



دوره DBA هوش مصنوعی در کسب و کار

حکمرانی هوش مصنوعی در سازمان

از تصمیم برای ورود به هوش مصنوعی
تا استقرار، کنترل ریسک و پاسخ‌گویی سازمانی

مهران ضیابری

مدیرعامل شرکت پردازش هوشمند ترگمان
رئیس کمیسیون هوش مصنوعی سازمان نظام صنفی رایانه‌ای

مرداد ۱۴۰۵

یک دستور ساده، ده‌ها مسئله پیچیده



و از همین‌جا، حکمرانی هوش مصنوعی آغاز می‌شود...

هر واحد، یک مسئله متفاوت می بیند

حکمرانی هوش مصنوعی یعنی کنار هم نشان دادن مسئله ها در یک نظام تصمیم گیری



پشت گردونه مسائل، چهار کارکرد ساده است

حکمرانی هوش مصنوعی در عمل، چهار کار اصلی انجام می‌دهد



شناسایی زمینه MAP

شناسایی و تعریف مساله، کاربرد،
ذی‌نفعان، منافع و آثار



جهت‌دهی و راهبری GOVERN

تعیین سیاست‌ها، نقش‌ها،
حدود مسئولیت‌ها و اختیارات



مدیریت MANAGE

اولویت‌دهی، کنترل، واکنش،
و در صورت لزوم توقف



پایش و سنجش MEASURE

ارزیابی عملکرد، ریسک،
اعتمادپذیری و پیامدها

« هشت رکن اسلاید قبلی، صورت سازمانی همین چهار کارکردند »

مدیرعامل یک دستیار هوشمند سازمانی می خواهد، اولین تصمیم شما چیست؟

انتخاب پیمانکار



انتخاب مدل



گردآوری داده و اسناد



تأمین بودجه



ساخت نسخه پایلوت



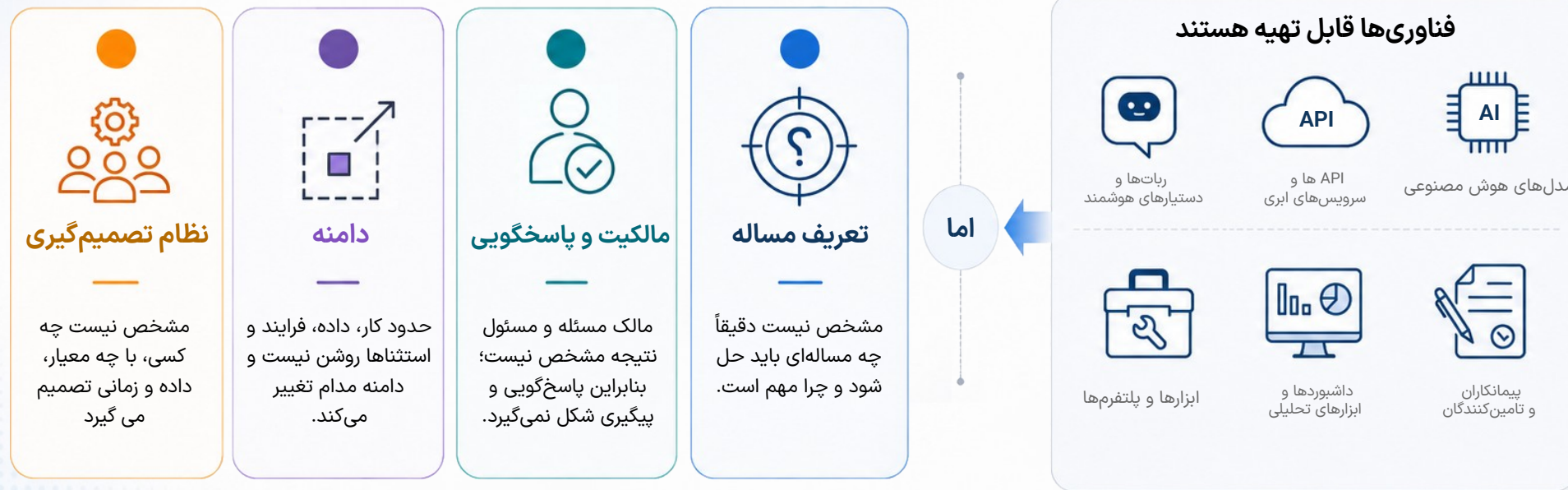
تعیین مالک و دامنه



همه تصمیم ها می توانند اولویت اصلی باشند

مشکل سازمان فقدان فناوری نیست

فقدان تعریف مساله، مالکیت، دامنه و نظام تصمیم گیری است



حکمرانی هوش مصنوعی سازمانی: تعریف عملیاتی

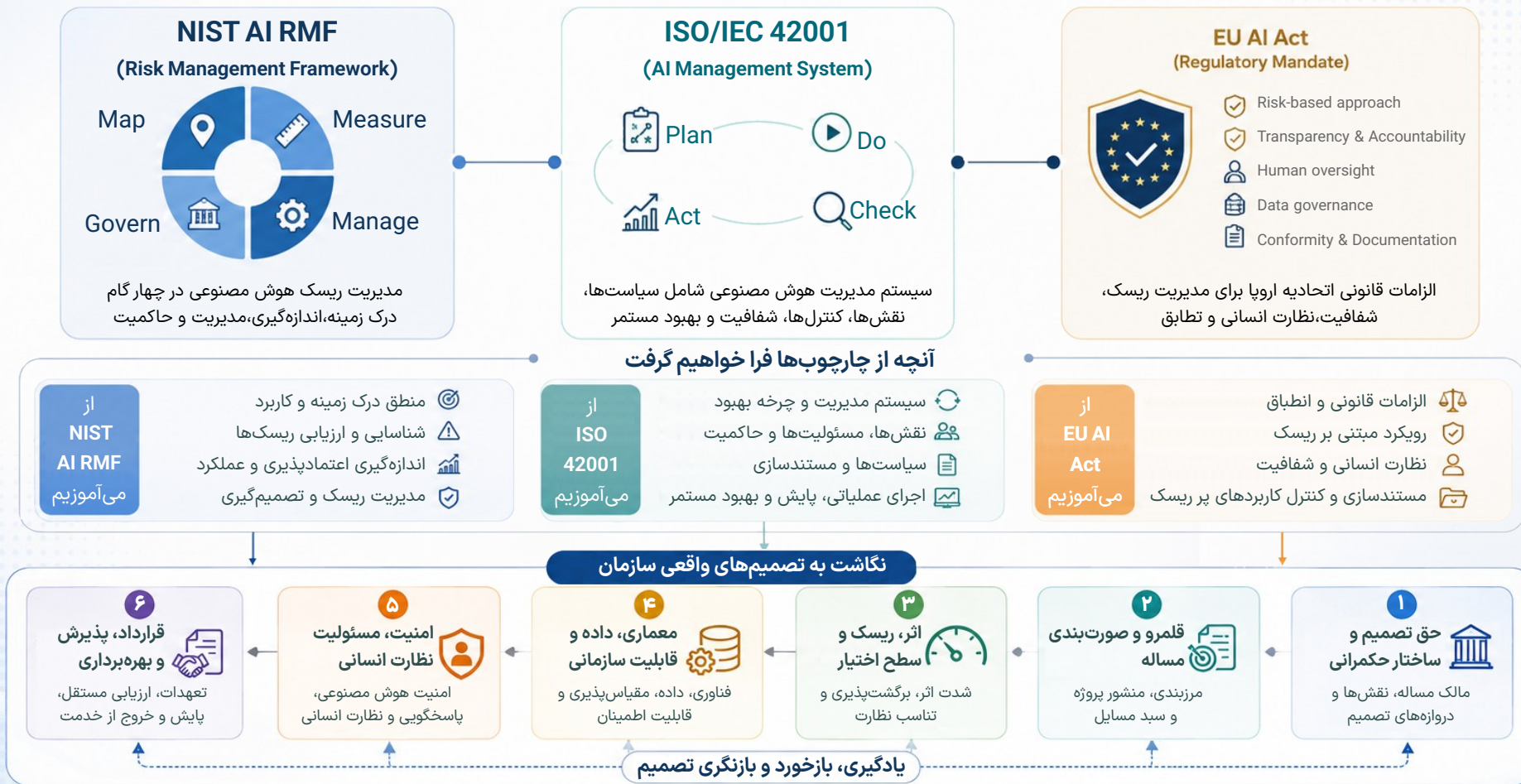
حکمرانی هوش مصنوعی سازوکاری است برای تعیین اینکه چه مساله‌ای، با چه فناوری، تحت مسوولیت چه کسی، با چه حدود و کنترل‌هایی و بر مبنای کدام معیار موفقیت حل می‌شود.



در این دوره حکمرانی هوش مصنوعی نه به عنوان مجموعه‌ای از اصول انتزاعی، بلکه به عنوان **منطق تصمیم‌گیری و کنترل** در سازمان ارایه می‌شود

رابطه طرح درس با چارچوب‌های معتبر

طرح درس، منطق NIST AI RMF، نظام ISO/IEC 42001 و الزامات حقوقی EU AI Act را در یک مسیر تصمیم‌گیری تلفیق می‌کند.





دانشگاه آزاد اسلامی
واحد تهران مرکزی

دوره DBA هوش مصنوعی

این پرونده هنوز خام است



هر مدیر، فقط بخشی از مسئله را می‌بیند.

بخش اول

مدل عملیاتی حکمرانی و حق تصمیم



پایلوت موفق، هنوز قابلیت سازمانی نیست

خروجی فنی: اجزا ساخته شده اند؛ قابلیت استفاده وجود ندارد.



راه حل های پراکنده: هر پایلوت کار می کند؛ اما سازمان هنوز
قابلیت مشترک ندارد.



محصول قابل استفاده : یک مسئله مشخص برای یک گروه کاربر
به صورت انتها به انتها حل شده است.



هر قابلیت سازمانی ممکن است با یک پروژه آغاز شود؛ اما با تحویل پروژه به وجود نمی آید.

دقیقاً چه چیزی را حکمرانی می‌کنیم؟

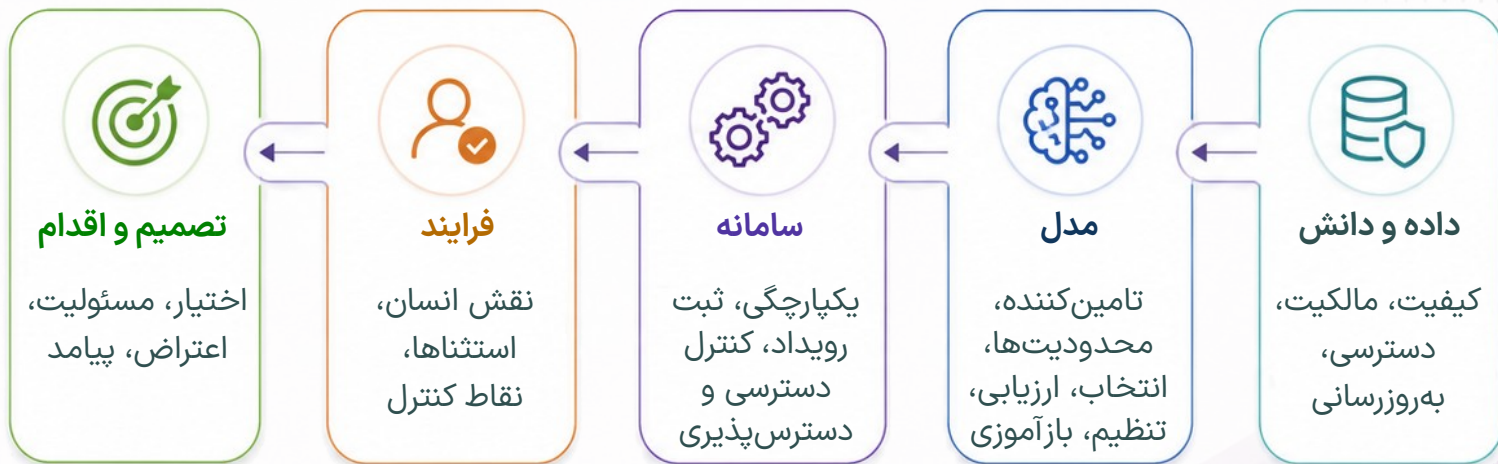
در حکمرانی هوش مصنوعی فقط درباره مدل تصمیم نمی‌گیریم؛ دامنه حکمرانی از داده تا تصمیم نهایی امتداد دارد

دامنه واقعی حکمرانی

در پرونده دستیار سازمانی



- منابع محتوایی و پایگاه دانش
- مدل زبانی و منطق پاسخگویی
- سامانه و رابط کاربر
- فرایند پاسخگویی و ارجاع
- تصمیم و اثر بر کاربر و سازمان



حکمرانی هوش مصنوعی کل زنجیره‌ای را در بر می‌گیرد که از داده و زمینه استفاده آغاز می‌شود و به تصمیم، اقدام و پیامد سازمانی می‌رسد.

حامی، مالک کسب و کار، مالک فنی و مالک ریسک

چهار نقش متفاوت؛ چهار نوع پاسخگویی



این ها نقش اند، نه الزاماً چهار نفر یا چهار واحد؛ برای هر تصمیم باید یک پاسخ گوی نهایی روشن باشد، نه یک مالک واحد برای کل سامانه.

نقشه حق تصمیم در چرخه عمر

در هر مرحله چه کسی پیشنهاد می‌دهد، چه کسی ارزیابی می‌کند، چه کسی تصویب می‌کند و چه کسی می‌تواند متوقف کند؟

تعریف مسئله	ارزیابی و تصویب	طراحی راه‌حل	توسعه و آزمون	استقرار	پایش و بازنگری
پیشنهاددهنده مالک کسب‌وکار	مالک کسب‌وکار	مالک فنی	مالک فنی	مالک فنی	مالک کسب‌وکار مالک فنی
ارزیاب مالک کسب‌وکار	مالک ریسک مالک فنی	مالک کسب‌وکار مالک ریسک	مالک فنی مالک کسب‌وکار	مالک ریسک مالک کسب‌وکار	مالک ریسک
تصویب‌کننده حامي	حامي	حامي مالک کسب‌وکار	مالک فنی	حامي	حامي
متوقف‌کننده حامي	مالک ریسک	مالک ریسک	مالک ریسک حامي	مالک ریسک	مالک ریسک حامي

مالک ریسک

مالک فنی

مالک کسب‌وکار

حامي

نمونه‌ای برای پرونده دستیار سازمانی



عبور از هر مرحله خودکار نیست؛ هر مرحله یک دروازه تصمیم دارد.

حکمرانی متمرکز، فدرال یا توزیع شده

کدام مدل برای سازمان، هلدینگ یا شرکت‌های تابعه مناسب‌تر است؟

اختیار محلی بیشتر

کنترل مرکزی بیشتر

توزیع شده



هر واحد تصمیم‌گیر و مجری اصلی راهکارهای خود است



سرعت، تناسب محلی و مالکیت بالاتر



دوباره‌کاری، پراکندگی و دشواری کنترل ریسک



واحدهای مستقل، تنوع بالا، بلوغ محلی قوی

مناسب برای:



فدرال



استاندارد و زیرساخت در مرکز، اجرا در واحدهای کسب‌وکار



تعادل میان کنترل مرکزی و چابکی محلی



نیاز به هماهنگی، نقش‌ها و قواعد روشن



هلدینگ‌ها، شرکت‌های تابعه، رشد مرحله‌ای

مناسب برای:



متمرکز



سیاست‌ها، پلتفرم و تصمیم‌های کلیدی در مرکز



یکپارچگی، کنترل و استاندارد مشترک



گلوگاه تصمیم و فاصله از نیاز واحدها



ریسک بالا، سازمان یکپارچه، بلوغ پایین‌تر

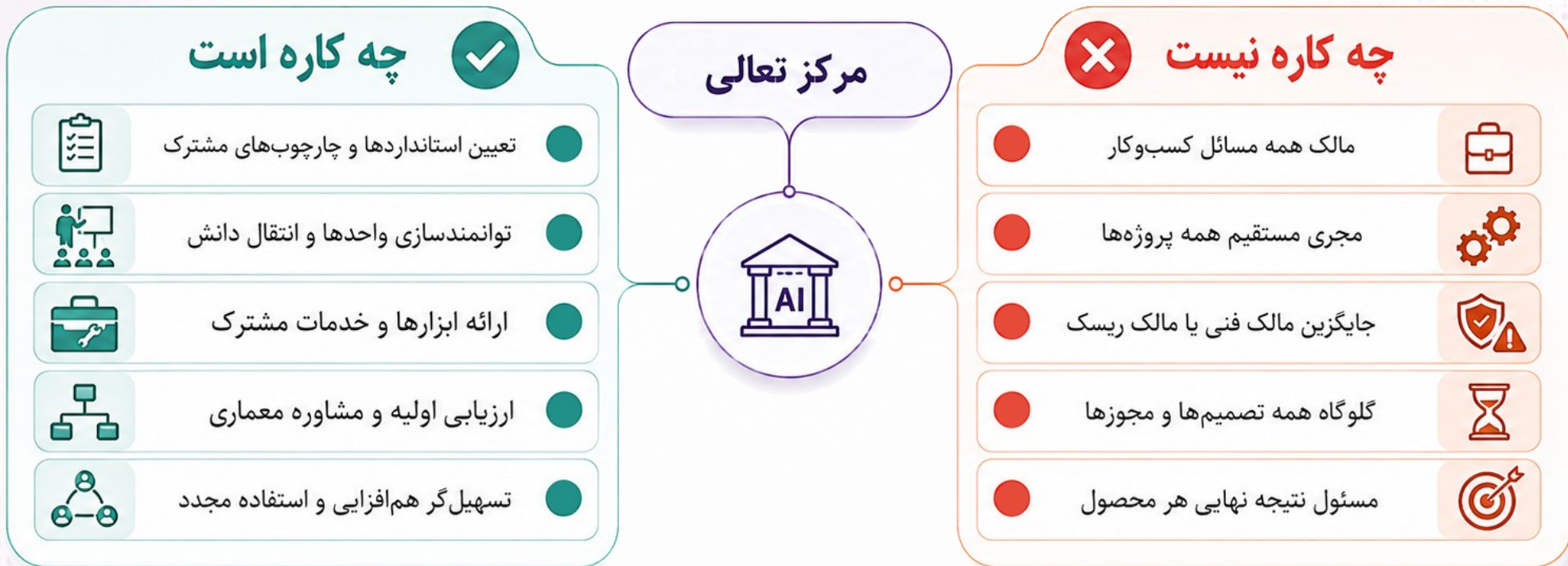
مناسب برای:



انتخاب مدل به اندازه سازمان، تنوع کسب و کار، سطح ریسک و بلوغ دیجیتال بستگی دارد.

مرکز تعالی هوش مصنوعی چه نقشی دارد؟

تفکیک نقش استانداردگذار، توانمندساز، مجری و ناظر



مرکز تعالی باید توانمندساز و هماهنگ کننده باشد؛ نه صاحب همه پروژهها و نه مجری همه راهکارها

چرخه عمر به مثابه مجموعه دروازه‌های تصمیم

چرخه عمر زمانی حکمرانی پذیر می‌شود که در نقاط ازپیش تعریف شده، صاحب تصمیم، شواهد لازم و گزینه‌های تصمیم روشن باشند.



در حکمرانی هوش مصنوعی، چرخه عمر فقط توالی کارهای فنی نیست؛ زنجیره‌ای از تصمیم‌ها، مجوزها و مسئولیت‌ها است.

کارت تحول ۱: هیچ کس به تنهایی مالک پاسخ نهایی نیست

دستیار سازمانی از پایلوت عبور کرده؛ اما هنوز مالک پاسخ نهایی روشن نیست

IT

زیرساخت و سامانه با ماست؛
مالک پاسخ محتوا نیستیم.

(HR) منابع انسانی

محتوای کارکنان با ماست؛
اما رفتار مدل با ما نیست.

دستیار سازمانی



تحول دیجیتال / کسب و کار

ما حامی پروژه بودیم،
نه پاسخگوی هر جواب.

حقوقی / ریسک و انطباق

اگر پاسخ اشتباه مبنای تصمیم شد،
مسئول کیست؟

ریسک ها	
ریسک انطباق	⚠
مسئولیت تصمیم	⚠
حفظ سوابق	⚠

مالک پاسخ نهایی:

?

- از اینجا به بعد، اختلاف IT، منابع انسانی و تحول دیجیتال آشکار می شود.



AI OR NOT AI
THIS IS THE PROBLEM

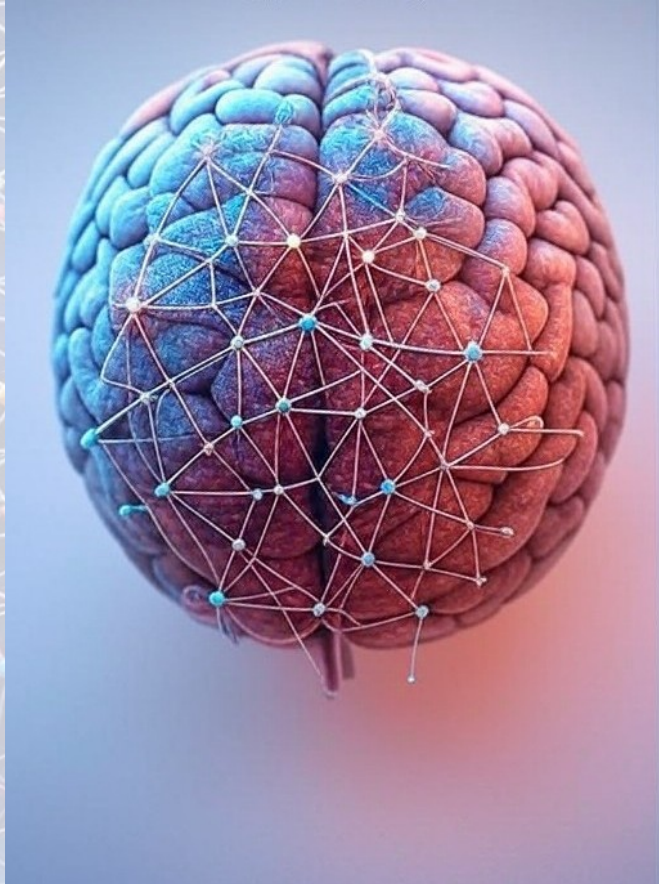
بخش دوم - قسمت اول:
چه چیزی واقعاً باید زیر حکمرانی هوش مصنوعی قرار گیرد؟

هوشمندسازی یا هوش مصنوعی

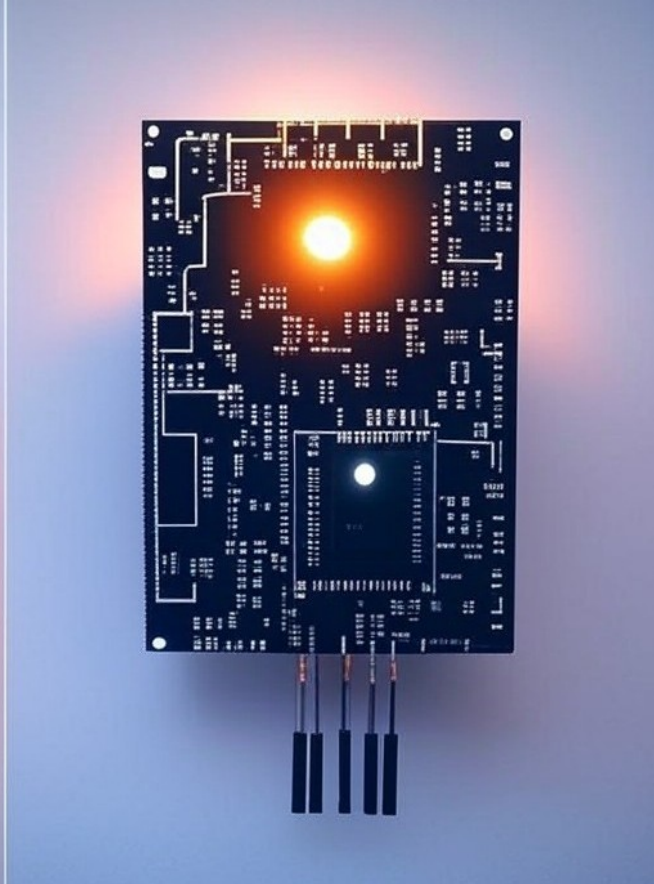


دانشگاه گیلان

دوره هوش مصنوعی



استنباط، پیش‌بینی، تولید و توصیه



اتصال، حسگر، کنترل و اتوماسیون

مقایسه هوش مصنوعی با هوشمندسازی از منظر حکمرانی

هوش مصنوعی	اینترنت اشیاء (هوشمندسازی)	
استنباط خروجی از ورودی با استفاده از یک مدل آموخته شده یا دانش محور	اجرای منطق و قواعد از پیش تعیین شده	ماهیت
نتیجه تحلیلی یا تصمیم یار/تصمیم گیر	اجرای فرایند یا واکنش از پیش تعریف شده	خروجی
خطای مدل، سوگیری، مسئولیت تصمیم، نظارت انسانی	ایمنی، تداوم عملیات، امنیت تجهیزات	ریسک اصلی
کسب و کار + فنی + ریسک	عملیات / مهندسی / فناوری	مالک اصلی
تمرکز بر رفتار سامانه، مسئولیت و کنترل تصمیم	تمرکز بر اتصال، ایمنی و پایداری	پیام حکمرانی



هرچه سامانه از «اجرای دستور» به «شکل دادن تصمیم» نزدیک تر شود، حکمرانی هوش مصنوعی پررنگ تر می شود.

