma jednak datę końcową, a mechanizm pozostaje: jakość kartoteki kontrahentów przestała być wewnętrzną sprawą działu księgowości i stała się warunkiem sprawnego obiegu dokumentów z każdym kontrahentem osobno.

289,5 mln

e-faktur wystawionych w KSeF od 1 lutego do 26 maja 2026 roku przez ponad 2 miliony wystawców

Dane Ministerstwa Finansów, za PAP, maj 2026

5

Wyzwania CIO: migracje, które czekają w kalendarzu

Kalendarze wsparcia największych dostawców ustawiają kolejną migrację systemów na drugą połowę dekadę. Doświadczenia rynku pokazują, że najczęściej niedoszacowanym elementem tych projektów są dane.

5.1. Kalendarz wsparcia: SAP i nie tylko

SAP utrzymuje podstawowe wsparcie dla aplikacji SAP Business Suite 7 do końca 2027 roku, z płatnym wsparciem rozszerzonym do końca 2030 roku (komunikat SAP z lutego 2020 roku, potwierdzany w dokumentacji wsparcia). Dla największych klientów SAP przygotował dodatkowo subskrypcyjną opcję przejściową na lata 2031-2033 (komunikat SAP, sierpień 2025).

Cezury wsparcia nie są przy tym specyfiką SAP. Microsoft ogłosił zakończenie rozwoju i wsparcia systemu Dynamics GP z końcem 2029 roku, z poprawkami bezpieczeństwa do kwietnia 2031 roku; jako jedną z docelowych ścieżek przejścia Microsoft wskazuje Dynamics 365 Business Central (komunikat z września 2024). Na polskim rynku, gdzie obok SAP silną pozycję mają Comarch i dostawcy krajowi (IDC, 2024), ten sam moment przychodzi przy zmianie wersji systemu albo konsolidacji po przejęciu. Każda taka operacja jest w praktyce migracją danych i egzaminem z jakości kartotek.

5.2. Ile firm czeka migracja

Tymczasem według badania niemieckojęzycznej grupy użytkowników SAP DSAG z lutego 2026 roku 54 procent ankietowanych organizacji nadal korzysta z SAP ECC lub starszych wersji Business Suite (pytanie wielokrotnego wyboru), a jedynie 37 procent planuje zejść z tych systemów do końca 2027 roku. Gartner szacował na początku 2025 roku, że ponad 60 procent klientów SAP pozostaje na ECC bez podjętej decyzji o migracji (za The Register, luty 2025). Dla rynku polskiego brak publicznych danych o liczbie firm na ECC; największy polski partner SAP szacuje liczbę wszystkich użytkowników SAP ERP w Polsce na ponad 2000 firm (All for One, materiał firmowy bez daty).

Rachunek jest prosty: znaczna część konwersji odbędzie się w latach 2026-2030, w warunkach konkurencji o te same zasoby wdrożeniowe.

5.3. Migracja danych jako główne ryzyko

Badanie ISG „The State of SAP Migration” (styczeń 2026, 200 organizacji zatrudniających ponad 1000 osób) pokazuje stan realizacji: 58 procent programów migracji SAP przekracza budżet i harmonogram, a w terminie i budżecie mieści się 15 procent. Według polskiego omówienia tego badania złożoność migracji danych wskazuje 24 procent respondentów (erp-view.pl, luty 2026). Autorzy badania formułują wniosek istotny dla dalszej części tego raportu: przeniesienie do nowego systemu wieloletnich, nieuporządkowanych danych podstawowych podważa wartość analityki i AI po uruchomieniu.

5.4. Przykład SAP: Business Partner i CVI

W S/4HANA model Business Partner jest obowiązkowy: klient i dostawca przestają być odrębnymi bytami, a przejście na wspólny model realizuje mechanizm Customer Vendor Integration, wykonywany jeszcze w systemie źródłowym jako obowiązkowy krok przygotowawczy (SAP, nota 2265093 i oficjalne materiały SAP o konwersji).

Zadaniem CVI jest przeniesienie danych do nowej struktury, nie poprawa ich jakości. Trzy warianty zapisu tego samego kontrahenta pozostaną po konwersji trzema partnerami biznesowymi, o ile wcześniej nie zostaną scalone. Rekord z brakiem w polu wymaganym przez nowy model wraca do zespołu jako wyjątek do ręcznej obsługi, zwykle w oknie przełączenia systemów, czyli wtedy, gdy zespół ma najmniej czasu. Rozstrzygnięcie, czy dwa podobne rekordy opisują ten sam podmiot, wymaga przy tym wiedzy biznesowej, nie technicznej.

Stąd kolejność, którą rekomendujemy w projektach: pomiar duplikatów i braków przed konwersją, deduplikacja z udziałem właściciela danych, dopiero potem CVI. Ta sama praca wykonana po uruchomieniu kosztuje więcej, bo odbywa się pod presją produkcyjną.

Kolejność, która się broni

TYDZIEŃ 1-2

Pomiar

ile duplikatów i braków
jest w kartotece

TYDZIEŃ 3-8

Deduplikacja

scalenia podpowiada
system,
zatwierdza człowiek

DOPIERO TERAZ

Migracja

do nowego systemu
wchodzą
dane uporządkowane

Pomiar i deduplikacja przed migracją; po uruchomieniu ta sama praca odbywa się pod presją produkcyjną. Opracowanie SNOK.

Tę sekwencję stosujemy niezależnie od docelowego systemu. Pomiar zajmuje dni i nie wymaga dostępu do środowiska produkcyjnego, więc może wyprzedzić decyzję o projekcie. Deduplikacja na żywej kartotece toczy się równolegle z przygotowaniem migracji i nie blokuje jej harmonogramu. Sama migracja domyka wtedy pracę na danych, zamiast ją dopiero zaczynać - a zespół w oknie przełączenia zajmuje się systemem, nie rozstrzyganiem, czy dwa rekordy to ta sama firma.

58%

programów migracji SAP przekracza budżet i harmonogram;
w terminie i budżecie mieści się 15 procent

ISG, The State of SAP Migration (n=200), styczeń 2026

6

Sztuczna inteligencja zaczyna się od kartoteki

Modele językowe i agenci nie sprawdzają, czy dane są prawdziwe. Odpowiadają na podstawie tego, co otrzymają - równie stanowczo przy danych spójnych i sprzecznych.

Mechanizm jest łatwy do wyobrażenia. Agent zapytany o saldo kontrahenta, który występuje w systemach w czterech wariantach, poda odpowiedź natychmiast; nie zasygnalizuje, że wybrał jeden z czterech możliwych rekordów. Automatyzacja zbudowana na niespójnych danych nie usuwa nieporządku, lecz zwiększa tempo i zasięg jego skutków.

Prognozy analityków od dwóch lat mówią to samo. Gartner ocenia, że do końca 2026 roku organizacje porzucą 60 procent projektów AI, które nie zostały wsparte danymi gotowymi do pracy z AI, i wskazuje, że 63 procent organizacji nie ma odpowiednich praktyk zarządzania danymi lub nie jest ich pewne (komunikat prasowy z 26 lutego 2025 roku).

Polskie dane układają się w tę prognozę bez naciągania. Skoro technologie AI stosuje 8,7 procent przedsiębiorstw (GUS, 2025), a kompletną infrastrukturę danych pod AI deklaruje 9 procent średnich i dużych firm (EY, 2026) - to badania o różnych populacjach, ale o wspólnym kierunku - trudno bronić tezy, że bariera leży w dostępności modeli. Modele są dostępne od ręki; przygotowane dane nie.

Warstwa danych podstawowych pełni w architekturze AI rolę, której nie zastąpi ani hurtownia, ani platforma analityczna: rozstrzyga, który rekord jest właściwy i kto odpowiada za jego poprawność. Hurtownia i lakehouse gromadzą oraz modelują dane, natomiast nie orzekają o ich prawdziwości; bez warstwy wzorcowej analityka konsoliduje niespójności zamiast je usuwać.

60%

projektów AI zostanie porzuconych do końca 2026 roku, jeżeli nie zostaną wsparte danymi gotowymi do pracy z AI

Gartner, komunikat prasowy, 26.02.2025



7

Dlaczego programy porządkowania danych nie przynoszą efektów

Paradoks tej dziedziny polega na tym, że o porażkach programów zarządzania danymi wiadomo od lat, a mimo to kolejne organizacje powtarzają ten sam scenariusz.