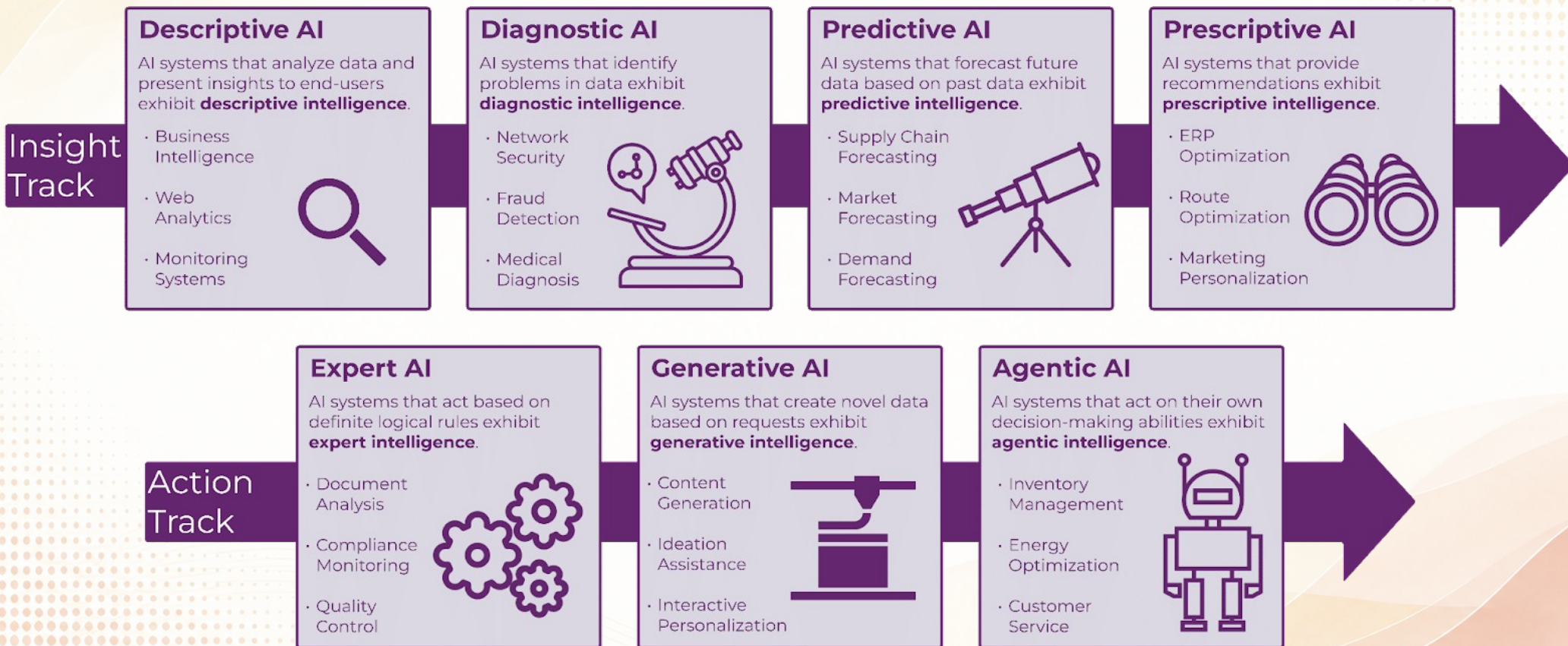
هوش تجاری، تحلیل یا تصمیم خودکار؟

هوش مصنوعی	هوش تجاری	
پیش‌بینی، پیشنهاد، خودکارسازی	گزارش و تفسیر	نقش
امتیاز، پیش‌بینی، توصیه، اقدام	داشبورد، گزارش، KPI	خروجی
تصمیم‌گیر، تأییدکننده یا ناظر؛ متناسب با ریسک کاربرد	تصمیم‌گیر اصلی	جایگاه انسان
علاوه بر کنترل‌های داده: عملکرد مدل، سوگیری، حدود استفاده، نظارت انسانی و مسئولیت پیامد	کیفیت و تبار داده، دسترسی، تعریف شاخص و تفسیر گزارش	پیام حکمرانی

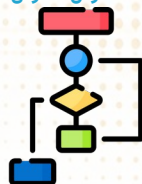
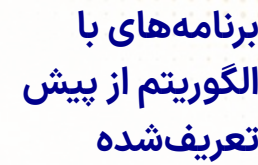
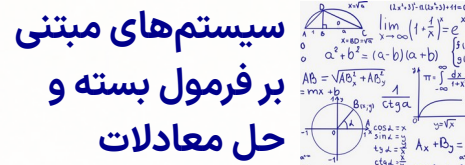


مرز شدت حکمرانی، میزان اثر سامانه بر تصمیم و اقدام است؛ نه صرفاً داشتن الگوریتم یا داشبورد.

کارکرد سامانه، به‌تنهایی هوش مصنوعی بودن آن را ثابت نمی‌کند



چه چیزهایی هوش مصنوعی نیستند؟



خودکارسازی فرایندها و اتوماسیون صنعتی در سطوح پایه هوش مصنوعی نیست.

داده‌های ذخیره‌شده در نوعی و پاسخ‌های از پیش تعیین‌شده یا برنامه‌ریزی‌شده هوش مصنوعی نیستند

مستقل از حجم و پیچیدگی هر برنامه‌ای که از یک فرمول مشخص استفاده کند و به ازای ورودی مشخص خروجی کاملاً مشخص بدهد هوش مصنوعی نیست

برنامه‌های رایانه‌ای هرچقدر هم که پیچیده باشند اگر اجرای صرف یک نگاشت یا دستور از پیش تعیین‌شده معمولاً هوش مصنوعی نیست

ماشین حساب
بازی رایانه‌ای
اتوماسیون صنعتی
مترجم قاعده‌مند

سیستم آبیاری خودکار
دستگاه‌های خط تولید
ربات‌های صنعتی برنامه‌ریزی‌شده
سیستم‌های کنترل روشنایی برنامه‌ریزی‌شده

پایگاه‌های دانش
سؤال و جواب‌های کوتاه (FAQ)
پایگاه قوانین
چت‌بات مبتنی بر سؤالات از پیش تعیین‌شده

تصورسازی رایانه‌ای
طراحی سه بعدی
واقعیت افزوده



خودکارسازی یا دیجیتال‌سازی فرایندها و اتوماسیون صنعتی در سطوح پایه هوش مصنوعی نیست.

هر آنچیزی که ثابت باشد یا تغییرات آن قابل پیش‌بینی و یا تصادفی باشد هوش مصنوعی محسوب نمی‌شود.

با وجودی که تصادف پایه اصلی هوش است اما سیستم‌هایی که صرفاً بر اساس تصادف خروجی بدهند هوش مصنوعی نیستند

اگر قواعد فعالیت از پیش مشخص و قابل حدس باشد مستقل از میزان پیچیدگی یا عمق هوش مصنوعی نیست

بازیکن خودکار شطرنج*
سامانه کنترل خط تولید برنامه‌ریزی‌شده
سیستم کنترل تردد مبتنی بر RFID
دیواره آتش بدون تحلیل

خدمات هوشمند ثبت‌نام!
خودکار هوشمند! چراغ هوشمند!
تلفن هوشمند! کتری هوشمند! ساعت هوشمند!
صنعت هوشمند! دولت هوشمند!

وب‌سایت‌های عمومی
تارنماهای خدماتی
پورتال‌ها
نمایشگرهای وضعیت

سامانه‌های قرعه‌کشی
بازی‌های کامپیوتری مبتنی بر تصادف
مولد رمز عبور

غربال اولیه قلمرو و شدت حکمرانی

ابتدا هوش مصنوعی بودن را تشخیص دهید؛ سپس کنترل را متناسب با کاربرد و پیامد تعیین کنید

۱ دروازه اول: آیا با سامانه AI روبه‌رو هستیم؟

آیا سامانه با یک مدل آموخته‌شده یا دانش‌محور،
از ورودی برای تولید خروجی استنباط می‌کند؟

خیر

خارج از فهرست AI

اما کنترل‌های IT، داده، امنیت، حریم خصوصی و
ایمنی همچنان برقرارند

بله یا نامشخص

۲ دروازه دوم: کاربرد مشخص سامانه را بررسی کنید

از راست به چپ، نخستین وضعیت منطبق را مبنا قرار دهید

۱ کاربرد ممنوع

توقف و ارجاع حقوقی

۲ پریسک یا جزء ایمنی

الزامات سخت‌گیرانه

۳ مشمول شفافیت

اطلاع‌رسانی یا نشانه‌گذاری

۴ سایر کاربردها

کنترل متناسب با ریسک سازمانی

شدت پیامد | مقیاس استفاده | حساسیت داده | میزان اتکای انسان | امکان بازگشت و جبران



هوش مصنوعی بودن فقط ورود به فرایند است؛ کاربرد و پیامد، شدت حکمرانی را تعیین می‌کنند.

کارت تحول ۲: واحدها از «دستیار» تعریف یکسانی ندارند

هر واحد، دامنه متفاوتی را تصور می‌کند



تا وقتی درباره معنای «دستیار» توافق نکنیم، درباره دامنه پروژه هم توافقی نداریم.



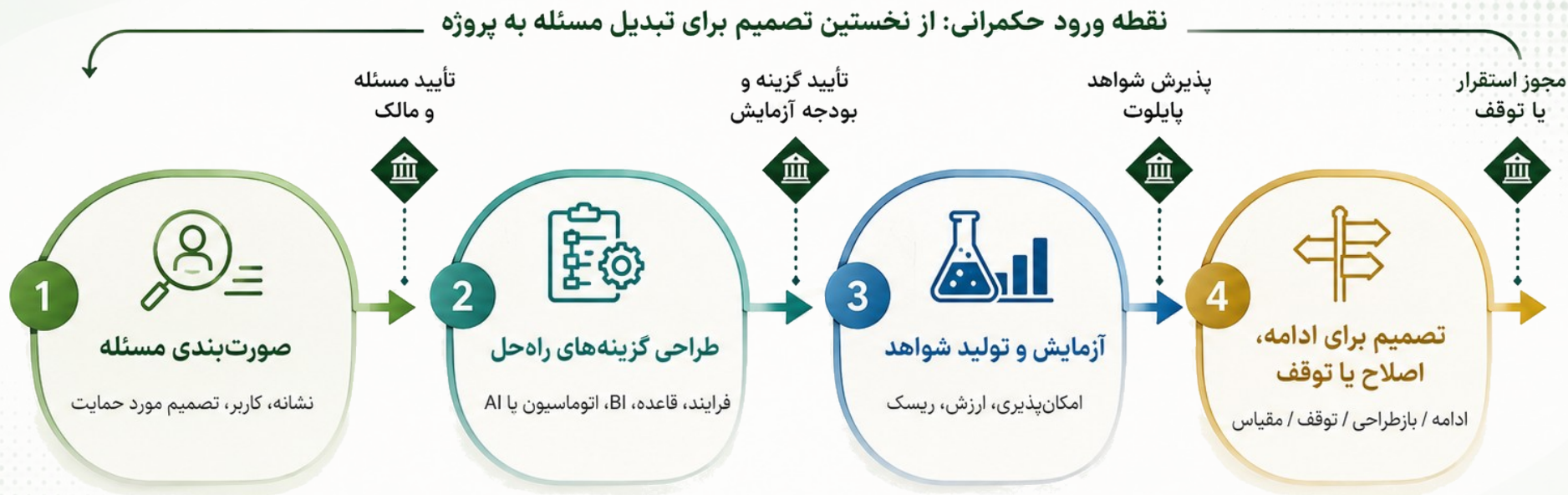
بخش دوم - قسمت دوم: حکمرانی مسئله و سبد سرمایه گذاری

فناوری آخرین حلقه حل مساله است



اگر شناخت سازمان و مساله را حذف کنیم، فناوری فقط پاسخ دقیق تری به سؤال اشتباه می دهد!

روش حل مساله و نقطه ورود حکمرانی



حکمرانی در پایان پروژه وارد نمی‌شود؛ از لحظه‌ای آغاز می‌شود که سازمان تصمیم می‌گیرد یک مساله را به پروژه تبدیل کند

همه مسئله‌ها به هوش مصنوعی نیاز ندارند

خوش فرم

- قابل حل از طریق جستجو، قاعده یا فرمول
- الگوریتم و هوش مصنوعی کلاسیک

بد فرم

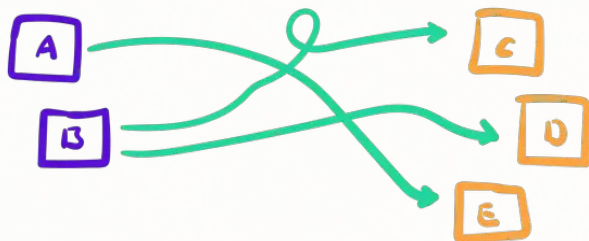
- بدون راهکار مستقیم
- نیازمند حل در ابعاد مختلف و با نگاه متفاوت
- مناسب هوش مصنوعی نوین

لاینحل!

Well-defined



Ill-defined



Wicked



تله‌های صورت‌بندی مسئله

Narrowing

چارچوب‌بندی مساله
به نحوی که فقط
حالات خاصی را تحت
پوشش قرار دهد

Hamstering

به صورت مصنوعی
مفروضاتی برای
مساله در نظر گرفته
شود

Solutioneering

چارچوب‌بندی مساله
به نحوی به یک
فناوری خاص وابسته
شود

Anchoring

چارچوب‌بندی مساله
به نحوی که یک
پاسخ خاص داشته
باشد

Catastrophizing

حل مسایلی که فعلاً
وجود ندارند با فرض
کردن بدترین حالات

Presuming

پیش‌فرض‌گیری در
مورد رفتار کاربر نهایی

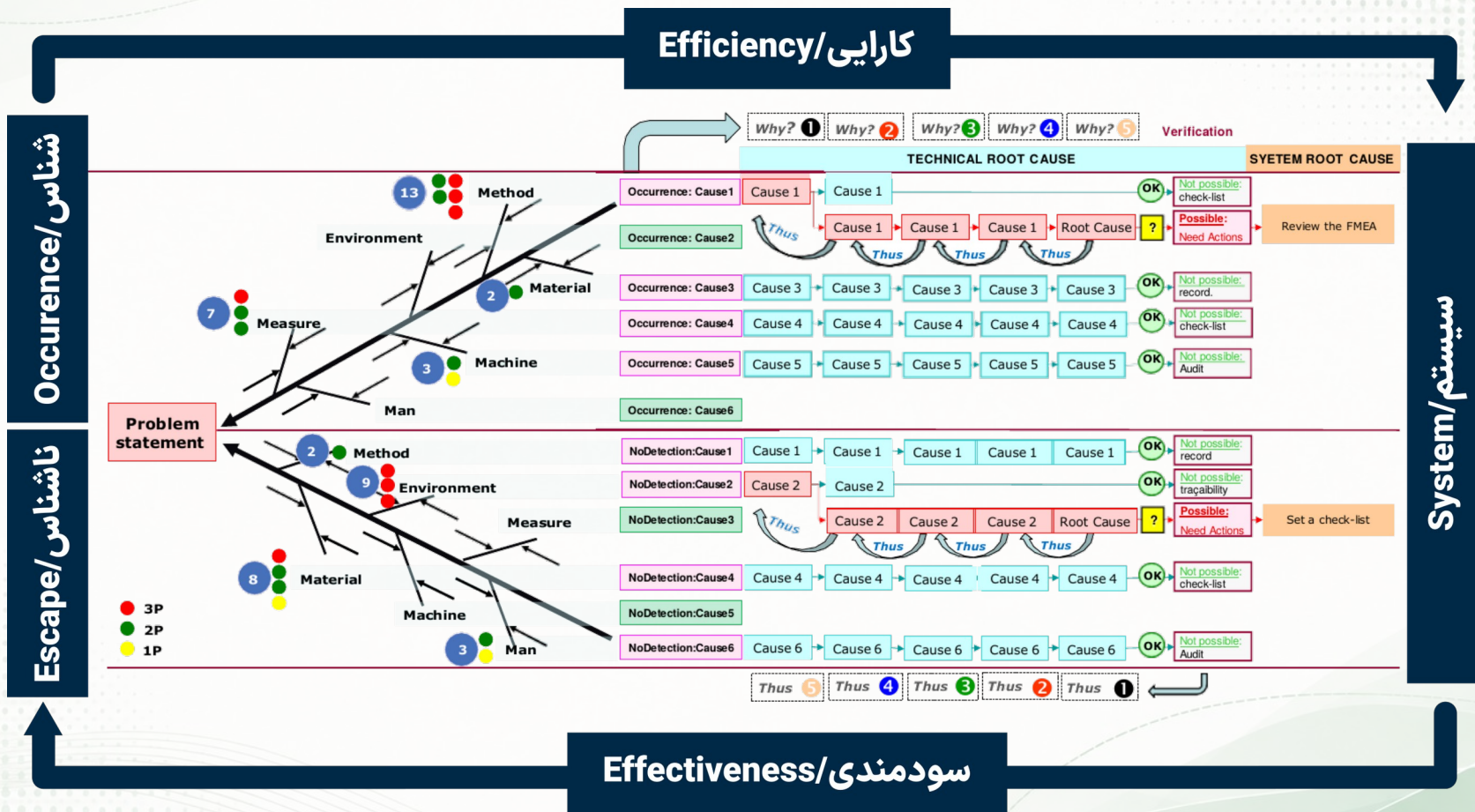
Buzzwording

چارچوب‌بندی مساله
مبتنی بر برخی
محصولات یا
سرویس‌های معروف

Wishlisting

چارچوب‌بندی مساله
به نحوی که مجموعه
ویژگی‌های دلخواه را
پوشش دهد

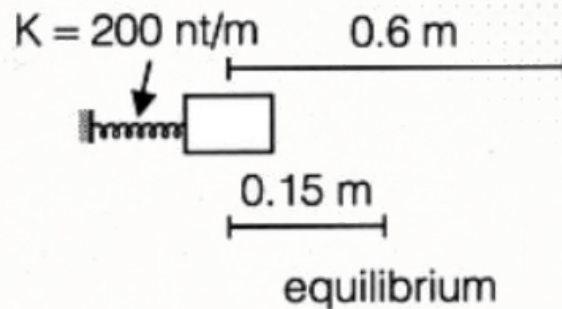
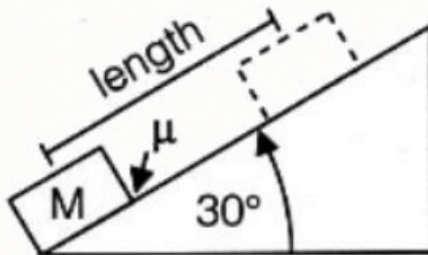
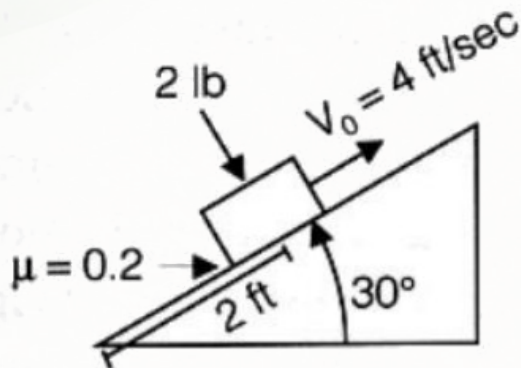
قاعده پیچیدگی مساله



Double Ishikawa and Naze Naze Analysis

یک مسئله، چند صورت بندی

زاویه نگاه، متغیرهای مهم و فضای راه حل را تغییر می دهد



نظر گروه آماتورها

نظر گروه متخصصان

مساله در مورد يك جعبه روی سطح شیبدار است
مساله در مورد اصطکاک روی سطح شیبدار است

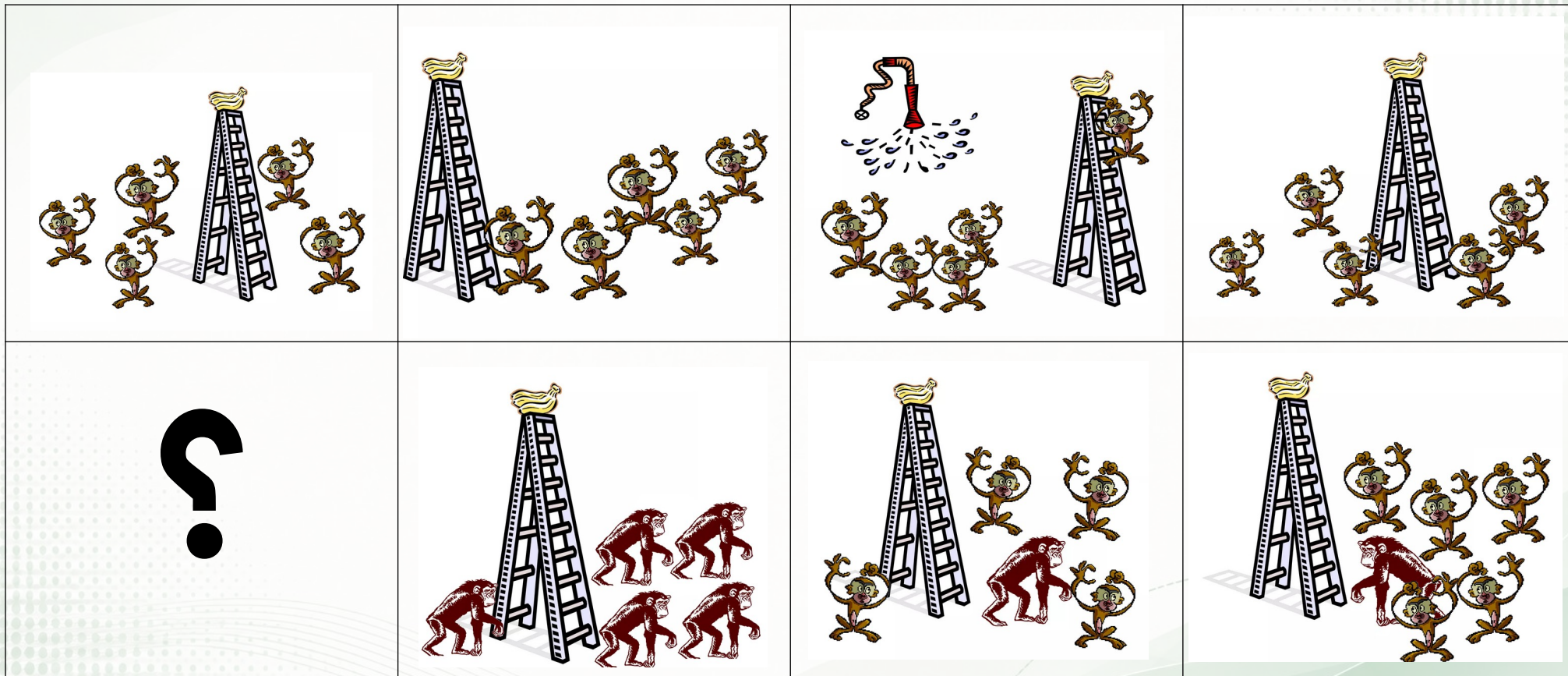
مساله انرژی پتانسیل است
مسایل مرتبط با انرژی پتانسیل فرمول مستقیم دارند



سازمان ها معمولاً بر سر راه حل اختلاف ندارند؛ ابتدا مسئله را به شکل های متفاوتی می بینند.

تمرین میمون: محدودیت واقعی کجاست؟

کدام محدودیت در صورت مسئله وجود دارد و کدام را خودمان به آن اضافه کرده ایم؟



کارت تحول ۳: مشکل فقط جستجو نیست؛ اسناد متناقض هستند

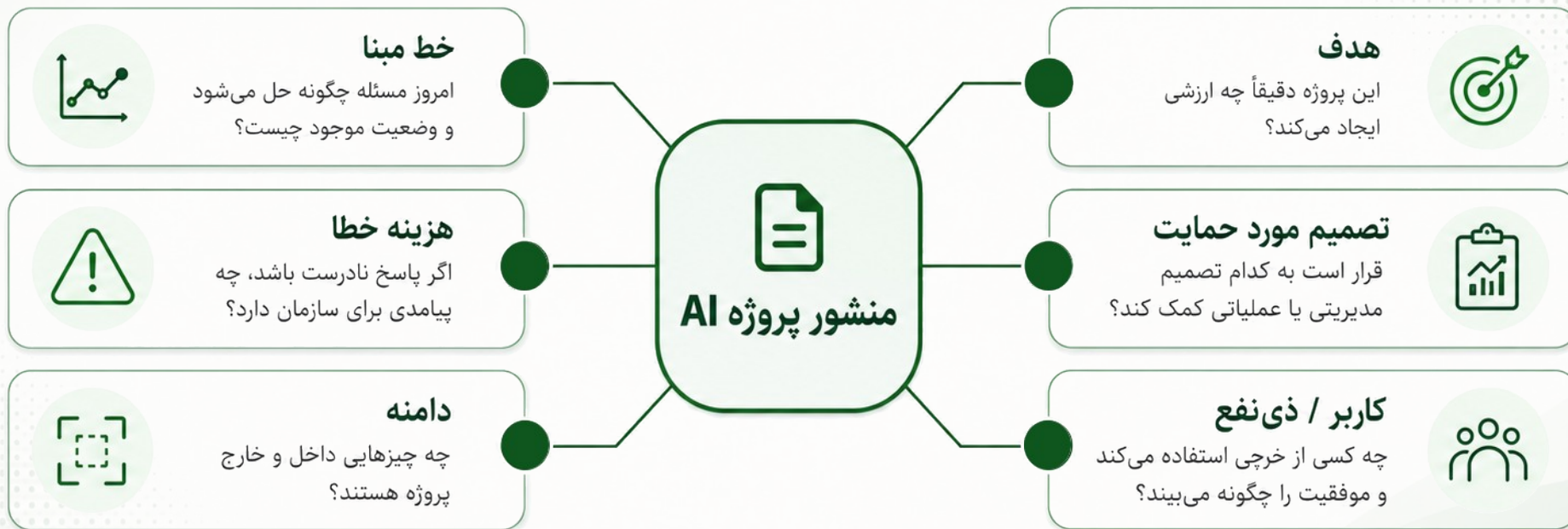
یک سؤال واحد، چند سند، چند پاسخ



صورت مسئله از «ساخت چت بات» به «حکمرانی دانش، نسخه اسناد و مالکیت محتوا» تغییر کرد.

منشور پروژه هوش مصنوعی باید چه چیزهایی را روشن کند؟

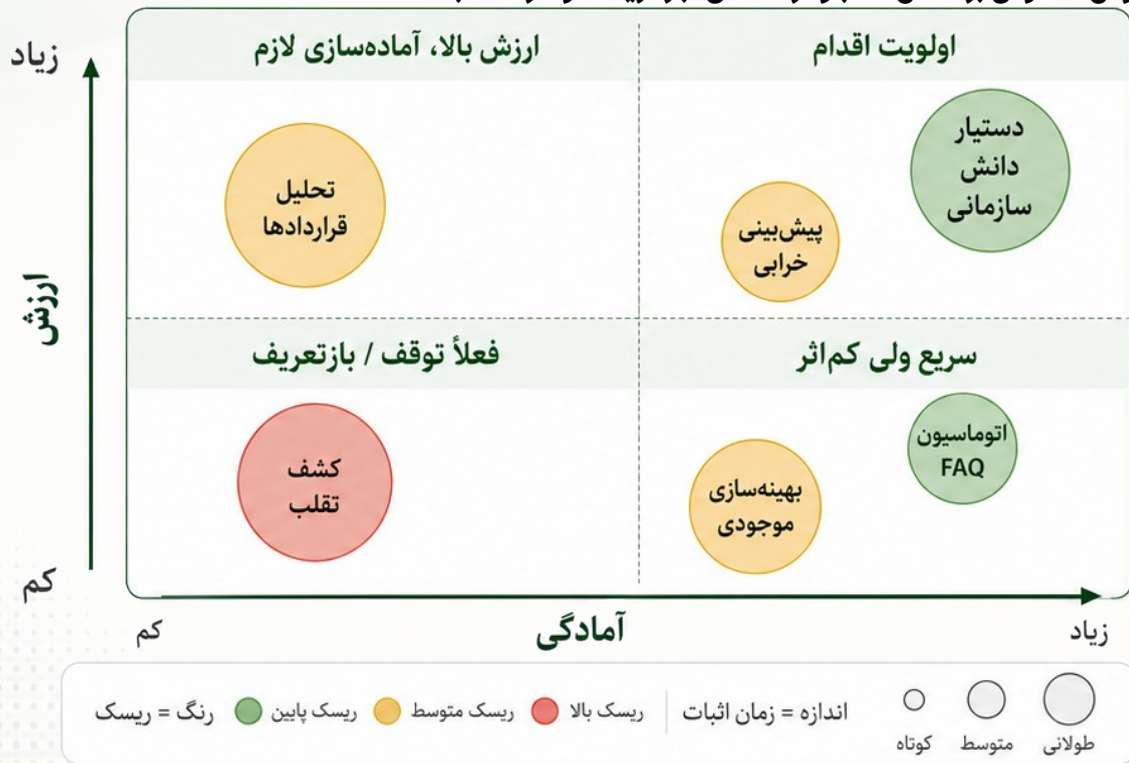
پیش از شروع هر پروژه، باید چند ابهام کلیدی به صورت صریح روشن بشوند



منشور پروژه، یک پروژه هوش مصنوعی را از یک ایده مبهم به یک مساله قابل حکمرانی تبدیل می‌کند.

سبد پروژه: ارزش، آمادگی، ریسک و زمان اثبات

اولویت‌بندی پروژه‌های هوش مصنوعی بر اساس کسب‌وکار، آمادگی اجرا، ریسک و سرعت اثبات



۱) ارزش

اثر بر درآمد، هزینه، کیفیت یا تصمیم



۲) آمادگی

داده، فرایند، مالک، دسترسی و پذیرش



۳) ریسک

حقوقی، عملیاتی، شهرتی و تغییر سازمانی



۴) زمان اثبات

سرعت رسیدن به نتیجه قابل اندازه‌گیری



پروژه مناسب فقط جذاب نیست؛

باید هم ارزش بسازد، هم آماده اجرا باشد، هم ریسک آن قابل مدیریت باشد و هم بتواند در زمانی معقول ارزش خود را اثبات کند.



پایلوت چه چیزی را باید اثبات کند؟

پایلوت برای نمایش صرف یا دمو نیست؛ باید درباره ادامه سرمایه‌گذاری تصمیم‌ساز باشد.



پایلوت خوب فقط نمی‌پرسد «کار می‌کند یا نه؟»؛ بلکه نشان می‌دهد «می‌ارزد؟ قابل بهره‌برداری است؟ آیا ارزش ادامه دادن دارد؟»

صورت مساله حالا روشن تر شده است

ما دیگر درباره یک چت بات حرف نمی زنیم؛ با یک مساله سازمانی روبه رو هستیم.



گام بعدی: ارزیابی اثر، سطح ریسک و میزان اختیار



دانشگاه آزاد اسلامی
واحد تهران مرکزی

دوره DBA هوش مصنوعی

بخش سوم:

ارزیابی اثر، سطح ریسک و اختیار



همه کاربردهای هوش مصنوعی هم سان نیستند

اثر، سطح ریسک و میزان اختیار، سطح حکمرانی را تغییر می دهد



پاسخ عمومی

اطلاع رسانی و محتوای عمومی



تحلیل

تحلیل داده و بینش



توصیه

پیشنهاد و راهنمایی



تصمیم اجرایی

تصمیم خودکار با اثر واقعی

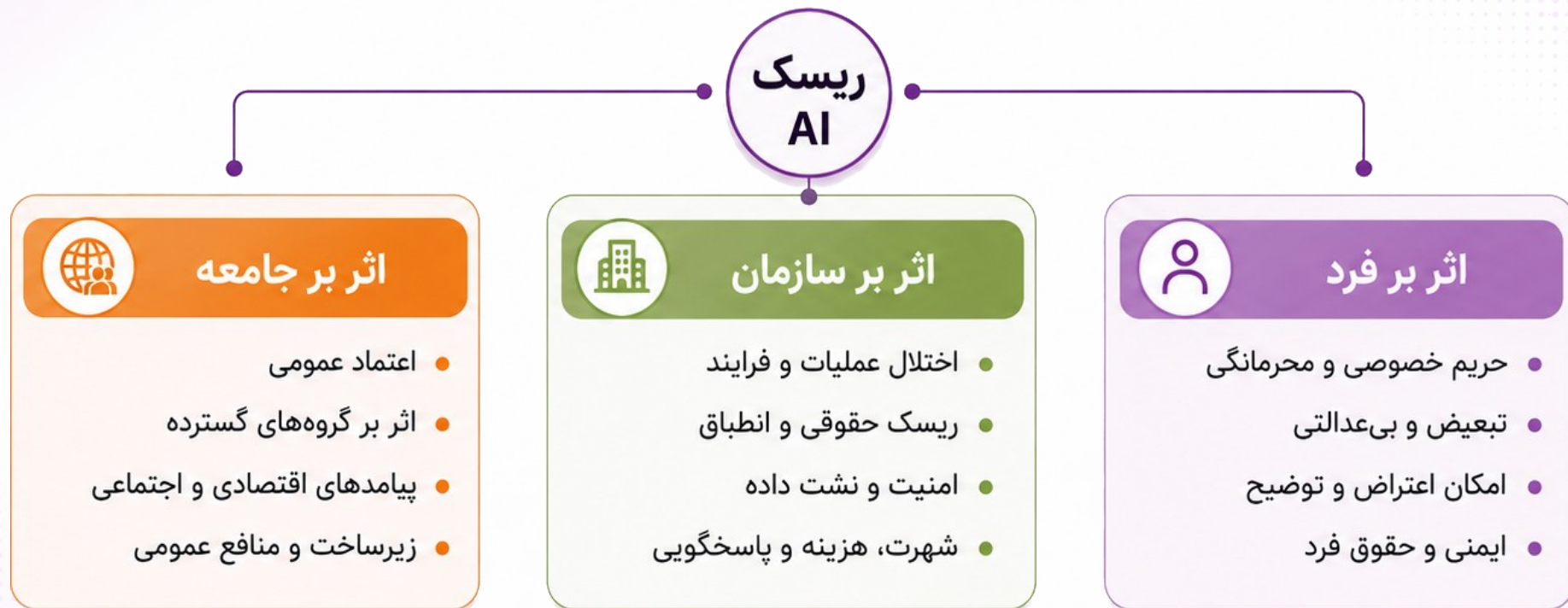


برای ارزیابی هوش مصنوعی باید فراتر از فناوری، به پیامد، اختیار و امکان کنترل نگاه کنیم.



ریسک فقط احتمال خطای مدل نیست

برای ارزیابی یک کاربرد هوش مصنوعی باید اثر آن را بر افراد، سازمان و جامعه ببینیم



خطای مدل فقط یکی از منابع ریسک است: داده، فرآیند، اختیار، کاربر و زمینه استفاده هم مهم هستند



ارزش خطا در کاربردهای مختلف یکسان نیست

معیار درست را خطای پر هزینه تعیین می کند، نه عادت ما به اعداد

Precision

فرد بی گناه نباید متهم شود

Recall

خطر واقعی نباید جا بماند



هرکدام از خطاها در مسایل مختلف ریسک متفاوت دارند و برخی ریسک ها جبران ناپذیرند



شدت، گستره، برگشت پذیری و امکان اعتراض

چهار پرسش ساده برای ارزیابی اثر یک کاربرد هوش مصنوعی



هرچه شدت و گستره بیشتر، و برگشت پذیری و امکان اعتراض کمتر باشد، سطح کنترل و نظارت باید بالاتر برود



از پاسخ عمومی تا تصمیم اجرایی

هرچه پاسخ به تصمیم حساس تر و اثرگذارتر نزدیک شود، سطح کنترل، شواهد و نظارت باید بیشتر شود



دستیار سازمانی اگر فقط پاسخ عمومی بدهد یک مساله دارد؛ اگر مبنای اقدام اجرایی شود، مساله ای کاملاً متفاوت خواهد داشت.



طبقات ریسک در هوش مصنوعی

هرچه اثر سامانه بر حقوق، ایمنی و تصمیم‌های مهم بیشتر باشد، سطح حکمرانی و الزامات نیز بالاتر می‌رود

۱ ریسک بالاترین		ریسک غیرقابل قبول ممنوع	سامانه‌هایی که حقوق بنیادین یا آزادی انسان را به‌طور جدی تهدید می‌کنند.	نمونه‌ها • امتیازدهی اجتماعی • دستکاری شناختی و رفتار • بهره‌کشی از آسیب‌پذیری افراد
۲ ریسک بالا		ریسک بالا الزامات سخت‌گیرانه	سامانه‌هایی که بر فرصت‌های مهم، ایمنی یا حقوق افراد اثر جدی می‌گذارند.	نمونه‌ها • زیرساخت‌های حیاتی • استخدام و منابع انسانی • آموزش و ارزیابی • اعتبارداری و خدمات ضروری • سلامت و تجهیزات پزشکی
۳ ریسک محدود		ریسک محدود الزام شفافیت	کاربر باید بداند با AI یا محتوای تولیدشده توسط AI مواجه است.	نمونه‌ها • چت‌بات‌ها • دستیارها • محتوای مولد
۴ ریسک حداقلی		ریسک حداقلی استفاده آزاد با کنترل‌های عمومی	کاربردهای کم‌خطر که معمولاً مشمول تعهدات ویژه نیستند.	نمونه‌ها • فیلتر هرزنامه • پیشنهادهای ساده • کاربردهای کمکی داخلی

برچسب «هوش مصنوعی» به تنهایی تعیین‌کننده نیست؛ سطح ریسک دامنه حکمرانی را تعیین می‌کند.



هرچه اختیار بیشتر، نیاز به شواهد و نظارت بیشتر

هرچه اثر سامانه بر حقوق، ایمنی و تصمیم‌های مهم بیشتر باشد، سطح حکمرانی و الزامات نیز بالاتر می‌رود

افزایش اختیار



افزایش شواهد و نظارت

**قاعده ساده: سامانه‌ای که فقط پاسخ می‌دهد، به کنترل سبک‌تر نیاز دارد؛
سامانه‌ای که اثر می‌گذارد یا اقدام می‌کند، به شواهد قوی‌تر و نظارت سخت‌گیرانه‌تر نیاز دارد**



ماتریس ریسک دستیار سازمانی

همه پاسخ‌های دستیار هم‌ریسک نیستند؛ سطح ریسک به موضوع، اثر و میزان اختیار بستگی دارد



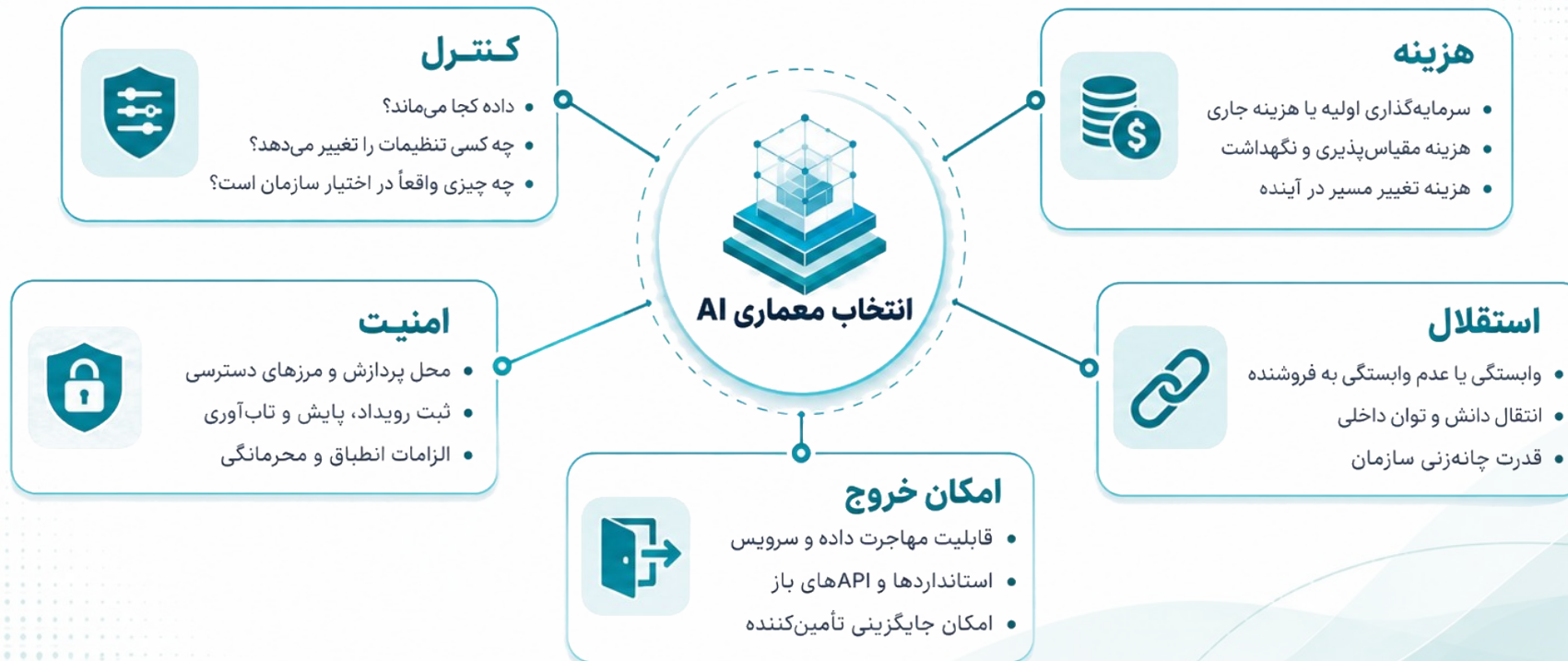
هرچه پاسخ به تصمیم نزدیک‌تر شود، نیاز به شواهد، نظارت و مسئولیت‌پذیری انسانی بیشتر می‌شود.



بخش چهارم: معماری، داده و تأمین کننده

معماری یک انتخاب فنی نیست

هر انتخاب معماری بر سطح کنترل، هزینه، امنیت، استقلال و امکان خروج اثر می‌گذارد.



معماری فقط انتخاب ابزار نیست؛ انتخاب میزان کنترل سازمان بر داده، تصمیم و آینده سامانه است



چهار رکن زیرساخت هوش مصنوعی و مالک هر رکن

زیرساخت هوش مصنوعی فقط سرور و مدل نیست؛ هر رکن مالک، مسئولیت و تصمیم‌های حاکمیتی متفاوتی دارد



همه این ارکان زیر مسئولیت حکمرانی سازمان هستند.

شکست بسیاری از پروژه‌های هوش مصنوعی ناشی از فناوری نیست؛ بلکه از جایی آغاز می‌شود که داده، مدل، زیرساخت یا کاربرد، مالک مشخص و مسئول پاسخ‌گو ندارد



هوش مصنوعی بخشی از یک سامانه بزرگتر است

هوش مصنوعی در خلأ کار نمی‌کند؛ در بستر فرایندها، داده‌ها، سیستم‌ها و افراد، خلق ارزش می‌کند.



برای موفقیت هوش مصنوعی باید کل سامانه دیجیتال را به صورت یکپارچه طراحی، مدیریت و پایش کنیم.