Gartner prognozował w grudniu 2021 roku, że ponad 75 procent programów zarządzania danymi podstawowymi nie osiągnie zakładanych celów biznesowych w horyzoncie 2025 roku (Magic Quadrant for MDM Solutions). W obiegu branżowym krąży ta sama liczba z wcześniejszą datą i bez źródła; w niniejszym raporcie stosujemy wyłącznie wersję udokumentowaną.

Przyczyny rzadko leżą w technologii. W projektach i rozmowach przedwdrożeniowych obserwujemy cztery powtarzalne wzorce:

1. Zbyt szeroki zakres na start

Program obejmuje od pierwszego dnia wszystkie domeny danych. Powstaje komitet, komitet powołuje zespoły robocze, a po roku organizacja ma dokumentację warsztatów zamiast działającej warstwy danych i pierwszego mierzalnego wyniku.

2. Komitet zamiast właściciela

Nikt nie ma mandatu, aby rozstrzygnąć sporny rekord merytorycznie. Spór o to, czy dwa zapisy opisują ten sam podmiot, jest sporem biznesowym; bez właściciela danych po stronie biznesu każda decyzja czeka na kolejne spotkanie.

3. Scalanie w pełni automatyczne

Dopasowanie rekordów jest wynikiem probabilistycznym. Błędne scalenie dwóch odrębnych podmiotów jest trudniejsze do wykrycia i droższe w naprawie niż duplikat, dlatego decyzję o połączeniu powinien zatwierdzać człowiek.

4. Czyszczenie bez reguł przy wejściu

Jednorazowe uporządkowanie kartoteki bez reguł walidacji nowych rekordów daje efekt na kwartał. Potem kartoteka wraca do stanu wyjściowego, a organizacja do przekonania, że „już to robiliśmy i nie zadziałało”.

Cztery podejścia, jeden warunek

Cztery podejścia spotykane na rynku różnią się nie jakością, lecz zakresem odpowiedzialności. Jednorazowe czyszczenie rozwiązuje stan kartoteki na dzień migracji, nie chroni przed ponownym zabrudzeniem. Hurtownia i lakehouse odpowiadają za gromadzenie i analizę, nie za rozstrzyganie, który rekord jest właściwy. SAP Master Data Governance jest osadzony w ekosystemie SAP; w krajobrazach wielosystemowych trzeba osobno ocenić, czy potrzebna jest również warstwa niezależna od ERP. Platforma danych podstawowych niezależna od ERP utrzymuje rekord wzorcowy ponad systemami, natomiast nie przejmuje procesów operacyjnych, które słusznie zostają w systemach źródłowych. Te podejścia się uzupełniają; problem zaczyna się wtedy, gdy jedno z nich ma w planie projektu zastąpić pozostałe.

75%

programów zarządzania danymi podstawowymi nie osiągnie zakładanych celów biznesowych w horyzoncie 2025 roku

Gartner, Magic Quadrant for MDM Solutions, grudzień 2021

8

Podejście przyrostowe: metoda i mierniki

Odpowiedzią na statystykę porażek nie jest większy program, lecz mniejszy: jedna domena danych, jedno kryterium sukcesu i wynik mierzony na rzeczywistych danych organizacji w tygodniach, nie kwartałach.

8.1. Kolejność prac

1. Wybór domeny i kryterium sukcesu

Zaczynamy od domeny, która generuje najwięcej pracy ręcznej i sporów o liczby; najczęściej są to kontrahenci albo indeksy materiałowe. Kryterium sukcesu ustalamy przed startem.

2. Pomiar stanu

Na rzeczywistych danych organizacji ustalamy liczbę duplikatów, braki w polach krytycznych i rozbieżności między systemami. Pomiar pozwala oprzeć decyzję na liczbach zamiast na opiniach.

3. Reguły rekordu wzorcowego

Normalizacja nazw i form prawnych, zasady rozstrzygania konfliktów między źródłami, wersjonowanie oraz właściciel danych z mandatem do decyzji.

4. Scalanie z nadzorem i mierniki

Modele językowe trafnie podpowiadają, że dwa zapisy mogą opisywać ten sam podmiot, i uzasadniają odpowiedź; decyzję zatwierdza opiekun danych (data steward), a każda zmiana jest rejestrowana w śladzie audytowym. Po pilocie wynik porównujemy z pomiarem wyjściowym.