چه چیزی باید در مالکیت سازمان بماند

موفقیت و استقلال بلندمدت هوش مصنوعی به دارایی‌هایی وابسته است که اگر در اختیار سازمان نباشد، ریسک، هزینه و وابستگی ایجاد می‌کند



چه چیزهایی نباید به‌طور کامل خارج از سازمان قرار گیرد؟



هرآنچه مزیت رقابتی، کنترل، امنیت و استقلال تصمیم‌گیری سازمان را می‌سازد، باید در مالکیت و کنترل خود سازمان باقی بماند

ساخت، خرید، API، متن باز یا معماری ترکیبی؟

هیچ گزینه‌ای بهترین مطلق نیست؛ انتخاب درست به کنترل، هزینه، سرعت، امنیت و استقلال مورد نیاز سازمان بستگی دارد

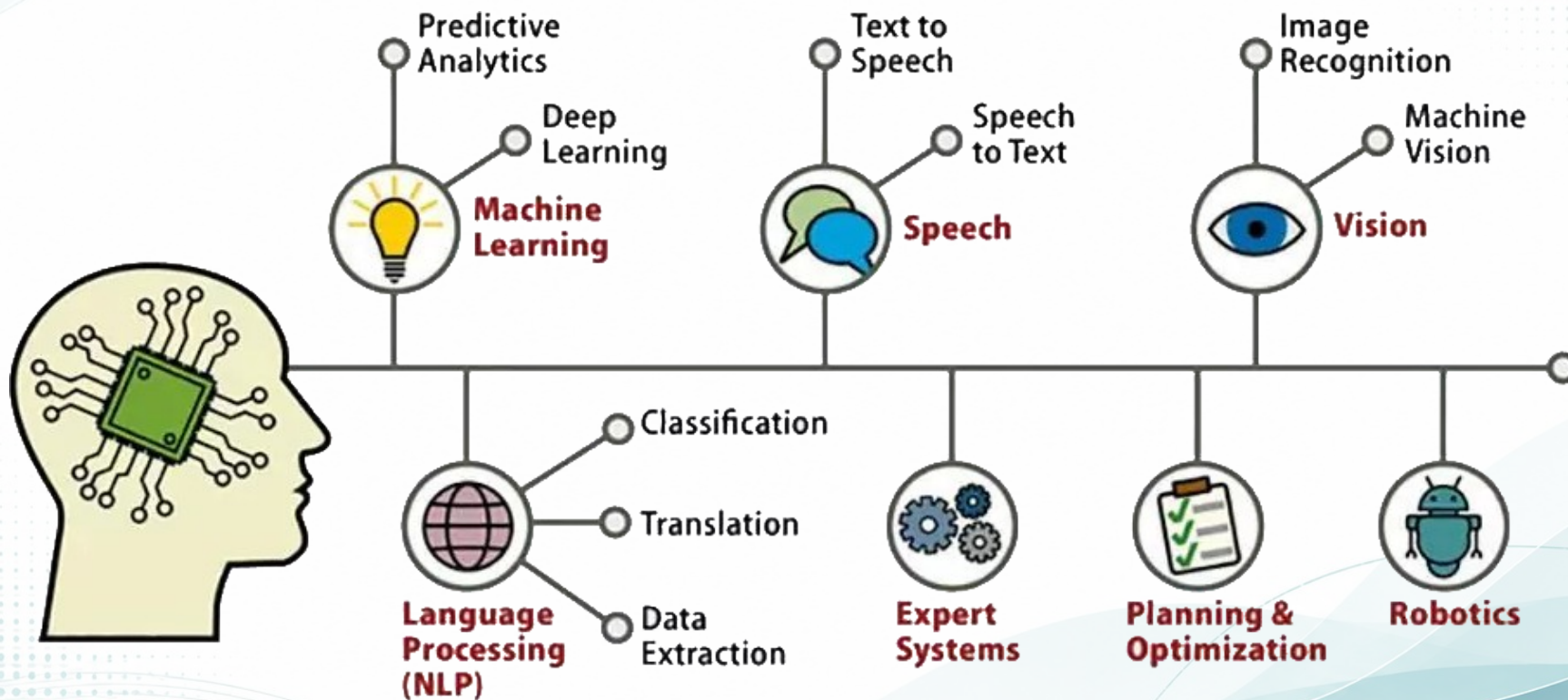
	ساخت (Build)	خرید (Buy)	API (API)	متن باز (Open Source)	معماری ترکیبی (Hybrid)
	مناسب برای: مزیت رقابتی سازمان، نیازهای کاملاً اختصاصی، محصولات AI	مناسب برای: نرم‌افزارهای آماده سازمانی، CRM، ERP، راهکارهای AI آماده	مناسب برای: LLM، پردازش گفتار، بینایی ماشین، ترجمه، استخراج داده و ...	مناسب برای: مدل‌های متن‌باز (مثل llama، Mistral، Llama) و ابزارها و پلتفرم‌های متن‌باز	ایده: ترکیب بهترین گزینه‌ها برای هر جزء بر اساس نیاز، ریسک و ارزش کسب‌وکار
سرعت استقرار	کند	متوسط تا سریع	بسیار سریع	متوسط	سریع تا بسیار سریع
هزینه اولیه	بالا	متوسط تا بالا	پرداخت به ازای مصرف	کم تا متوسط	بهینه و متوازن
کنترل و مالکیت	بسیار بالا	متوسط	کم	بالا	بالا
انعطاف و سفارشی‌سازی	بسیار بالا	متوسط	متوسط	بالا	بسیار بالا
امنیت داده و حریم خصوصی	بسیار بالا	متوسط	کم (وابسته به ارائه‌دهنده)	بالا (در استقرار داخلی)	بالا (قابل مدیریت)
وابستگی به دیگران	کم	زیاد	بسیار زیاد	متوسط	متوسط
امکان سفارشی‌سازی عمیق	بسیار بالا	کم تا متوسط	کم	بالا	بالا
نیاز به تخصص و منابع	بسیار زیاد	کم تا متوسط	کم	زیاد	متوسط تا زیاد
مزایا	کنترل کامل بر مدل و سامانه سفارشی‌سازی نامحدود استقلال و مالکیت کامل	هزینه ریز توسعه و نگهداری زمان توسعه طولانی نیاز به تیم تخصصی فنی	سرعت‌ترین راه برای شروع بدون نیاز به زیرساخت دسترسی به مدل‌های پیشرفته	نیاز به زیرساخت (GPU و ...) مسئولیت نگهداری و به‌روزرسانی هدف هزینه‌های مصرفی API	بهترین تعادل سرعت، هزینه و کنترل کاهش وابستگی و ریسک انعطاف و مقیاس‌پذیری بالا
معایب	هزینه ریز توسعه و نگهداری زمان توسعه طولانی نیاز به تیم تخصصی فنی	محدودیت در سفارشی‌سازی وابستگی به فروشنده هزینه لایسنس و تمدید	وابستگی به ارائه‌دهنده هزینه مصرفی در طول زمان محدودیت داده‌های حساس	نیاز به زیرساخت (GPU و ...) مسئولیت نگهداری و به‌روزرسانی نیاز به تیم تخصصی حساس	نیاز به زیرساخت و استاندارد کاهش وابستگی و ریسک محدودیت پیچیدگی بیشتر



سرعت استقرار	→	ساخت
کنترل و مالکیت	→	خرید
وابستگی به دیگران	→	API
نیاز به تخصص و منابع	→	ترکیبی

سؤال اصلی این نیست که «بهترین فناوری چیست؟»؛ بلکه این است که چه چیزی باید تحت کنترل سازمان بماند و چه چیزی را می‌توان برون‌سپاری کرد؟ معماری‌های موفق معمولاً از ترکیب هوشمندانه ساخت، خرید، API و متن‌باز به صورت هدفمند استفاده می‌کنند.

راه کارهای هوش مصنوعی غیر مولد

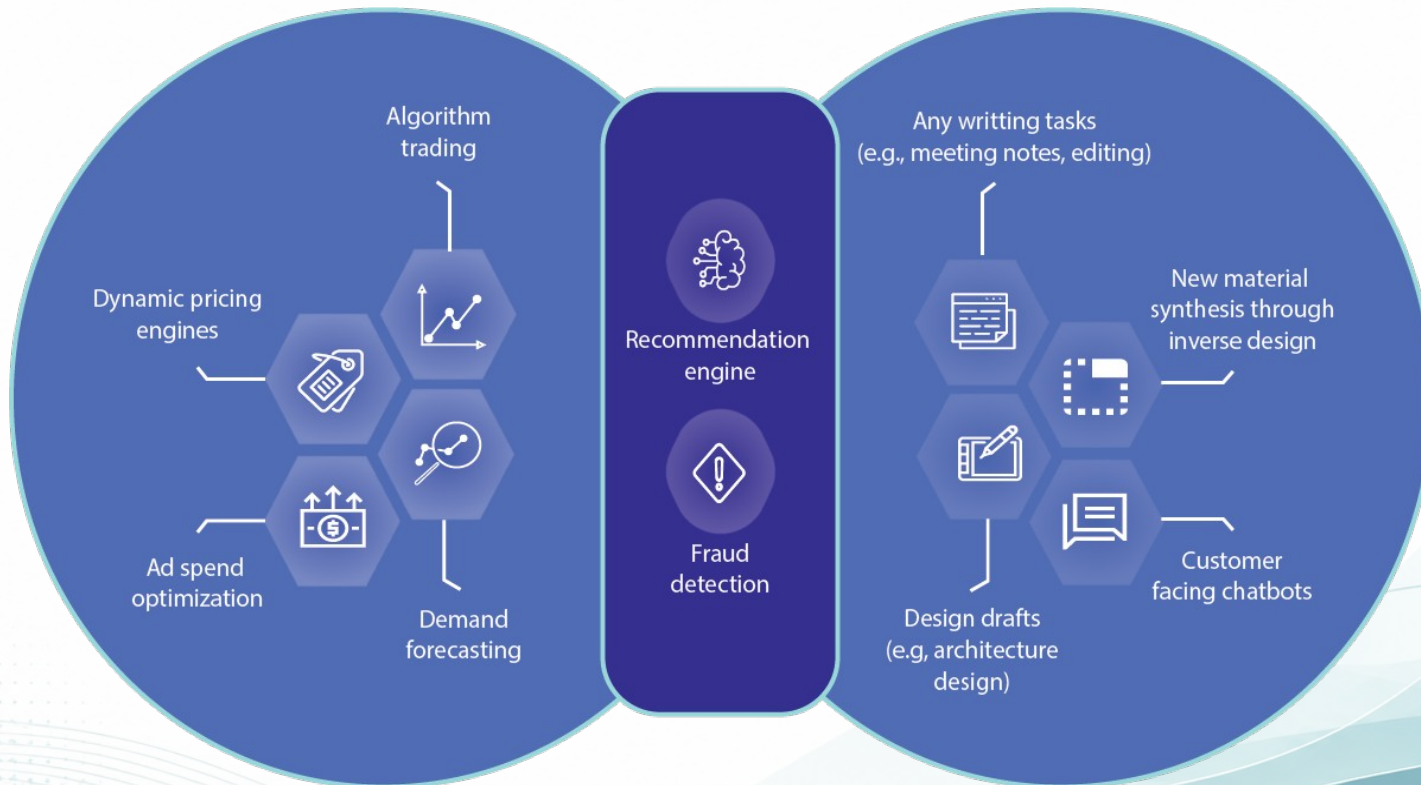


هوش مصنوعی مولد سر و صدای زیادی ایجاد کرده اما کاربردهای هوش مصنوعی غیر مولد به مرور به مرور بیشتر است



مقایسه هوش مصنوعی غیر مولد و مولد

هیچ گزینه‌ای بهترین مطلق نیست؛ انتخاب درست به کنترل، هزینه، سرعت، امنیت و استقلال مورد نیاز سازمان بستگی دارد



هوش مصنوعی غیر مولد

هوش مصنوعی مولد

مشکلات هوش مصنوعی مولد

هوش مصنوعی مولد جذاب، محیرالعقول اما با بالاترین سطح ریسک است



Misinformation



Hallucinations



Profanity Changes



Biased Outputs



Intellectual Property
Infringement



Privacy Concerns



Manipulation and
Impersonation



Unintended
Consequences



سه معماری پیشنهادی برای دستیار سازمانی

۱. معماری لایه‌ای (Layered)

سریع برای شروع، محدود در کنترل



- مزایا:
- راه‌اندازی سریع و هزینه کمتر
- مناسب کاربردهای عمومی



سطح کنترل پایین

- محدودیت‌ها:
- کنترل داده و جریان اطلاعات محدود
- ریسک خروج داده به بیرون

۲. معماری متمرکز (Centralized)

تعادل میان کنترل و سرعت



- مزایا:
- کنترل بهتر داده و دسترسی
- یکپارچگی با سیستم‌های سازمان



سطح کنترل متوسط

- محدودیت‌ها:
- پیچیدگی پیاده‌سازی متوسط
- نیاز به تیم فنی و زیرساخت

۳. معماری فدراتیو (Federated)

بیشترین کنترل و انطباق



- مزایا:
- انطباق کامل با سیاست‌ها و مقررات
- کنترل محلی داده‌ها در هر واحد



سطح کنترل بالا

- محدودیت‌ها:
- پیچیدگی و هزینه پیاده‌سازی بالا
- نیاز به بلوغ سازمانی و هماهنگی قوی



انتخاب معماری به اهداف کسب‌وکار، سطح ریسک، الزامات انطباقی و بلوغ دیجیتال سازمان بستگی دارد.



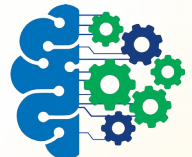
دانشگاه آزاد اسلامی
واحد تهران مرکزی

دوره DBA هوش مصنوعی

پیرانتز فنی:

پشت صحنه دستیار هوشمند سازمانی



**BIGDATA****IOT****AI / ML****Digital Twin**

چهار مفهوم نزدیک، اما نه هم معنا

کلان داده: داده‌هایی که به دلیل حجم، سرعت، تنوع یا پیچیدگی، به معماری‌ها و ابزارهای مقیاس‌پذیر نیاز دارند.

اینترنت اشیا: مجموعه‌ای از حسگرها، عملگرها و پردازنده‌ها که از طریق شبکه محلی یا اینترنت به یکدیگر متصل هستند.

هوش مصنوعی و یادگیری ماشین: الگوریتم‌ها و سیستم‌هایی که توانایی یادگیری و اجرای وظایف پیچیده را دارند.

همزاد دیجیتال: نمایشی رقومی و مجازی از یک سیستم فیزیکی که با آن همگام شده است.

کلان داده فقط «داده زیاد» نیست

مجموعه‌ای عظیم از داده‌های گردآوری‌شده از منابع مختلف که در سه حوزه حجم، تنوع و سرعت رشد در مقیاس عظیم باشند و شش ویژگی زیر به آن‌ها مترتب است:

VOLUME



حجم

حجم داده‌های
تشکیل‌دهنده
کلان‌داده

VARIETY



تنوع

انواع داده‌های موجود
در کلان‌داده اعم از
ساخت‌یافته، خام، سند،
چندرسانه‌ای و ...

VELOCITY



سرعت

سرعت ایجاد یا
تغییرات در کلان‌داده

VERACITY



اطمینان

میزان دقت و
اطمینانی که به صحت
داده‌ها وجود دارد

VALUE



ارزش

ارزش داده‌های
خلق‌شده برای
کسب‌وکار هدف

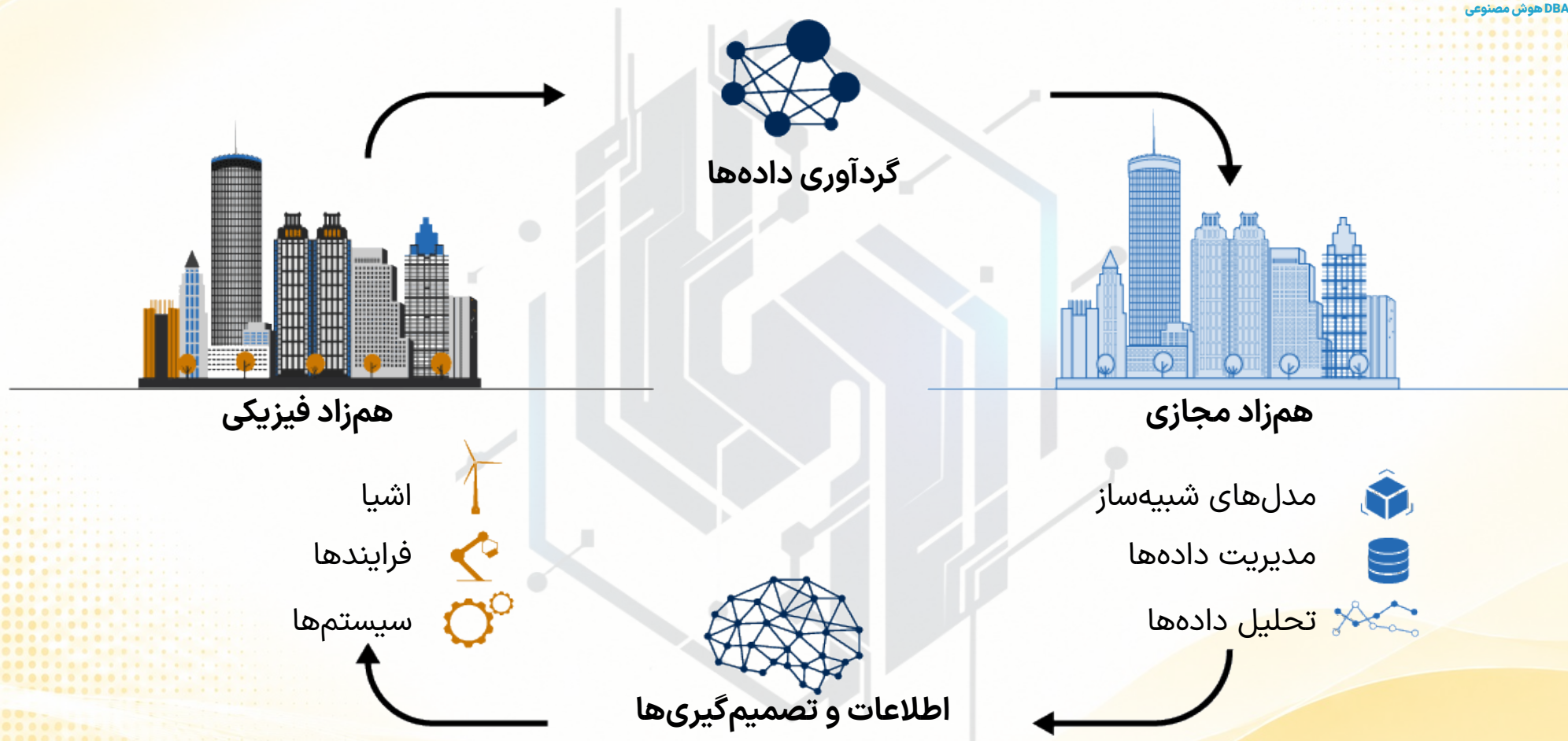
VARIABILITY



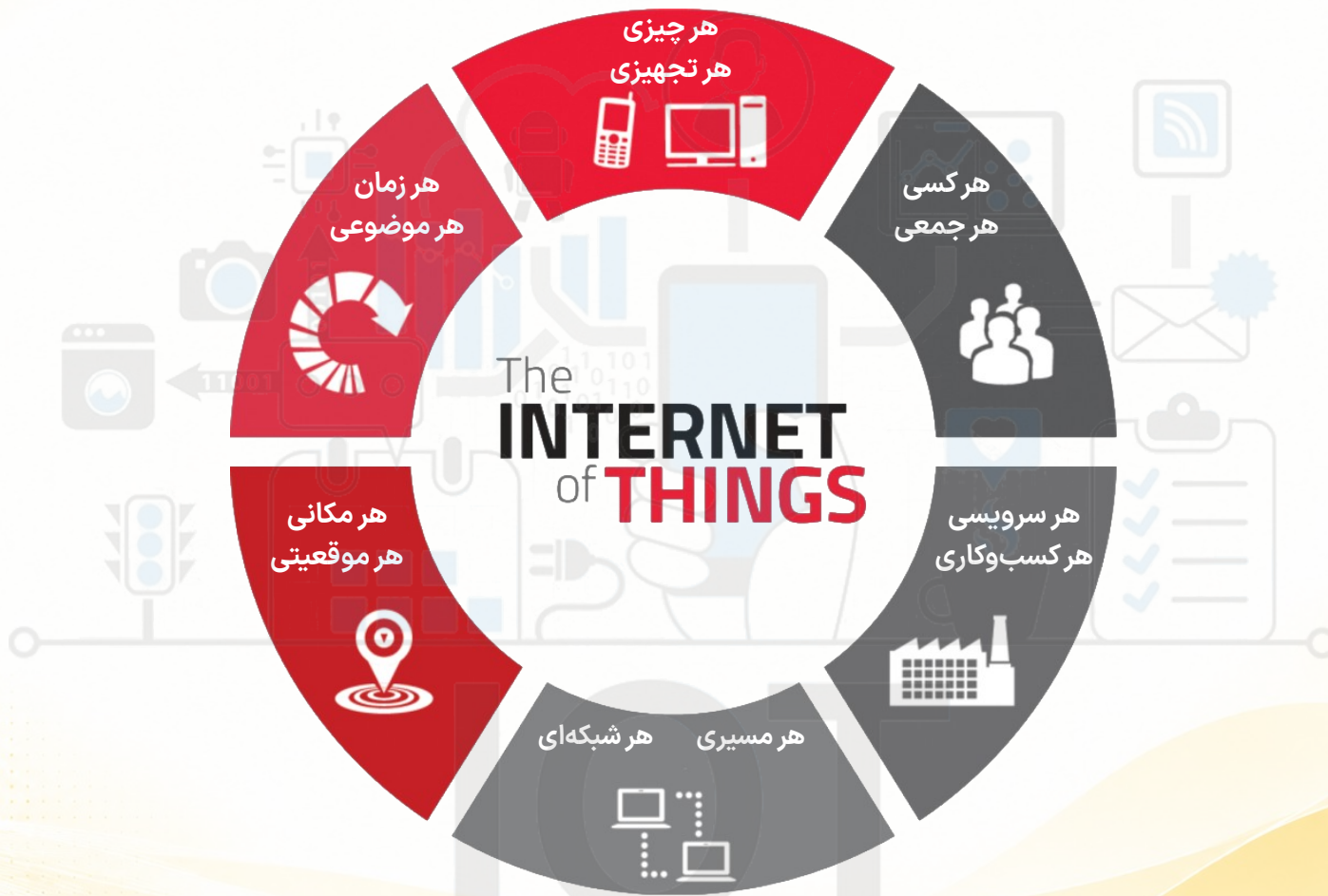
تغییرپذیری

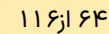
روش‌هایی که می‌توان
از کلان‌داده‌ها بهره‌برد

همزاد دیجیتال؛ پیوند داده، مدل و سیستم فیزیکی



از اینترنت اشیا تا اینترنت همه چیز





علم داده، هوش مصنوعی و یادگیری ماشین؛ مرزها کجاست؟

یادگیری ماشین

شاخه‌ای از هوش مصنوعی که از داده‌ها و الگوریتم‌ها برای ایجاد توانمندی یادگیری در ماشین‌ها استفاده کرده و آموخته‌های ماشین در گذر زمان و با داده‌های جدید افزایش می‌یابد



هوش مصنوعی

یکی از شاخه‌های علوم کامپیوتر که در آن سعی می‌شود ماشین‌ها همانند انسان‌ها فکر و عمل کنند و در نهایت به عنوان یک عامل خودکار خودشان تصمیم بگیرند

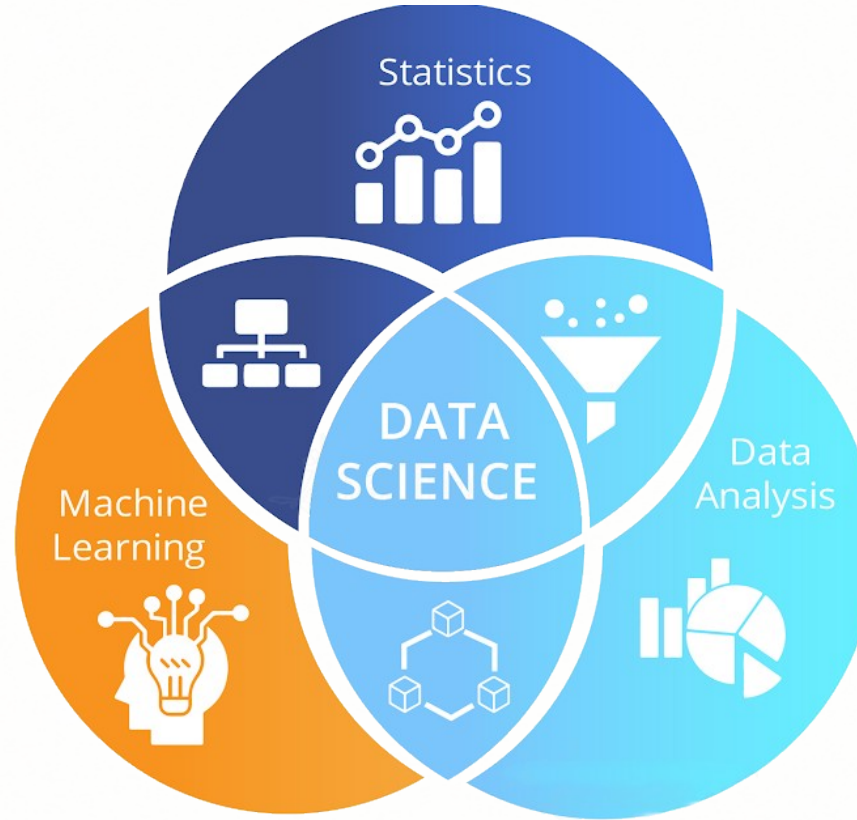


علم داده

فرایند استخراج برداشت‌های بامعنی و کاربردی از داده‌های حجیم خام (کلان داده) برای حل مسایل واقعی و توسعه راهکارهای تجاری و عملیاتی



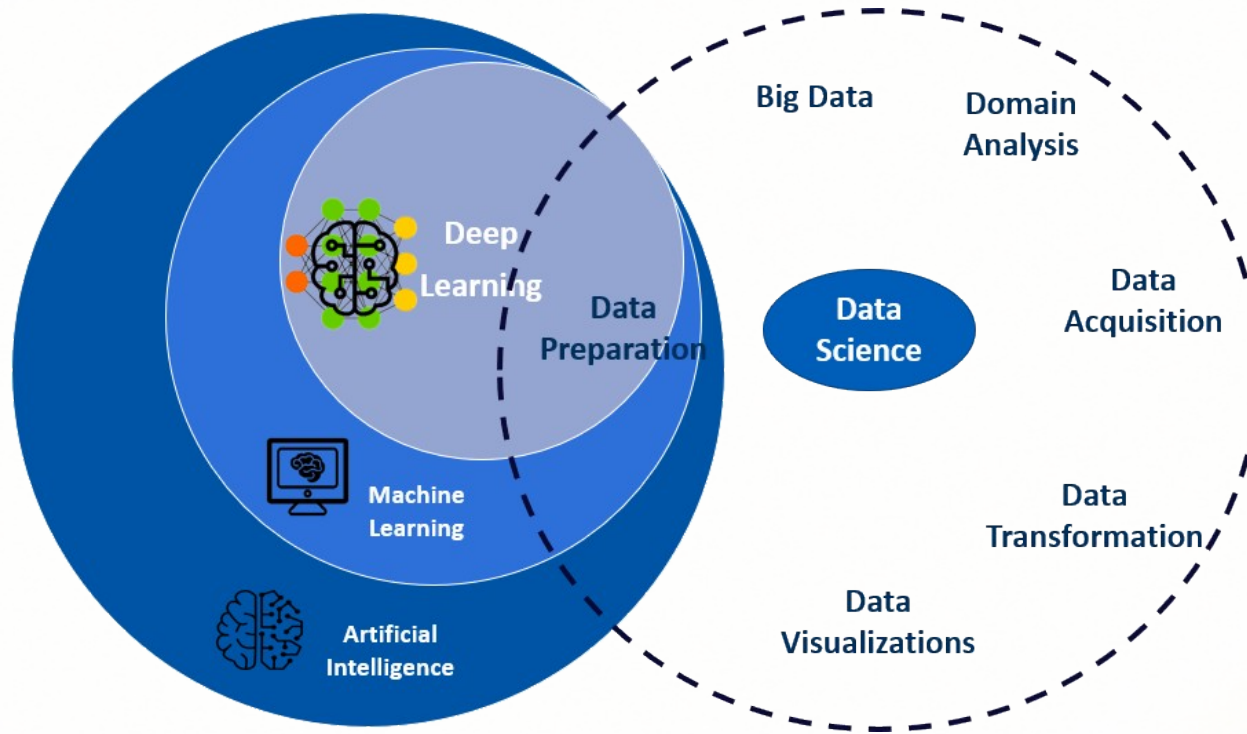
علم داده از هم‌پوشانی چند قابلیت شکل می‌گیرد



علم داده بدون آمار، تحلیل، مهندسی داده و فهم مسئله کسب‌وکار ناقص است.



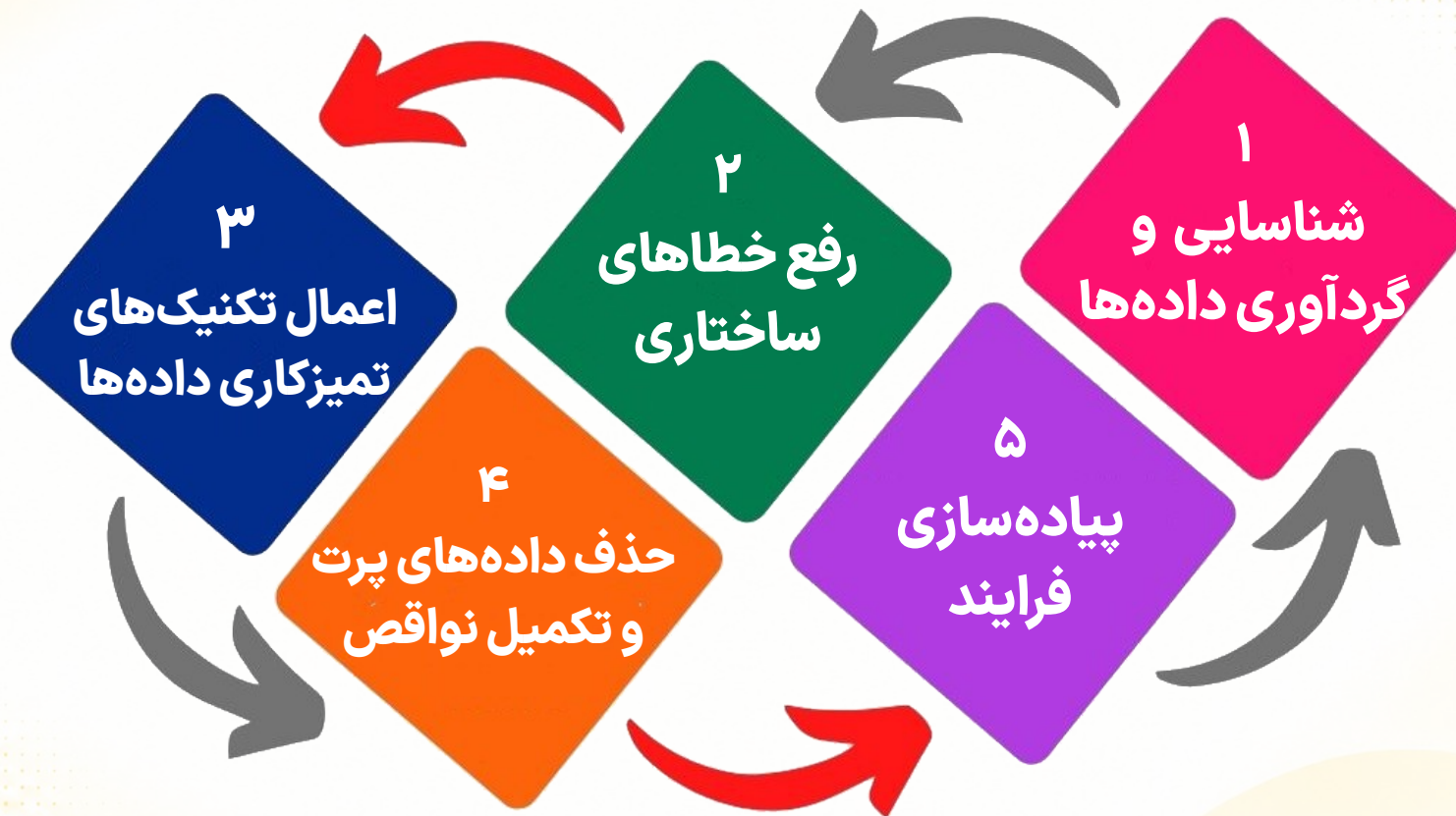
رابطه علم داده با ML، AI و DL



این حوزه‌ها هم‌پوشان‌اند، اما مترادف نیستند و هر پروژه به ترکیب متفاوتی از آنها نیاز دارد.



از داده خام تا داده قابل استفاده



کیفیت مدل از جایی آغاز می‌شود که هنوز هیچ مدلی وجود ندارد



پاک سازی، یکپارچه سازی و آماده سازی داده



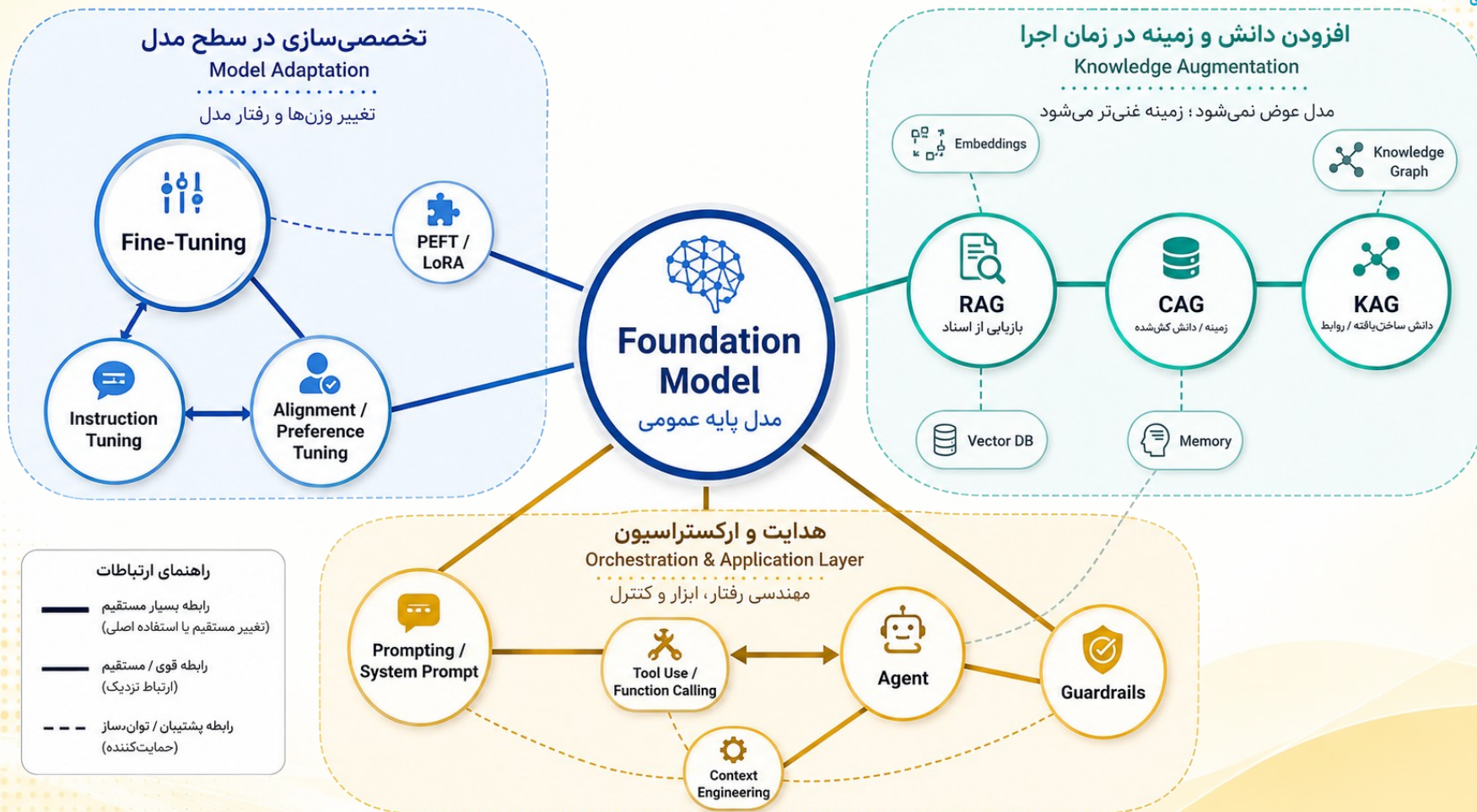
داده پاک، لزوماً داده قابل اتکا نیست



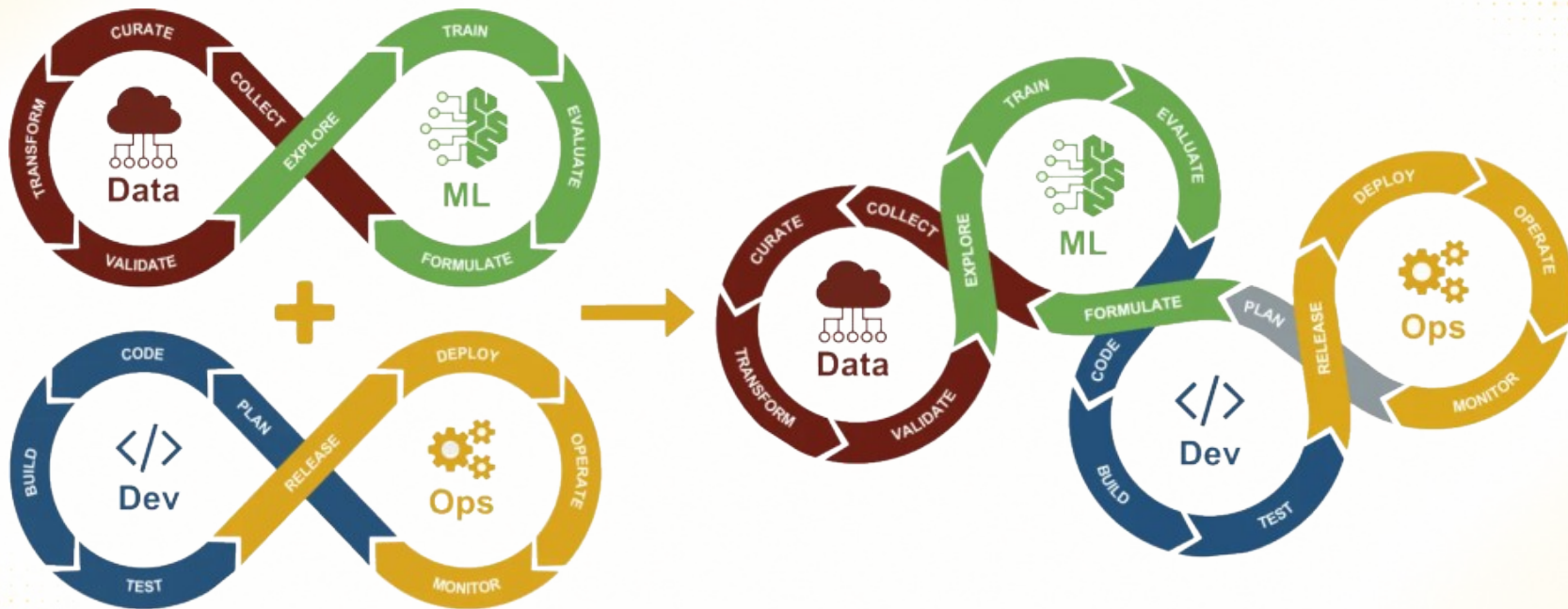
داده ممکن است تمیز باشد، اما قدیمی، غیرمجاز، بی مالک یا متعلق به نسخه اشتباه باشد



چگونه مدل پایه را برای کاربرد سازمانی آماده می کنیم؟



MLOps؛ حلقه اتصال داده، مدل، توسعه و عملیات

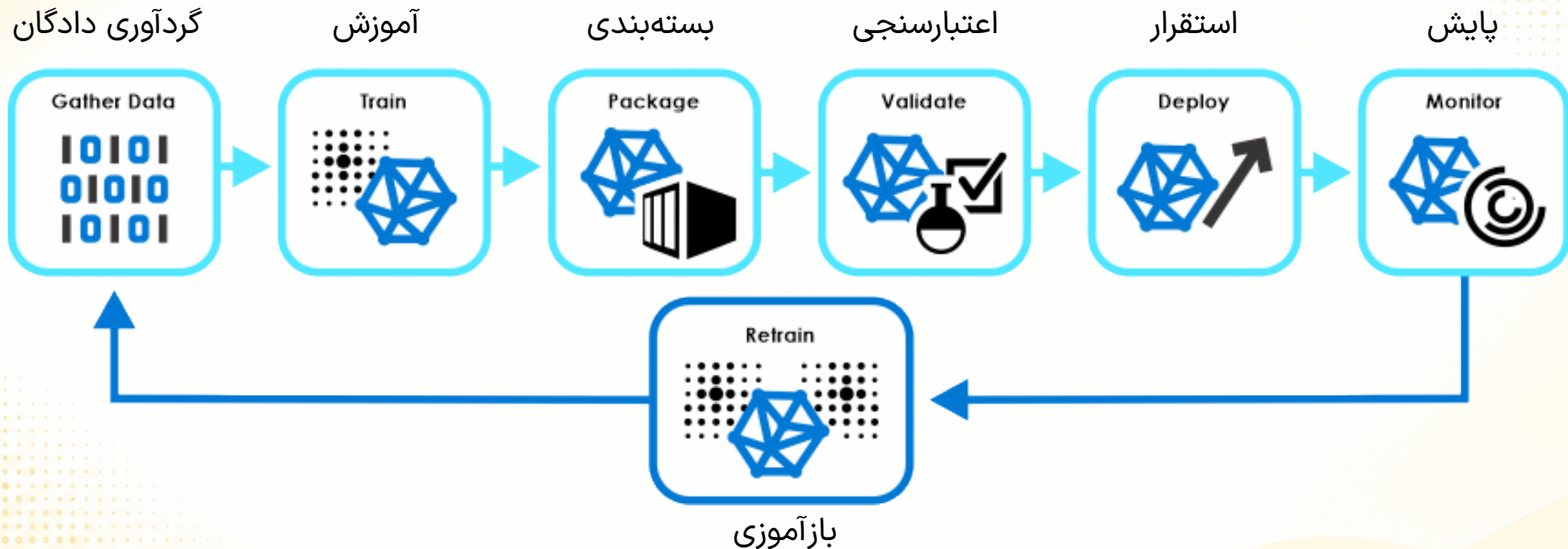


$$\text{MLOps} = \text{DataML} + \text{DevOps}$$

MLOps پیوند چرخه داده، مدل، توسعه نرم افزار و عملیات است.



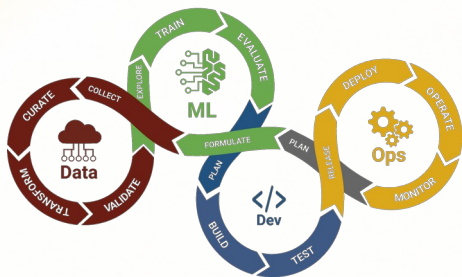
مدل یک بار ساخته نمی شود: چرخه MLOps



استقرار پایان چرخه نیست؛ پایش، بازاعتبارسنجی و در صورت نیاز بازآموزی، بخشی از خدمت هستند.



زیرساخت هوش مصنوعی فقط GPU نیست



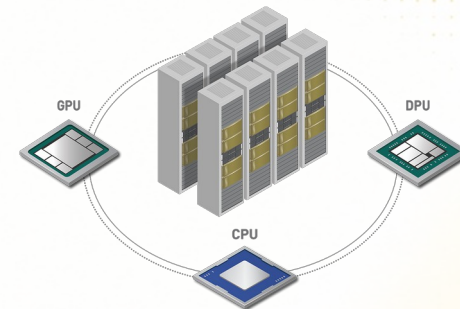
زیرساخت مهندسی هوش



زیرساخت یادگیری ماشین



زیرساخت داده‌ای

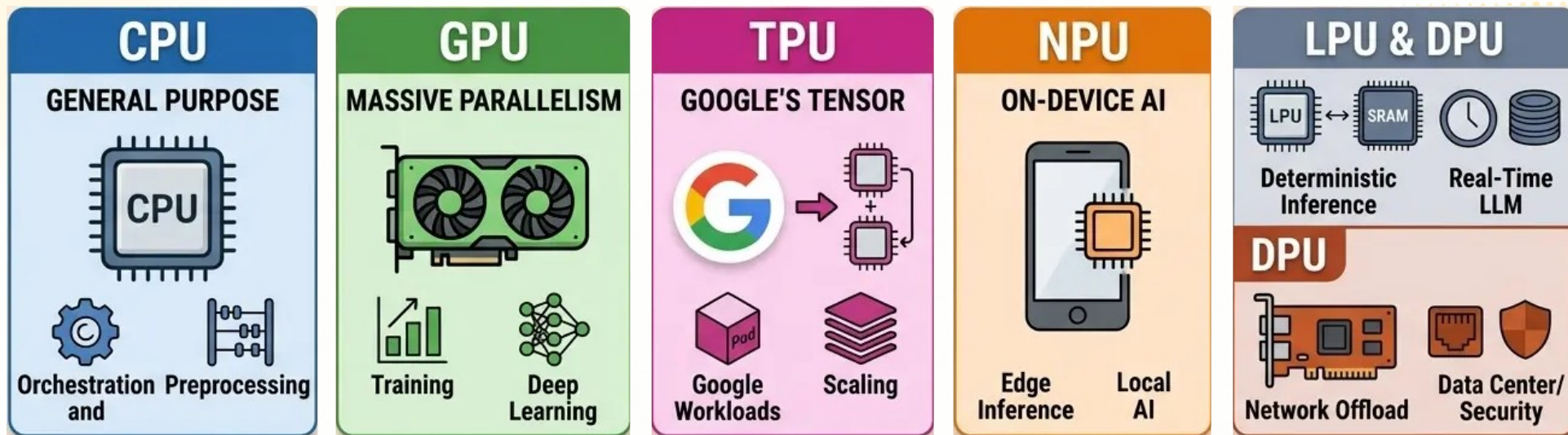


زیرساخت پردازشی

زیرساخت داده، پردازش، یادگیری و مهندسی باید همزمان و متناسب با بارکاری طراحی شوند.



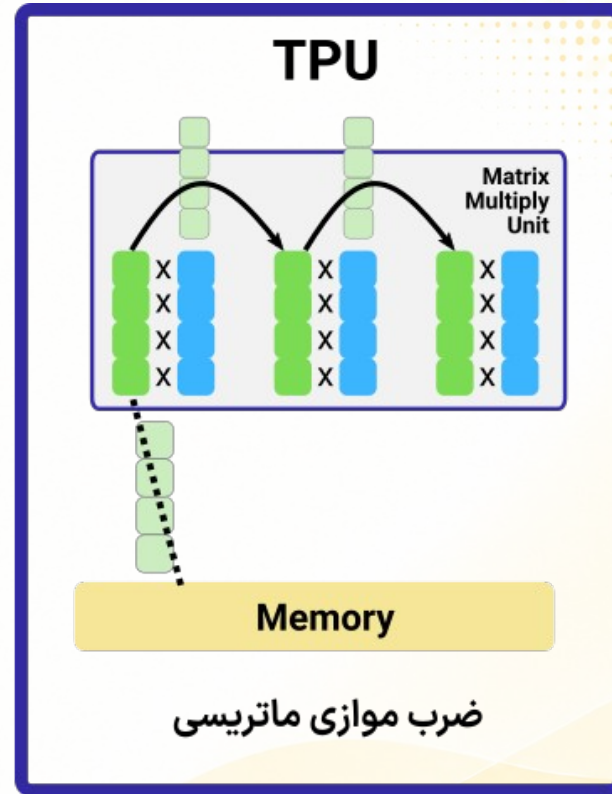
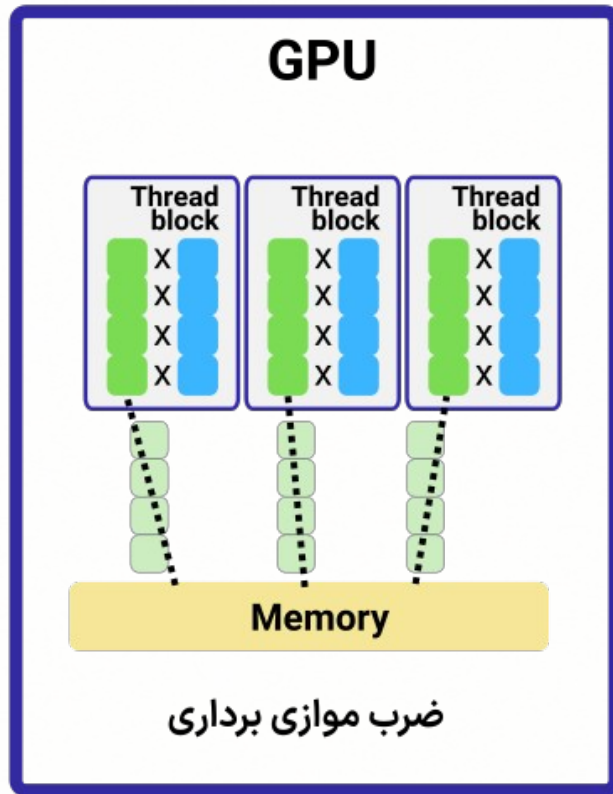
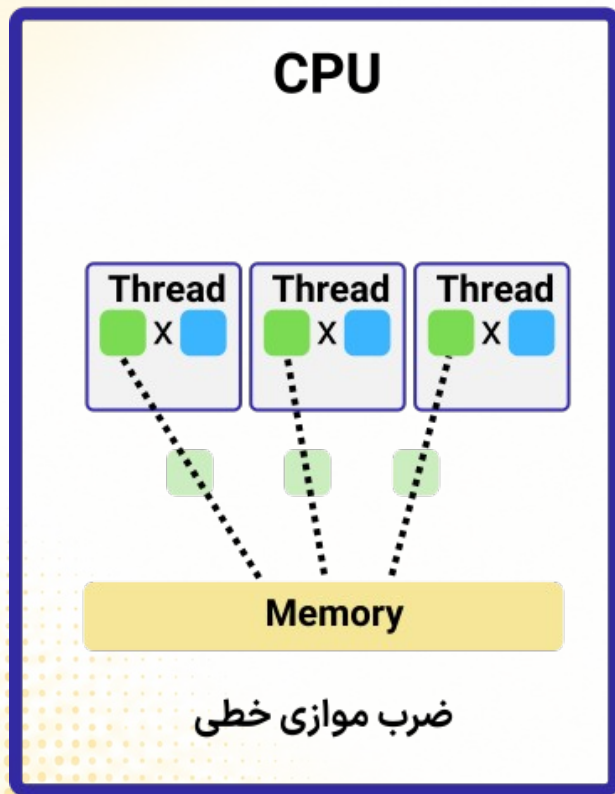
هر پردازنده برای یک بارکاری



پردازنده برتر مطلق وجود ندارد؛ تناسب با نوع عملیات، تأخیر، مقیاس، مصرف انرژی و هزینه تعیین کننده است.