Pilot jednej domeny zamykamy zwykle w cztery do ośmiu tygodni, zależnie od stanu danych i liczby systemów do zasilenia. W ten sposób pracujemy między innymi z Medicoverem w domenie danych pracowniczych: 18 krajów, około 45 tysięcy pracowników i około 100 tysięcy zdarzeń na danych rocznie, z walidacją, obiegiem akceptacji i pełną historią zmian. Do typowych mierników pilota należą: udział duplikatów w domenie, kompletność pól krytycznych, liczba rozbieżności między systemami oraz czas rozstrzygnięcia spornego rekordu.

8.2. Pomiar, który można wykonać samodzielnie

Skalę problemu w jednej domenie można oszacować własnymi siłami w kilka dni, zanim padnie decyzja o jakimkolwiek projekcie. Wystarczy wyciąg z kartoteki kontrahentów i cztery pytania:

- 1 Ile rekordów ma ten sam NIP przy różnych nazwach lub adresach?
- 2 Ile rekordów ma braki w polach, które procesy uznają za krytyczne?
- 3 Ile pozycji różni się między ERP a CRM dla tych samych podmiotów?
- 4 Kto rozstrzyga, która wersja jest właściwa, i ile czasu to zajmuje?

Odpowiedź „nie wiemy” na którekolwiek z pytań również jest wynikiem pomiaru, i to najbardziej pouczającym.

8.3. Dla kogo ta praca ma sens teraz

Priorytet mają organizacje, które przygotowują konwersję do S/4HANA, konsolidują spółki po przejęciach, uruchamiają automatyzację lub AI na danych operacyjnych albo muszą wykazać wobec audytora, gdzie leżą dane i kto ma do nich dostęp. Odradzamy natomiast start tym, którzy oczekują jednorazowego czyszczenia bez wskazania właściciela danych po stronie biznesu, planują objąć wszystkie domeny naraz przed pierwszym potwierdzonym wynikiem albo nie mają nikogo, kto rozstrzygnie sporny rekord merytorycznie. Niezależnie od zakresu program obejmuje warunki brzegowe: uprawnienia i kontrolę dostępu, ochronę danych osobowych oraz zasady retencji - szczególnie przy danych pracowniczych i danych kontrahentów. Jasna kwalifikacja oszczędza obu stronom kwartał rozmów bez wyniku.

„W pilotażach danych podstawowych najrzadziej zaskakuje technologia. Najczęściej to, jak szybko zapadają decyzje, gdy zamiast opinii na stole leży pomiar”

Michał Korzeń, CTO SNOK

9

Podsumowanie

Trzy procesy zbiegły się w czasie: rynek wycenił zarządzanie danymi podstawowymi w transakcjach wartych miliardy dolarów, polski regulator uruchomił system sprawdzający kartoteki kontrahentów wszystkich podatników, a kalendarz wsparcia SAP ustawił kolejkę konwersji, podczas których dane należą do najczęściej niedoszacowywanych ryzyk. Równoległe badania krajowe pokazują, że przygotowanie danych, nie dostępność technologii, jest wąskim gardłem wdrożeń AI w polskich firmach.

Wniosek praktyczny jest mniej efektowny niż okładki raportów o sztucznej inteligencji: przewagę zbuduje nie ten, kto pierwszy kupi narzędzia, lecz ten, kto wcześniej uporządkuje warstwę, na której te narzędzia pracują. Tę pracę można zacząć od pomiaru, który mieści się w dniach i nie wymaga żadnej decyzji inwestycyjnej.

10

Nota metodyczna i źródła

Raport opracował zespół SNOK na podstawie źródeł publicznych oraz własnych doświadczeń projektowych. Liczby rynkowe weryfikowano w lipcu i sierpniu 2026 roku; przy każdej podajemy źródło i datę. Tam, gdzie źródła różnią się metodyką (GUS i Eurostat) albo pochodzą od pojedynczego dostawcy (dane o błędach faktur), zaznaczamy to wprost. Odniesienia do projektów SNOK dotyczą prac zrealizowanych; nazwę klienta podajemy wyłącznie za jego zgodą. Raport ma charakter informacyjny i nie stanowi porady podatkowej ani prawnej; decyzje w sprawach KSeF i rozliczeń warto konsultować z doradcą podatkowym.