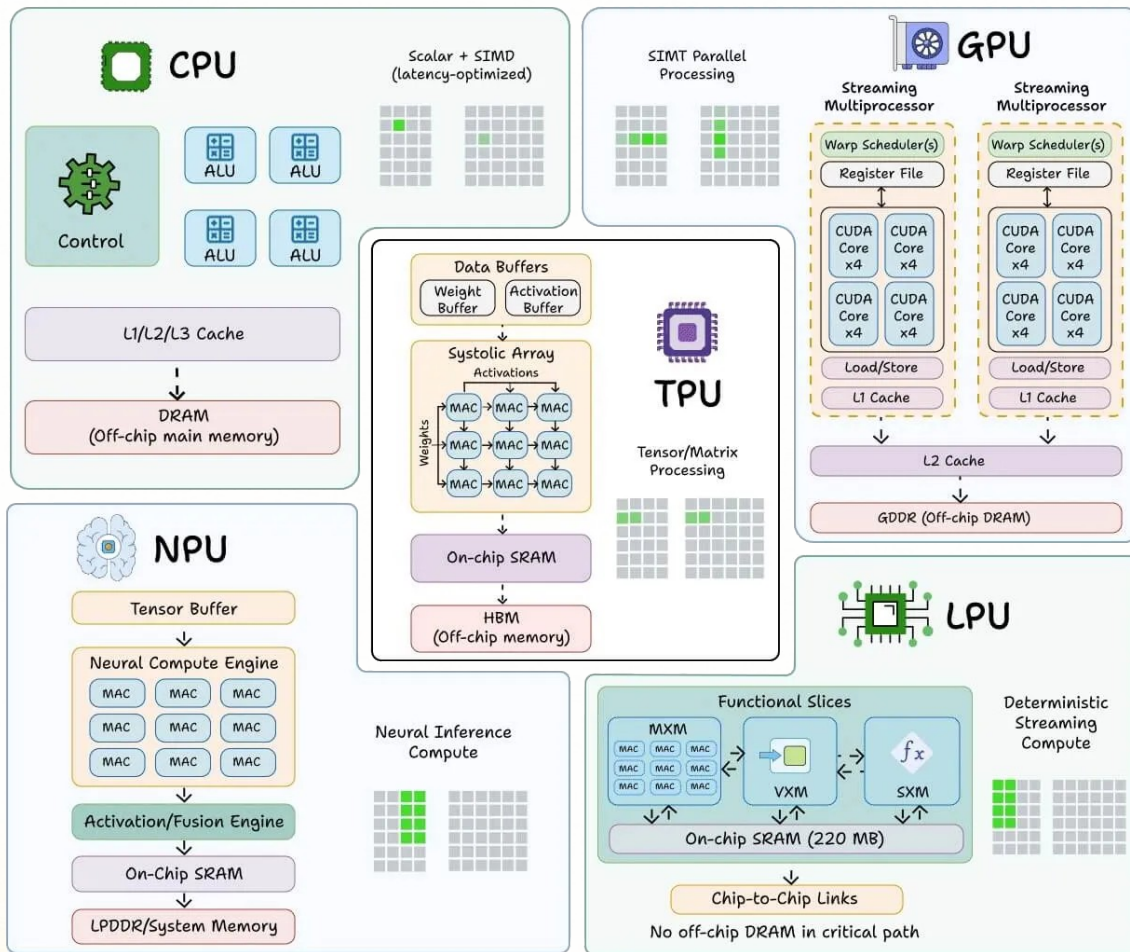
CPU، GPU و TPU؛ تفاوت در الگوی پردازش



تفاوت اصلی در نوع موازی سازی و جابه جایی داده است؛ نه اینکه یکی همیشه از دیگری بهتر باشد.



درون پردازنده‌ها چه می‌گذرد؟



آموزش و استنتاج دو مساله زیرساختی متفاوت هستند

آموزش معمولاً بار سنگین و دوره‌ای می‌خواهد، استنتاج با تأخیر پاسخ، تعداد کاربران همزمان، تداوم خدمت و هزینه هر درخواست سنجیده می‌شود.

آموزش مدل



- بار پردازشی سنگین و معمولاً دوره‌ای
- تمرکز بر توان عملیاتی و زمان تکمیل آموزش
- نیاز احتمالی به چند GPU و ارتباط پرسرعت
- وابستگی جدی به حجم داده و فضای ذخیره‌سازی
- تولید وزن‌ها، checkpoint، لاگ و artifacts

امکان اجاره ظرفیت موقت برای اوج مصرف

استنتاج و ارائه خدمت



- بار پیوسته یا متغیر بر اساس مصرف
- تمرکز بر تأخیر پاسخ و کاربران همزمان
- دسترسی پذیری، افزونگی و بازبازی خدمت
- مقیاس پذیری خودکار و نزدیکی به کاربر یا داده
- هزینه هر پاسخ، هر توکن یا هر تراکنش

↔ ممکن است یک سخت‌افزار واحد هر دو کار را انجام دهد، اما معیار طراحی آن‌ها یکسان نیست.

زیرساخت آموزش برای تولید و تغییر مدل طراحی می‌شود؛ زیرساخت استنتاج برای ارائه پایدار، سریع و اقتصادی خدمت

مدل‌های تأمین و تخصیص ظرفیت پردازشی

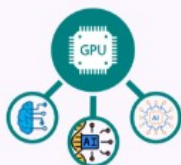
زیرساخت اختصاصی، ابری، مجازی و اشتراکی از نظر هزینه، کنترل، جداسازی و مقیاس‌پذیری یکسان نیستند.

سرورهای اختصاصی به صورت تکی و خوشه‌های پردازشی
ماشین‌های مجازی یا ابری با پردازنده گرافیکی اختصاصی به صورت تکی یا خوشه‌ای
مرکز داده اختصاصی به صورت مجازی شامل تمامی امکانات پردازشی و شبکه به صورت مدیریت‌شده

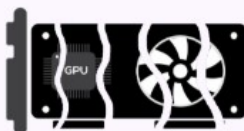
- **Dedicated**
- **VM/Cloud**
- **VPDC**



GPUDN



Shared GPU



Sliced GPU



VPDC



VM/Cloud



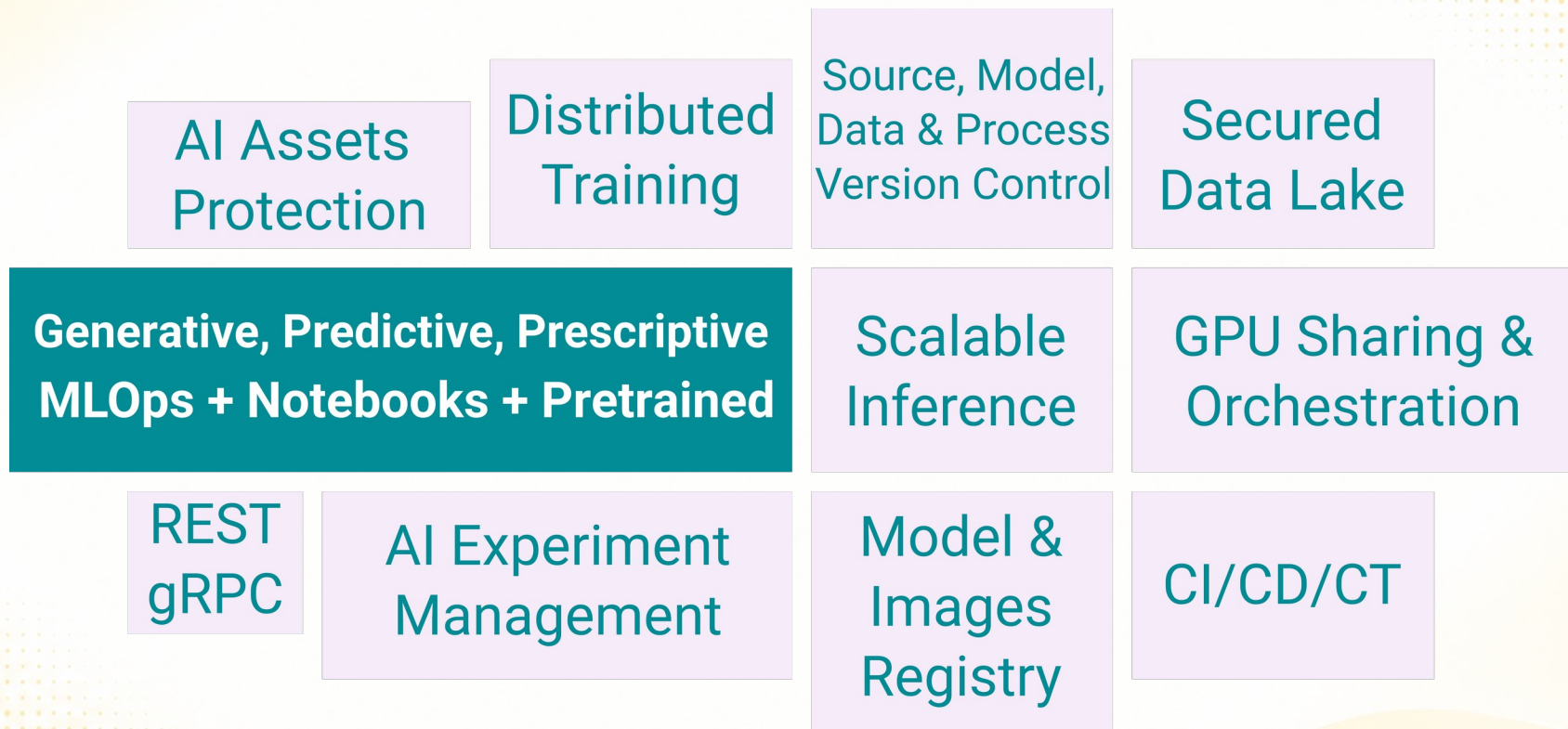
Dedicated

پردازنده گرافیکی قطعه‌بندی شده و تخصیص داده‌شده به صورت اختصاصی (صرفاً برای کارت‌های حرفه‌ای)
فناوری اشتراک‌گذاری واقعی پردازنده گرافیکی به صورت کاملاً پویا قابل بهره‌برداری بر روی تمامی کارت‌ها
شبکه پردازش هوش مصنوعی نزدیک به مصرف‌کننده مورد استفاده در مناطق محروم و با شبکه ضعیف

- **Sliced GPU**
- **Shared GPU**
- **GPUDN**

انتخاب خدمت پردازشی باید بر اساس نوع بارکاری، حساسیت داده، سطح جداسازی و الگوی مصرف انجام شود؛ نه صرفاً قیمت GPU.

سازمان چه قابلیت‌هایی نیاز دارد؟



MLOps نقطه شروع است سازمان به بیش از MLOps نیاز دارد



پایه‌های سرویس پایدار

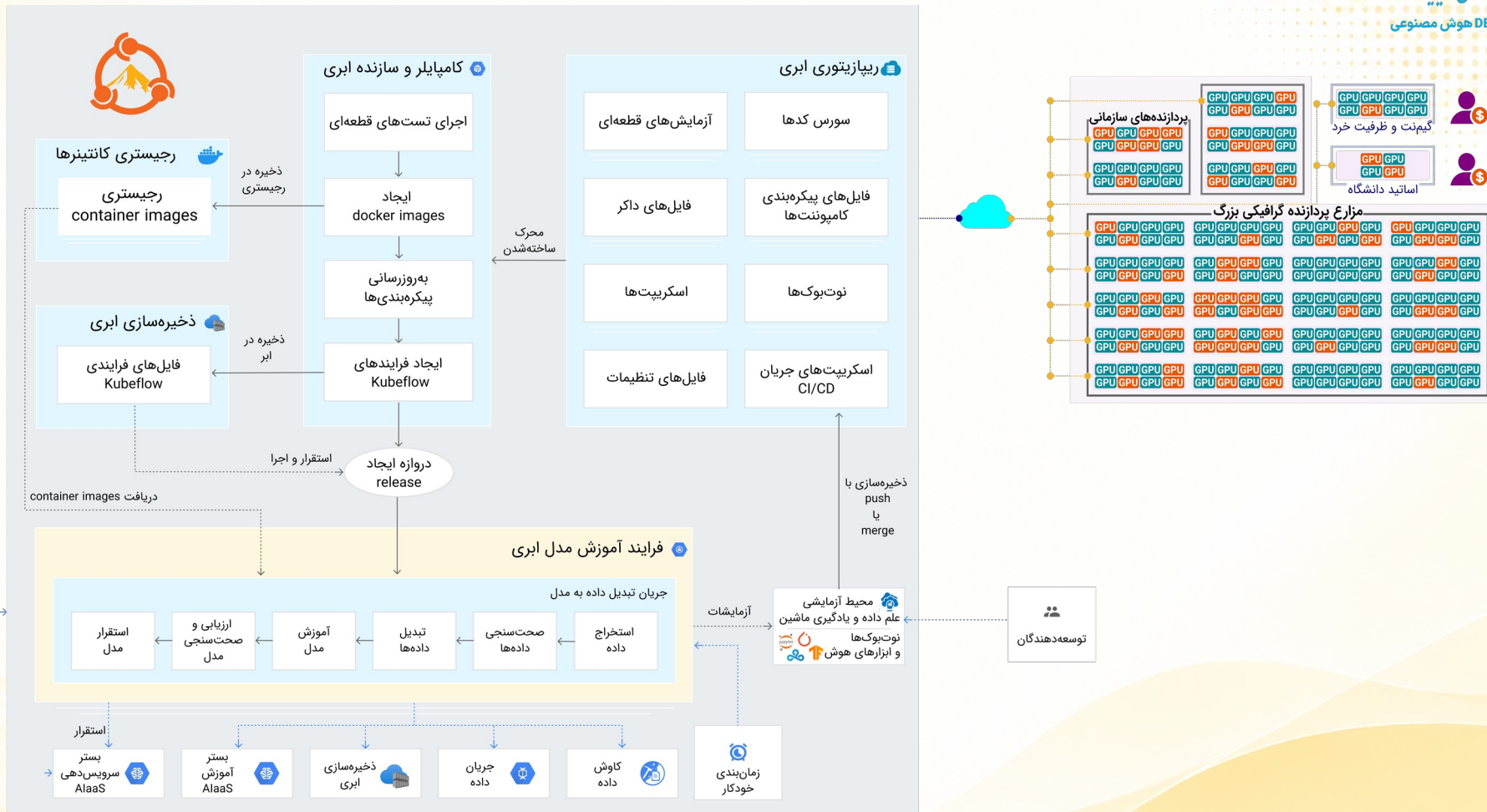
دسترس‌پذیری بالا، مقیاس‌پذیری و قابلیت اعتماد، پایه‌های هر سرویس هوش مصنوعی در دنیای واقعی هستند.

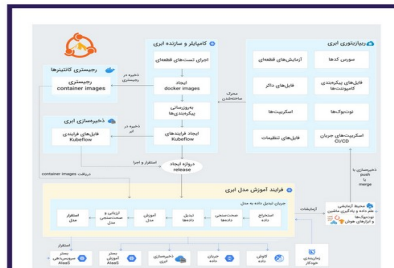
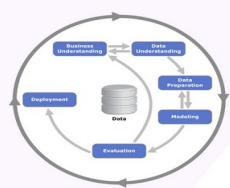
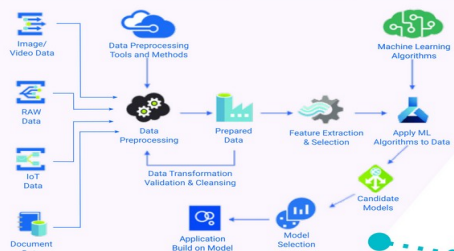


بدون دسترس‌پذیری بالا، سرویس قابل استفاده نیست؛ بدون مقیاس‌پذیری، رشد متوقف می‌شود و بدون قابلیت اعتماد کاربران از دست می‌روند.

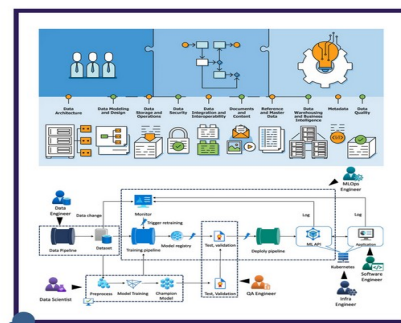
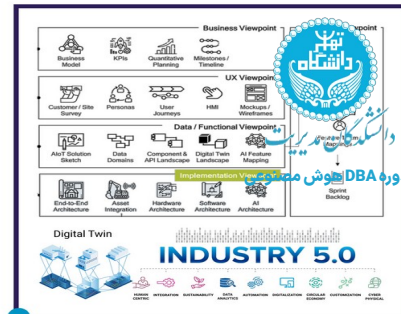


سکوی هوش مصنوعی

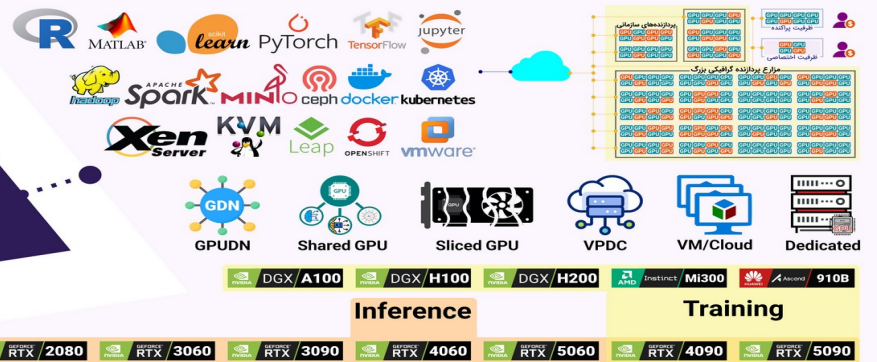




The diagram illustrates the integration of DevOps and MLOps. It features two circular flows. The top flow represents the DevOps cycle: Plan (blue), Code (blue with <code></code> icon), Build (blue), Test (blue), Deploy (yellow with gear icon), and Monitor (yellow). The bottom flow represents the MLOps cycle: Plan (green), Build (green), Deploy (yellow with gear icon), Monitor (yellow), Evaluate (green with ML icon), and Train (green). A central '+' sign indicates the integration of these two cycles. Arrows show the flow of data and code: Data (cloud icon) flows into ML (ML icon), which then flows into Ops (gear icon). Ops flows back into Data, completing a loop. The DevOps cycle also flows into the MLOps cycle, with Plan and Build phases overlapping.



ایپراتور
هوش مصنوعی
و هوشمندسازی
سازمانی/ملی





دانشگاه آزاد اسلامی
واحد تهران مرکزی

دوره DBA هوش مصنوعی

پرونده دستیار: حالا پشت صحنه دیگر یک جعبه سیاه نیست

پشت هر پاسخ دستیار، زنجیره‌ای از داده، مدل، دانش‌افزایی، عملیات و زیرساخت قرار دارد؛ کیفیت و پایداری سامانه به هماهنگی همه این اجزا وابسته است.



حالا دیگر دستیار یک جعبه سیاه نیست؛ می‌توان امنیت، مسئولیت و نظارت را روی اجزای واقعی آن تعریف کرد.

بخش پنجم: امنیت، مسئولیت و نظارت انسانی



امنیت هوش مصنوعی فراتر از امنیت IT است

امنیت هوش مصنوعی را نمی‌توان نسخه بزرگتری از امنیت فناوری اطلاعات دانست؛ چرخه حیات مدل، داده، ابزار و استقرار، ریسک‌های تازه‌ای ایجاد می‌کنند.



امنیت IT



هویت و دسترسی



شبکه و دستگاه



برنامه و API



داده



عملیات زیرساخت

تمرکز سنتی: محافظت از دسترسی، سامانه و داده

اما در AI
پیچیدگی
بیشتر می‌شود



در امنیت AI چه چیزهایی اضافه می‌شود؟



گردش کار پیچیده و چندتیمی
داده، مدل، MLOps و استقرار



حجم و حساسیت داده در کل چرخه
از جمع‌آوری تا آموزش و استنتاج



بلوغ پایین‌تر فرایندها
خارج از DevSecOps کامل



وابستگی‌های پرریسک ML
کتابخانه‌ها، مدل‌ها و زنجیره تأمین



ابزارهای امنیتی ناکافی
دید کم به دارایی‌ها و نقاط ضعف AI



مدل به‌عنوان مصنوع خاص
وزن‌ها، رفتار، تنظیمات و API

امنیت سنتی کافی نیست؛ در هوش مصنوعی باید داده، مدل، فرایند، زنجیره تأمین و رفتار سامانه هم‌زمان امن شوند.



تهدیدات امنیتی خاص هوش مصنوعی

در هوش مصنوعی تهدید فقط به شبکه و داده محدود نیست؛ داده، مدل، بهروزرسانی، ورودی و خروجی هم می توانند هدف حمله قرار گیرند.



حمله همیشه سامانه را از کار نمی‌اندازد

سامانه ممکن است روشن و در دسترس باشد، اما رفتار، خروجی یا تصمیم آن منحرف شده باشد.

وضعیت ظاهری سالم سامانه

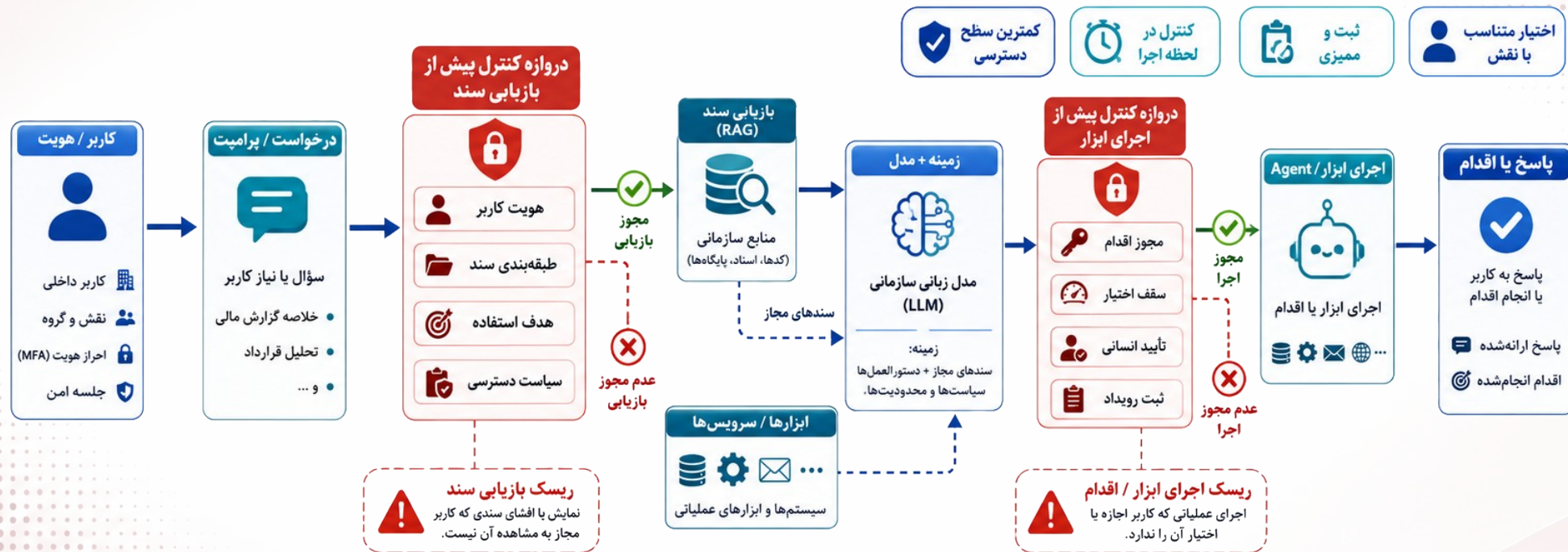


در امنیت هوش مصنوعی، «سامانه روشن است» به معنای «سامانه سالم است» نیست.



دستیارهای هوش مصنوعی ذاتاً امن نیستند مگر دسترسی در لحظه اجرا کنترل شود

مدل نباید سندی را ببیند یا ابزاری را اجرا کند که کاربر مجاز به آن نیست



کنترل پس از تولید پاسخ یا اجرای یک اقدام نوش داروی پس از مرگ سهراب است. باید پیش از بازیابی و پیش از اقدام کنترل کرد



مدل بلوغ حکمرانی هوش مصنوعی

سازمان‌ها از وضعیت موردی و پرسیک آغاز می‌کنند و به سمت کنترل‌های مستند، نظارت‌پذیر و سازگار با زمینه حرکت می‌کنند



سطح ۴: نوموتیک (خودآگاه)

سامانه‌ای هوشمند و تطبیق‌پذیر با مجوزدهی آگاه از زمینه و نظارت پویا.

سطح ریسک:

بهینه



سطح ۳: مدیریت شده

حکمرانی بر پایه سنج‌ها، پایش مستمر و همکاری مبتنی بر شواهد پیش می‌رود.

سطح ریسک:

پایین



سطح ۲: تعریف شده

سیاست‌ها مستند شده‌اند، مرزهای اختیار روشن است و مالکیت‌ها مشخص شده‌اند.

سطح ریسک:

متوسط

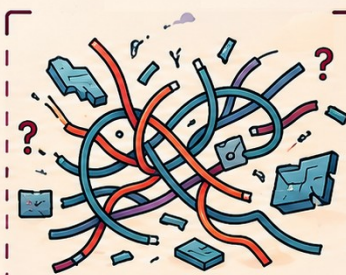


سطح ۱: واکنشی

پس از رخدادها، قواعد پایه شکل می‌گیرند؛ مجوزها ایستا و بازبینی‌ها دستی‌اند.

سطح ریسک:

بالا



سطح صفر: موردی

فاقد چارچوب رسمی، با مرزهای نامشخص و کشف واکنشی مسئله.

سطح ریسک:

بحرانی

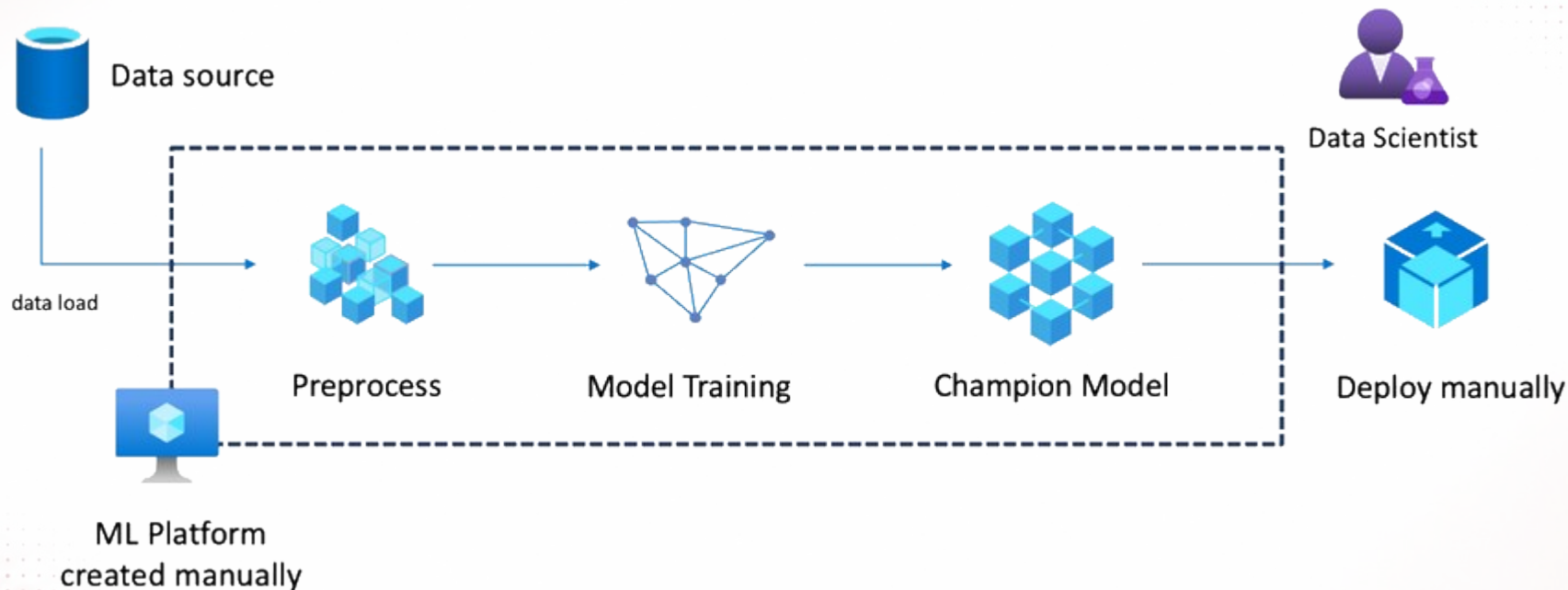


ریسک فقط با افزودن ابزار کم نمی‌شود؛ با بلوغ در سیاست، مجوزدهی، پایش و شواهد کنترل می‌شود.



بلوغ سطح صفر: دستی، موردی و فردمحور

وضعیت اغلب تیم‌های استارت‌آپی و شرکت‌های نوپا

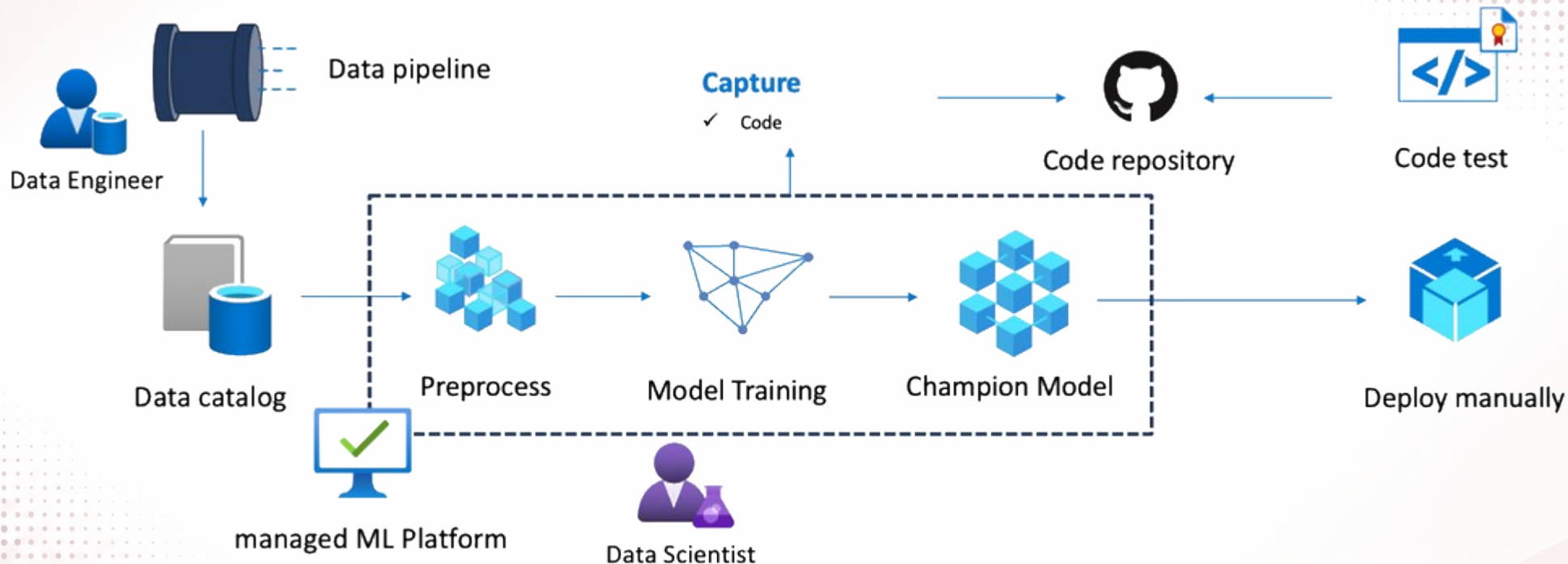


این مرحله به شدت به میزان تبحر دانشمندان و مهندسان داده وابسته است.
هر بار تغییرات نیازمند اعمال و استقرار دستی است و فرآیندها، در صورت وجود، متکی به اسکریپت‌های دستی هستند.



بلوغ سطح یک: کد و محیط مدیریت شده

مخزن کد، آزمون‌های پایه و سکوی مدیریت شده وارد فرایند می‌شوند

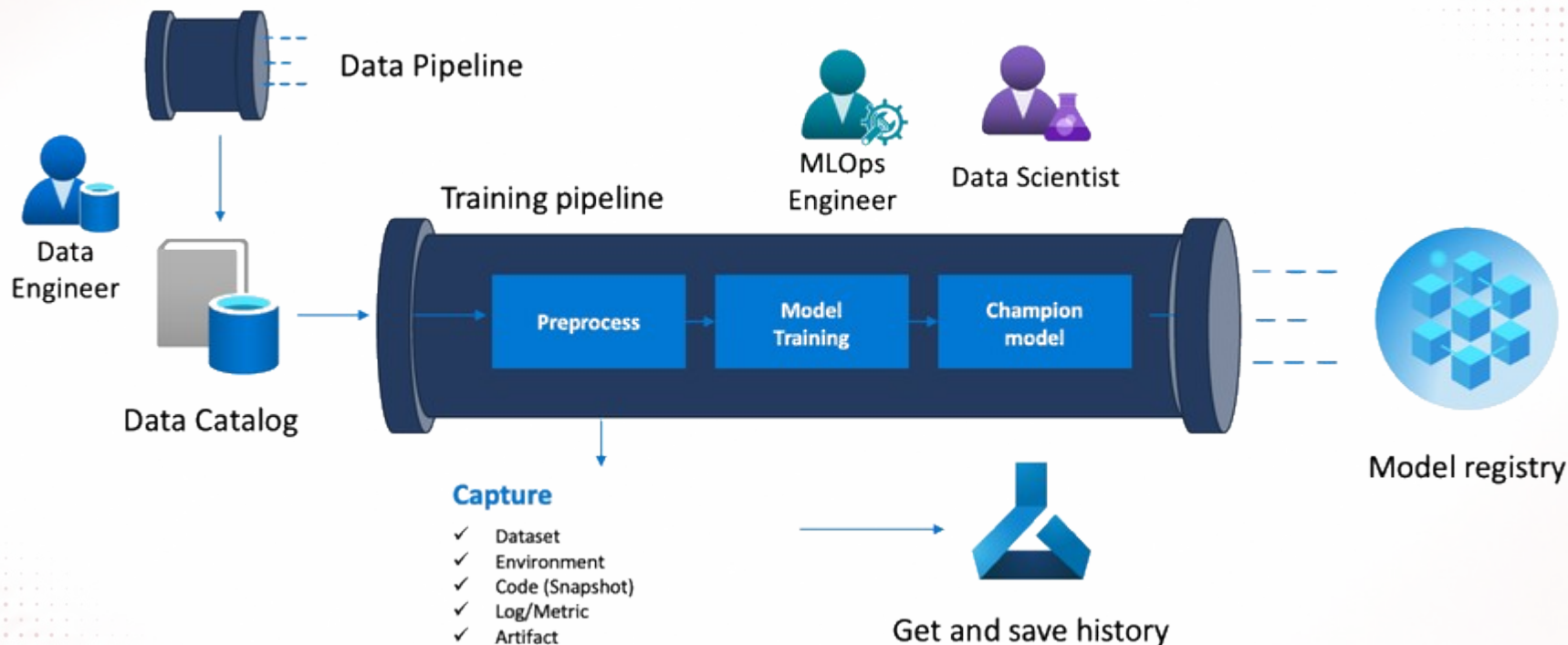


این مرحله به شدت به مهندسان و دانشمندان داده وابسته است. ایجاد نسخه‌های قابل انتشار (release) می‌تواند خودکار شود و معمولاً در فرآیند توسعه، آزمایش‌های قطعه‌ای نیز صورت می‌گیرد.



بلوغ سطح دو: خط لوله قابل تکرار و قابل ردیابی

ورودی حداقلی به دنیای MLOps

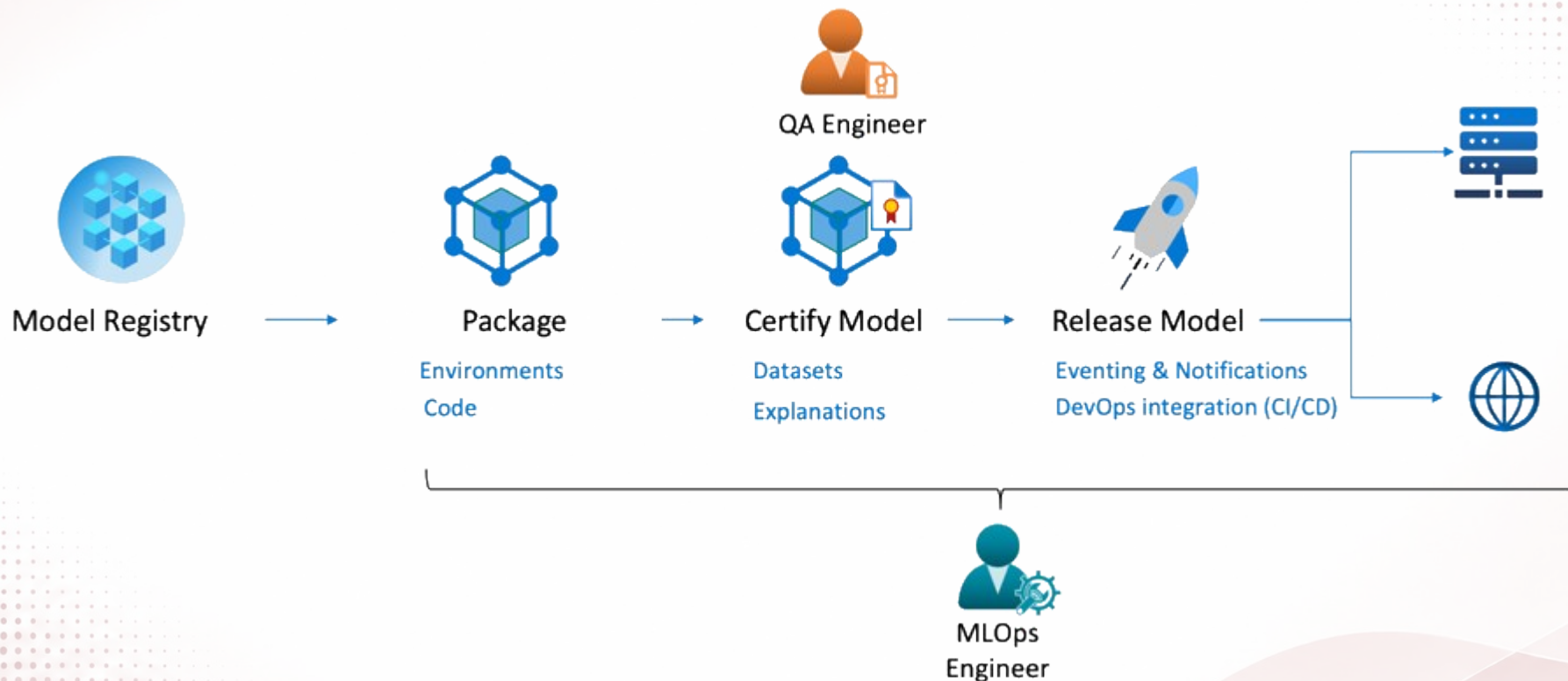


در این مرحله کیفیت محصول به شدت به مهندسان و دانشمندان داده وابسته است و در فرآیند توسعه آزمایش‌های قطعه‌ای صورت می‌گیرد. مکان دنبال کردن کیفیت برای بخش‌های مختلف فراهم آمده و سیستم قابلیت بررسی شفاف پیدا کرده، اما اثر آن در نسخه مستقر مشهود نیست.



بلوغ سطح سه: استقرار خودکار و CI/CD

بسته‌بندی، اعتبارسنجی و انتشار مدل به خط لوله CI/CD متصل می‌شوند

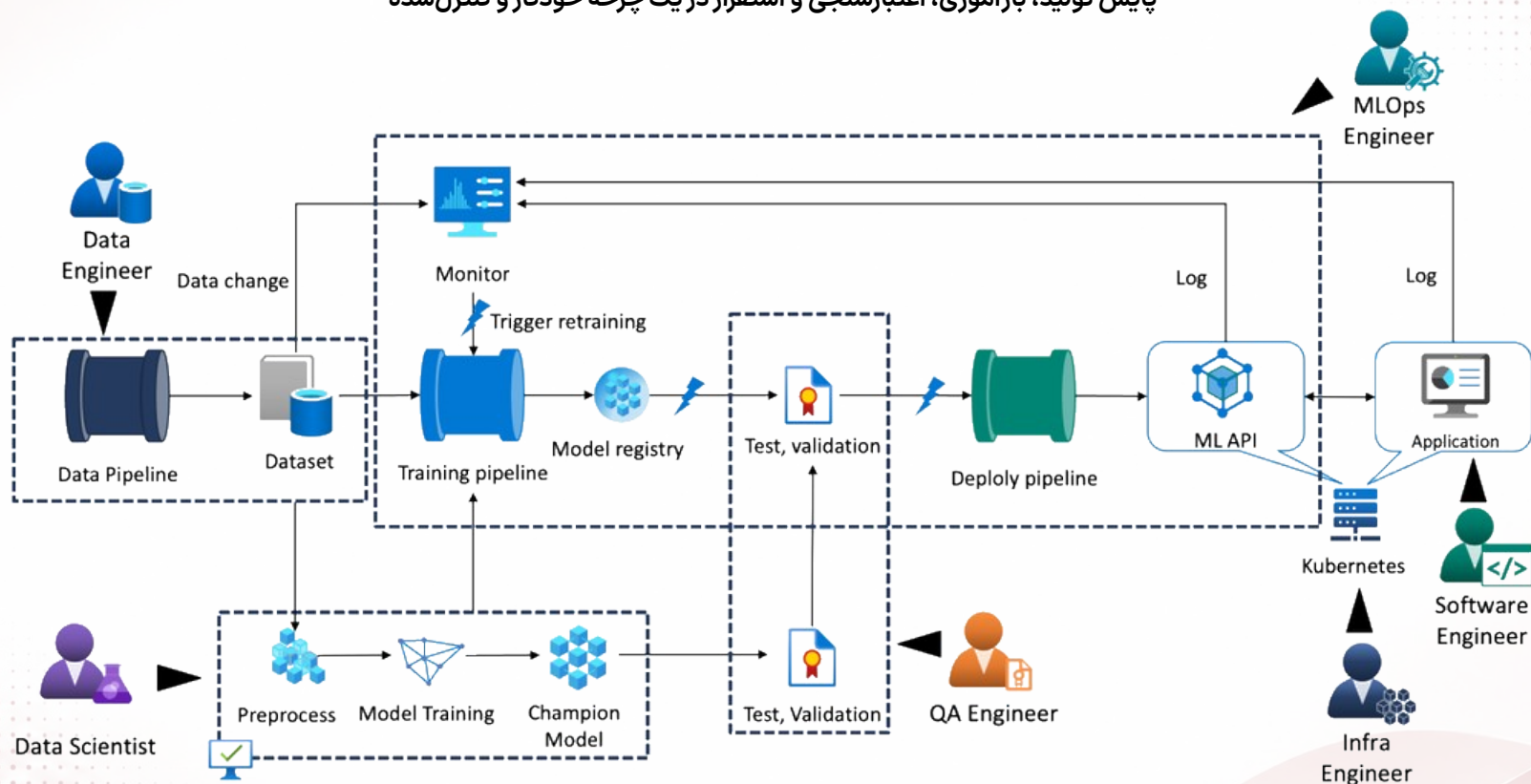


کیفیت محصول به دلیل خودکار شدن فرآیند آزمایش، چندان به دانشمندان داده وابسته نیست. سیستم قابلیت بررسی شفاف پیدا کرده و کیفیت محصول نهایی قابل کنترل شده است اما فرآیند یک‌طرفه بوده و بازخورد مدل مستقر تأثیری در فرآیندها ندارد.