- GUS, „Społeczeństwo informacyjne w Polsce w 2025 r.”, październik 2025
- Eurostat, wykorzystanie AI i analizy danych w przedsiębiorstwach UE, grudzień 2025 i luty 2026
- Komisja Europejska, „Poland - 2025 Digital Decade Country Report”, czerwiec 2025
- EY Polska, „Jak polskie firmy wdrażają AI” (badanie CubeResearch, ok. 500 firm, IV kw. 2025), publikacja 2026
- Algolytics, SW Research, „Dojrzałość polskich firm w obszarze analizy danych i AI” (ponad 700 firm), 2025
- Ministerstwo Finansów, ksef.podatki.gov.pl: harmonogram obowiązkowego KSeF i broszura struktury FA(3), 2026
- Dyrektor KIS, interpretacja indywidualna sygn. 0114-KDIP1-3.4012.44.2026.1.AMA, 24 marca 2026 (za omówieniem jpk.info.pl)
- Dane MF o liczbie e-faktur i wystawców w KSeF, za PAP, maj 2026
- Prawo.pl, „Drugi etap KSeF: system działa, ale firmy mają szereg problemów”, 2 kwietnia 2026
- Bankier.pl za danymi SmartKSeF Exorigo-Upos (dane jednego dostawcy), 29 lipca 2026
- SAP, strategia wsparcia Business Suite 7 (podstawowe do końca 2027, rozszerzone do 2030), luty 2020; opcja przejściowa 2031-2033, sierpień 2025
- SAP, nota 2265093 i oficjalne materiały o Business Partner i Customer Vendor Integration
- DSAG, Investitionsreport 2026 (n=198), luty 2026
- Gartner o klientach ECC, za The Register, luty 2025
- ISG, „The State of SAP Migration” (n=200), styczeń 2026
- All for One Poland, szacunek liczby użytkowników SAP ERP w Polsce (materiał firmowy, bez daty)
- Salesforce, komunikat o zamknięciu przejęcia Informatyky, 18 listopada 2025
- SAP, komunikaty o przejściu Reltio, 27 marca i 7 maja 2026
- Gartner, Magic Quadrant for MDM Solutions: XII 2021 (prognoza o programach MDM) oraz powrót raportu w IV 2026
- Mordor Intelligence (2025) i Grand View Research (2023), szacunki wielkości rynku MDM
- Gartner, komunikat prasowy o projektach AI i praktykach zarządzania danymi, 26 lutego 2025
- Fordata, Navigator Capital, „M&A Index Poland”: podsumowania 2024 i 2025
- Microsoft, komunikat o zakończeniu wsparcia Dynamics GP, 25 września 2024
- IDC, „Poland Enterprise Application Software Market” (udziały rynku ERP w Polsce), 2024

O SNOK

SNOK Sp. z o.o. jest polską firmą konsultingową IT z Warszawy. Prowadzimy projekty w obszarach technologii SAP, cyberbezpieczeństwa, automatyzacji oraz budowy oprogramowania i rozwiązań AI. Pracujemy w systemach zarządzania zgodnych z ISO 27001:2022 i ISO 9001:2015. Rozwijamy własną platformę do zarządzania danymi podstawowymi SNOK MDM, wdrażaną od 2022 roku, i doradzamy neutralnie w środowiskach SAP.

Od czego zaczynamy współpracę

Współpracę zaczynamy od pomiaru duplikatów: na próbce z jednej domeny sprawdzamy, ile rekordów opisuje ten sam podmiot, ile pozycji ma braki w polach krytycznych i gdzie powstają rozbieżności między systemami. Wynik wraz z rekomendacją zakresu przekazujemy w pięciu dniach roboczych. To najkrótsza droga do tego, żeby rozmowa o danych w Państwa organizacji zaczęła się od liczb, a nie od opinii.

Formularz pomiaru duplikatów i opis metody znajdują się na stronie snok.ai/pl/oferta/master-data-management. Pytania można kierować bezpośrednio do autorów raportu.

*Pomiar duplikatów · wynik i rekomendacja zakresu
w pięciu dniach roboczych*

Kontakt do autorów raportu:

Jacek Bugajski, CEO • jacek.bugajski@snok.ai

Michał Korzeń, CTO • michal.korzen@snok.ai