بلوغ سطح چهار: چرخه حلقه بسته با نظارت انسانی

پایش تولید، بازآموزی، اعتبارسنجی و استقرار در یک چرخه خودکار و کنترل شده

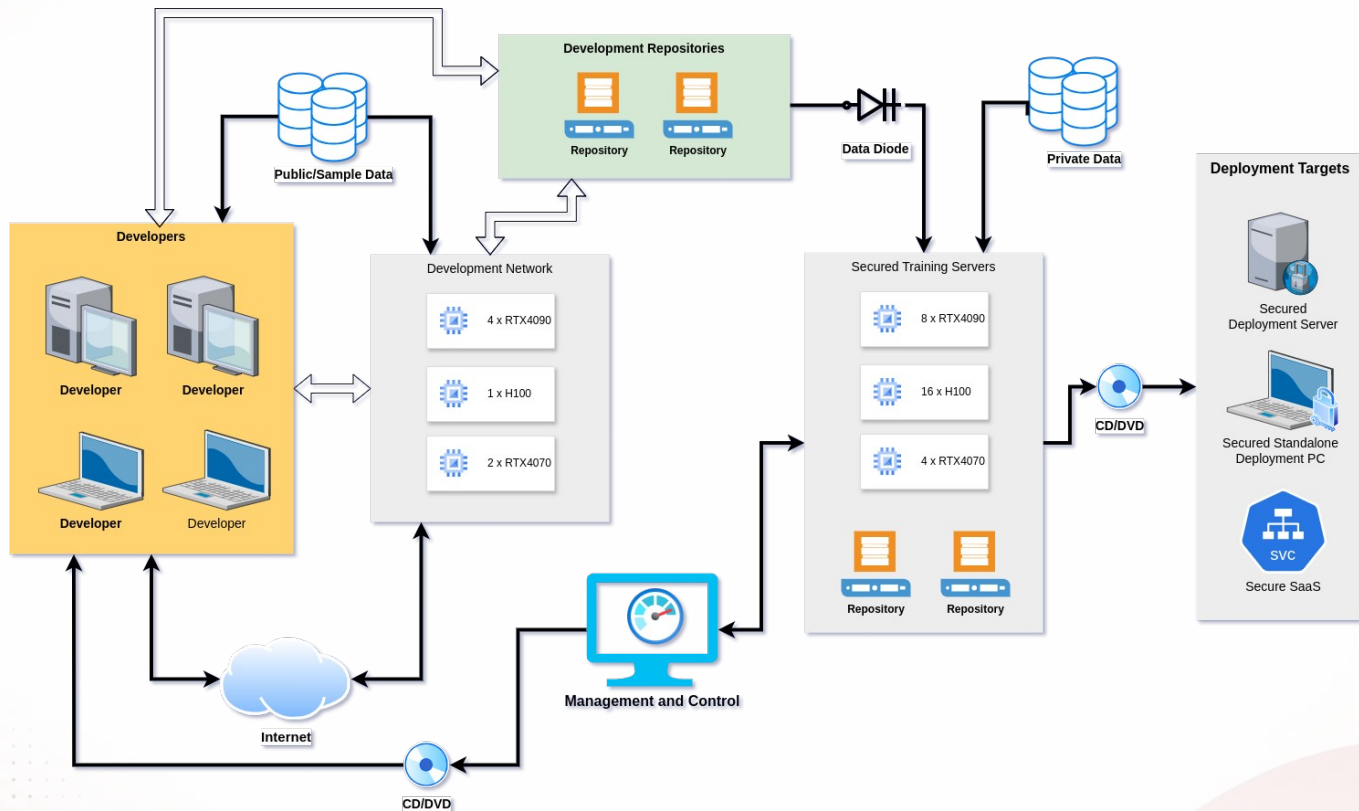


نقش عامل انسانی به ناظر و تعریف‌کننده فرایند تغییر می‌کند. انسان حکمرانی و ماشین اجرا می‌کند. فرآیندهای گردآوری داده‌ها خودکار است؛ فرآیند انتشار خودکار است؛ کیفیت محصول به دلیل خودکار شدن فرآیند آزمایش، چندان به دانشمندان داده وابسته نیست.



پیاده سازی سطح چهار در محیط حساس: استقرار جدا شده و کنترل شده

تفکیک محیطها، کنترل ورودی و خروجی، زنجیره تأمین امن و مقصدهای استقرار مجاز

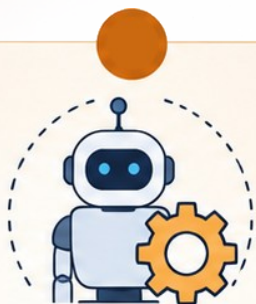


در کاربردهای حساس، اتوماسیون کافی نیست؛ امنیت باید از محیط توسعه تا آموزش و استقرار در خود معماری تعبیه شود.



نظارت انسانی یک کلید روشن یا خاموش نیست

سطح حضور انسان باید با اثر، اختیار و برگست‌پذیری تصمیم متناسب باشد



Human out of the Loop

اقدام مستقل در دامنه محدود
و ازپیش‌تصویب شده

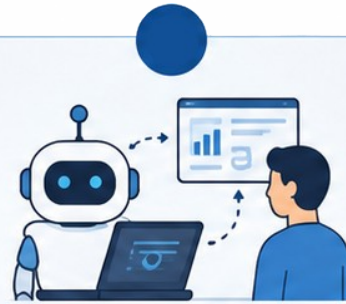
خودکارسازی در چارچوب مجاز



Human over the Loop

انسان سیاست، حدود
و حق توقف را تعیین می‌کند

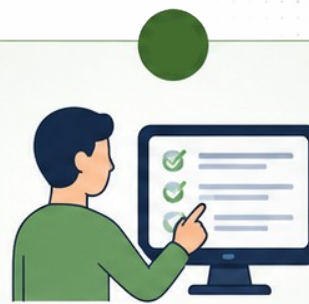
حاکمیت و حدود از پیش تعیین‌شده



Human on the Loop

AI اقدام می‌کند و
انسان پایش می‌کند

نظارت و پایش مستمر انسان



Human in the Loop

انسان تصمیم می‌گیرد

کنترل کامل توسط انسان

پنج شرط نظارت معنادار



امکان توقف



اختیار



دانش



اطلاعات



زمان

انسانی که فرصت، اطلاعات یا اختیار مداخله ندارد، نظارت واقعی ایجاد نمی‌کند.



خطا کجار رخ داد و چه کسی پاسخ گوست؟

خطای مدل، خطای سامانه، خطای استفاده و خطای تصمیم، منشأ و مالک یکسانی ندارند.

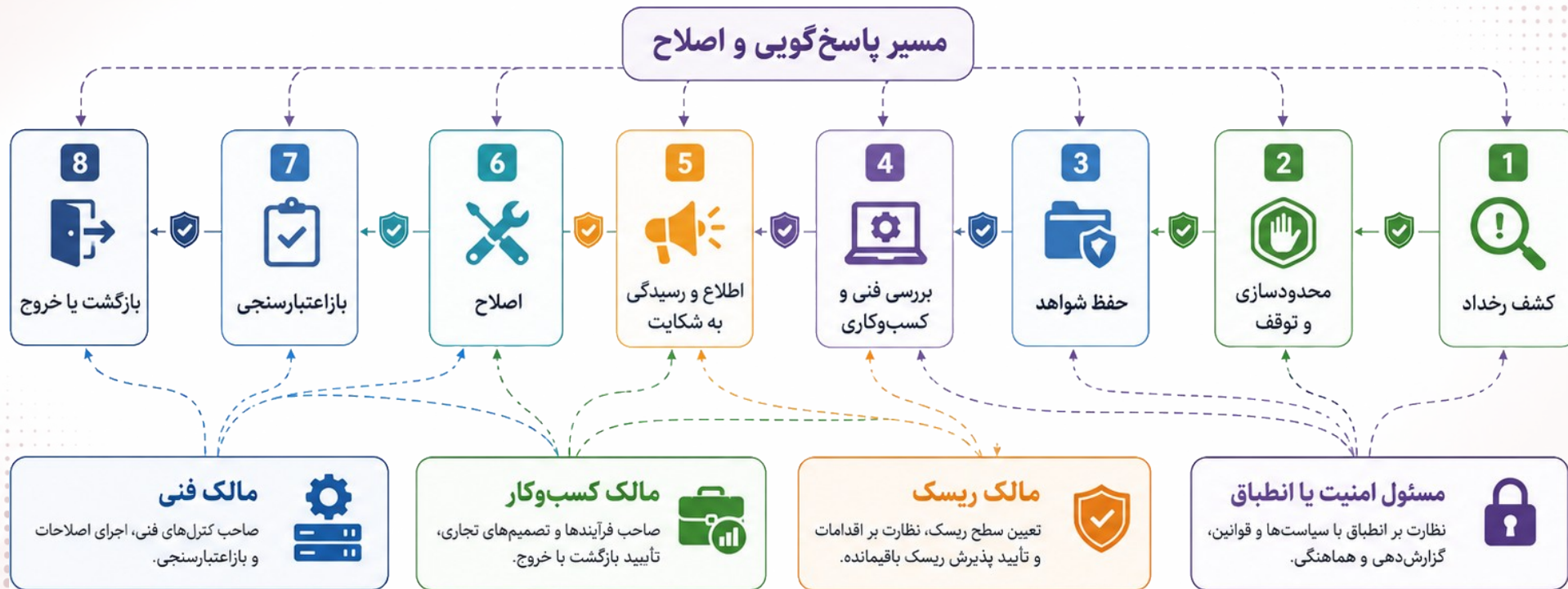


مسئولیت را نمی توان به طور کامل به سازنده مدل، واحد IT، واحد کسب و کار یا کاربر نهایی واگذار کرد.



امنیت وقتی حکمرانی می شود که مسیر پاسخ گویی روشن باشد

صاحب کنترل، حق توقف، شواهد، رخداد، شکایت، بازبینی و اصلاح باید از پیش تعیین شوند



امنیت یک وضعیت فنی نیست؛ نظامی برای تصمیم گیری، پاسخ گویی و اصلاح پس از تغییر یا رخداد است.



Shadow AI: سیاست، سواد و کنترل فنی باید همزمان باشند

کارکنان معمولاً پیش از تدوین سیاست سازمان، استفاده از ابزارهای عمومی را شروع می‌کنند.

ماتریس استفاده مجاز از ابزارها و داده‌ها

	ابزار عمومی	ابزار سازمانی	سامانه داخلی
داده عمومی	مجاز	مجاز	مجاز
داده داخلی	ممنوع	مجاز	مجاز
داده محرمانه	ممنوع	ممنوع	مجاز
داده بسیار حساس	ممنوع	ممنوع	ممنوع

- مجاز
- مشروط
- ممنوع

سیاست روشن

دامنه استفاده، انواع داده، ابزارهای مجاز و حدود مسئولیت

آموزش و سواد AI متناسب با نقش

آموزش عملی برای کارکنان، مدیران، تیم فنی و مالکان ریسک

کنترل فنی و ثبت استفاده

DLP، ثبت رخداد، کنترل دسترسی، پایش و گزارش‌گیری

سیاست بدون آموزش اجرا نمی‌شود و آموزش بدون کنترل فنی قابل اتکا نیست



پرونده دستیار: کاربر به سند غیرمجاز دست یافت

این رخداد فقط شکست فنی نیست؛ آزمون معماری مسئولیت و حکمرانی امنیت است.

مدیر امنیت / امنیت

فوراً دسترسی را قطع
و دامنه را مهار کنیم.

منابع انسانی

حریم خصوصی فرد و
حقوق او در اولویت است.



کاربر: محمد رضایی (کارمند)

دسترسی غیرمجاز !

چتها

اسناد

پروژه‌ها

تنظیمات

پرونده منابع انسانی

محرمانه

ارزیابی عملکرد سال ۱۴۰۲

اطلاعات پرسنلی

وضعیت حقوق و مزایا

مالک فنی

عیب در کنترل بازیابی
و دسترسی است؛ باید
سریع اصلاح شود.

مالک کسب‌وکار

تداوم خدمت حیاتی است؛
ریسک شهرت و توقف را
مدیریت کنیم.

تصمیمات و اقدامات پیشنهادی



توقف یا
محدودسازی خدمت



حفظ شواهد و بررسی
دامنه رخداد



اصلاح کنترل بازیابی



پاسخ به فرد متأثر
و شکایت



بازاعتبارسنجی پیش
از بازگشت



کنترل، مسئولیت و حق توقف وقتی واقعی هستند که در قرارداد، معیار پذیرش و تعهدات بهره‌برداری قابل الزام باشند.



بخش ششم قرارداد، پذیرش و بهره‌برداری

موضوع قرارداد دقیقاً چیست؟

مدل، سامانه، خدمت و نتیجه، چهار موضوع متفاوت در قرارداد هستند



ابهام در موضوع قرارداد، اختلاف در تحویل، مالکیت و مسئولیت را قطعی می کند.

ادعای فنی را به تعهد قابل آزمون تبدیل کنید

«دقت بالا»، «مدل اختصاصی»، «امنیت کامل» و «بهترین فناوری» هیچ کدام معیار پذیرش نیستند.



هر تعهد باید چهار جزء داشته باشد



هر وعده باید به یک مقدار قابل اندازه‌گیری و یکم روش آزمون مینقل تبدیل شود.

چه دارایی‌هایی باید در مالکیت یا کنترل سازمان باقی بمانند؟

قفل‌شدگی فقط به کد مربوط نیست، امکان بازسازی، ارزیابی، انتقال و خروج باید حفظ شوند.



آنچه برای بازسازی، ممیزب، انتقال یا ادامه خدمت لازم است نباید فقط در اختیار پیمانکار باشد.

زیرساخت آموزش باید کجا باشد؟

پاسخ یگانه ندارد؛ محل آموزش به حساسیت داده، تکرار آموزش، هزینه انتقال دانش و مسئولیت بهره‌برداری بستگی دارد.

سمت پیمانکار



- پروژه آزمایشی یا کوتاه‌مدت است
- داده غیرحساس یا ناشناس‌سازی شده است
- سرعت شروع از مالکیت فوری مهم‌تر است
- زیرساخت داخلی هنوز شکل نگرفته است

سرعت شروع بیشتر



مدل ترکیبی



- تخصص پیمانکار لازم است
- داده حساس است اما بخشی قابل جداسازی است
- آموزش پایه بیرون، نهایی‌سازی داخل
- انتقال تدریجی دانش و زیرساخت

توازن میان سرعت و کنترل



سمت کارفرما



- داده بسیار حساس یا محرمانه است
- آموزش مکرب و بلندمدت لازم است
- کنترل، ممیزی و انتقال‌پذیری مهم است
- مدل یک دارایی راهبردی سازمان است

کنترل و حاکمیت بیشتر



معیار تصمیم

مسئولیت بهره‌برداری



نیاز به انتقال



تکرار آموزش



هزینه GPU و مقیاس



حساسیت داده



هرچه داده حساس‌تر، آموزش مکررتر و وابستگی بلندمدت بیشتر باشد، زیرساخت آموزش باید بیشتر در کنترل کارفرما باشد



پاسخ مطمئن، لزوماً پاسخ درست نیست!

اطمینان مدل، لحن قاطع و صحت پاسخ سه مفهوم متفاوت هستند



مدل با اطمینان بالا، بخشنامه منسوخ را به عنوان پاسخ قطعی ارائه کرده است.

مثال پرونده

لحن: قاطع

Confidence: بالا

صحت: نامشخص/نیازمند بررسی



اعتماد به لحن پاسخ، ضریب اطمینان بالا، جای آزمون صحت و اعتبار منبع را نمی گیرد.

نوع خطا مهم‌تر از یک عدد دقت است

False Negative و False Positive در کاربردهای مختلف هزینه و ریسک یکسانی ندارند

False Positive

اعلام مزایای غیرمجاز به کارمند



- 1 سامانه به اشتباه اعلام می‌کند: کارمند واجد مزایا است.
- 2 هزینه: پرداخت یا تعهد غیرمجاز، نارضایتی پس از اصلاح، ریسک حقوقی یا بودجه‌ای.
- 3 در این سناریو، کنترل خطای اعطای نابه‌جا مهم‌تر است.

ریسک غالب: اعطای نابه‌جا

False Negative

نادیده‌گرفتن مزایای واقعی کارمند



- 1 سامانه به اشتباه اعلام می‌کند: کارمند واجد مزایا نیست.
- 2 هزینه: تضییع حق کارمند، شکایت، کاهش اعتماد و بار رسیدگی.
- 3 در این سناریو، کاهش خطای محروم‌سازی نابه‌جا مهم‌تر است.

ریسک غالب: نادیده‌گرفتن حق



مثال امنیتی



امنیتی: False Positive:
دسترسی اشتباه به سند محرمانه

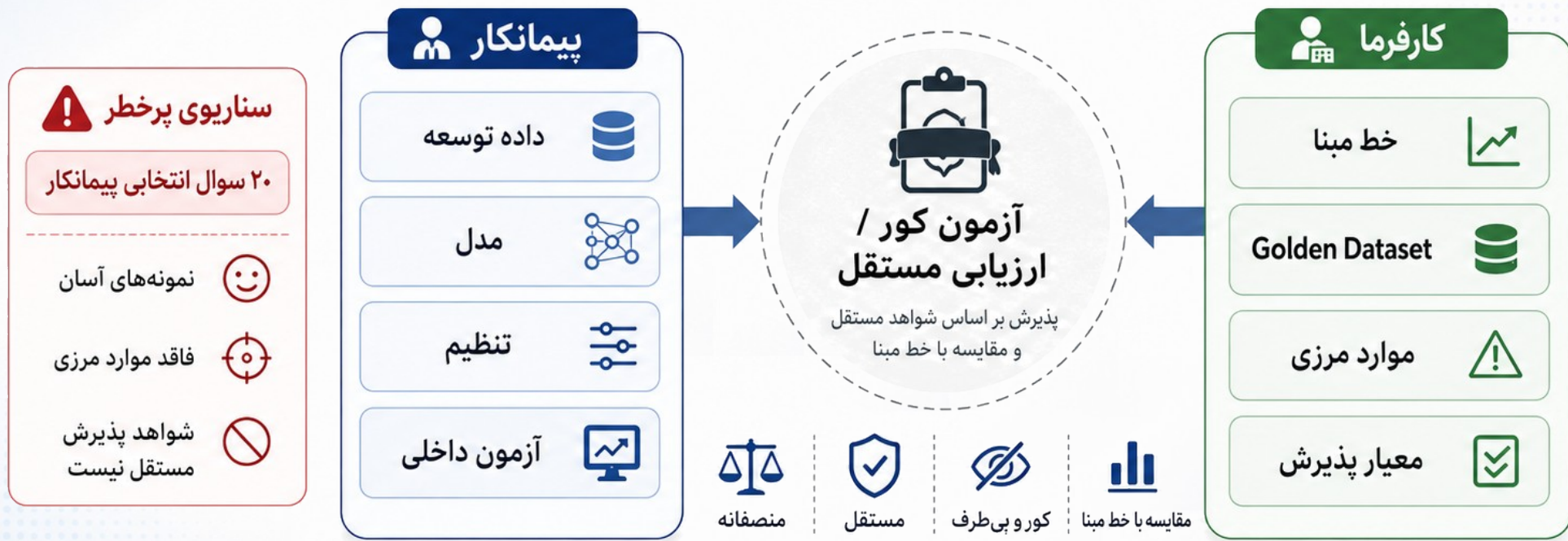


امنیتی: False Negative:
محروم شدن کاربر مجاز از دسترسی لازم

معیار موفقیت باید بر اساس هزینه خطا انتخاب شود، نه عددی که بزرگ‌تر یا حرفه‌ای‌تر به نظر می‌رسد.

خط مبنا، پیکره طلایی و ارزیابی مستقل

پیمانکار نباید هم سازنده سامانه، هم طراح آزمون و هم داور نتیجه باشد.



داده آزمونی که مدل پیشتر دیده یا پیمانکار انتخاب کرده باشد، شواهد مستقل پذیرش نیست.



دموی موفق، دلیل پذیرش پروژه نیست

دمو، پایلوت، MVP، و سامانه عملیاتی، ادعاهای متفاوتی را اثبات می‌کنند.



دموی زیبا امکان اولیه را نشان می‌دهد؛ دمو ضامن قابلیت بهره‌برداری، امنیت، مقیاس یا پایداری نیست.

پذیرش چندلایه و استقرار مرحله‌ای

پذیرش فقط دقت مدل نیست و استقرار نیز یک کلید روشن یا خاموش نیست

لایه‌های پذیرش



مسیر استقرار



سامانه باید هم درست پاسخ دهد، هم در شرایط واقعی قابل بهره‌برداری، امن، پذیرفتنی و اقتصادی باشد.

پرونده قابل حسابرسی: چه شواهدی باید باقی بماند؟

مسئولیت بدون تبار، مستندات، لاگ، و صبت تصمیمات قابل پذیرش نیست.



حکمرانی زمانی قابل دفاع است که سازمان بتواند نشان دهد چه تصمیمی، بر مبنای چه شواهدی و توسط چه کسی گرفته شده است.

حکمرانی با تحویل پروژه پایان نمی یابد

پایش، Drift، رخداد، بازآزمایی، توقف و خروج بخشی از قرارداد و عملیات اند.



قرارداد هوش مصنوعی باید تغییر سامانه و پایان عمر آن را اداره کند، نه فقط نسخه روز تحویل را



دوره DBA هوش مصنوعی

مصوبه نهایی دستیار سازمانی

اکنون درباره یک دمو تصمیم نمی‌گیریم؛ درباره یک قابلیت سازمانی تصمیم می‌گیریم.

تصمیم هیئت‌مدیره



تصویب کامل



استقرار محدود و مشروط ☐



تعليق برای رفع نواقص ☐



توقف ☐



بهره‌برداري
و خروج



قرارداد و
پذيرش



امنيت و
نظارت



اثر و ريسک



معماری و داده



مالک و
حق تصميم



مسئله و دامنه

پرونده کامل
دستیار سازمانی



حکمرانی هوش مصنوعی یعنی توان سازمان برای تصمیم‌گیری منظم، مستند و قابل بازنگری در سراسر چرخه عمر.

منابع و مراجع

- 1) NIST (2023), Artificial Intelligence Risk Management Framework (AI RMF 1.0), NIST AI 100-1.

مرجع اصلی دوره برای منطق Govern / Map / Measure / Manage، ریسک، مسئولیت و چرخه تصمیم.

- 2) NIST (2024), Artificial Intelligence Risk Management Framework: Generative AI Profile, NIST AI 600-1

مکمل ضروری برای LLM، هوش مولد، RAG، اعتمادپذیری، خطا و ریسک‌های جدید.

- 3) ISO/IEC 42001:2023, Artificial Intelligence Management System.

برای ساختار سازمانی حکمرانی، سیاست‌ها، نقش‌ها، مسئولیت‌ها، کنترل و بهبود مستمر.

- 4) ISO/IEC 23894:2023, Artificial Intelligence – Guidance on Risk Management.

مرجع مستقیم بخش ارزیابی اثر، سطح ریسک، کنترل و پذیرش ریسک.

- 5) ISO/IEC 38507:2022, Governance implications of the use of artificial intelligence by organizations.

به‌ویژه مناسب مباحث هیئت‌مدیره، حق تصمیم، پاسخ‌گویی و نقش مدیران.

- 6) European Union (2024), Regulation (EU) 2024/1689 – Artificial Intelligence Act (EU AI Act).

برای رویکرد مبتنی بر ریسک، کاربردهای پریسک، مستندسازی، نظارت انسانی و مسئولیت.

- 7) OECD, OECD AI Principles, updated 2024.

مرجع معتبر برای اصول اعتمادپذیری، پاسخ‌گویی، شفافیت، ایمنی و حکمرانی هوش مصنوعی

منابع و مراجع

8) ISO/IEC 5259 Series, Data Quality for Analytics and Machine Learning.

برای بخش‌های کیفیت داده، Data Cleansing، اعتبار داده، مدیریت و حکمرانی کیفیت داده.

9) Google Cloud, MLOps: Continuous Delivery and Automation Pipelines in Machine Learning.

منبع بسیار مناسب برای اسلایدهای MLOps، نسخه‌بندی، Retraining و Training Pipeline، Deployment، Monitoring.

10) Sculley, D. et al. (2015), "Hidden Technical Debt in Machine Learning Systems," NeurIPS.

یکی از بهترین منابع برای توضیح اینکه چرا «مدل» تنها جزء کوچکی از سامانه واقعی هوش مصنوعی است.

11) Lewis, P. et al. (2020), "Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks," NeurIPS.

مرجع اصلی RAG و مناسب برای بخش معماری دستیار سازمانی.

12) NIST SP 800-207 (2020), Zero Trust Architecture.

پایه نظری Zero Trust و نقطه شروع برای توضیح تفاوت امنیت سنتی با رویکرد ZTA.

13) NIST (2025), Adversarial Machine Learning: A Taxonomy and Terminology of Attacks and Mitigations, NIST AI 100-2e2025.

مهم‌ترین مرجع فنی برای Data Poisoning، Evasion، Model Theft، Privacy Attacks و سایر تهدیدهای خاص AI.

14) World Economic Forum (2020), AI Procurement in a Box – AI Government Procurement Guidelines.:

برای بخش پایانی دوره: تأمین‌کننده، قرارداد، پذیرش، تعهد قابل آزمون، مالکیت دارایی‌ها و خروج از خدمت.