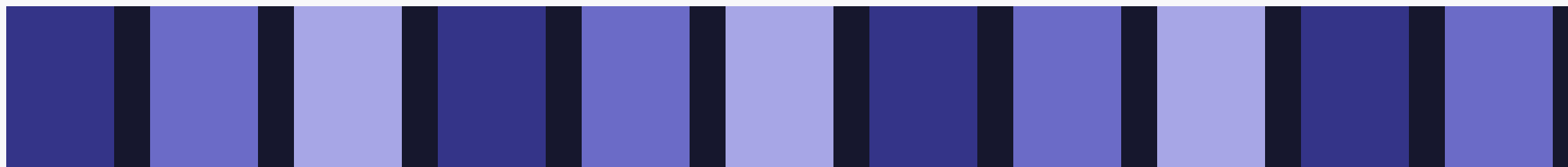


Physical AI v roce 2026

Od humanoidní robotiky k bezpečnému průmyslovému nasazení

Letní škola Czech AI Factory



Dvě otázky pro dnešních 90 minut

ÚVOD

První polovina vysvětlí, jak dnes Physical AI skutečně funguje. Druhá ukáže, jak stejný technologický stack převést do rozhodnutí, pilotu a bezpečného provozu.

45 min — dnešní Physical AI

- Co zůstává z klasické robotiky a co přináší učení z dat.
- Jak vznikají robotická data: VR, egocentrická videa, UMI, GELLO a ALOHA.
- Jak spolu souvisejí VLA modely, simulace, world modely, JEPA a hmat.
- Kde je dnes skutečná hranice generalizace a spolehlivosti.

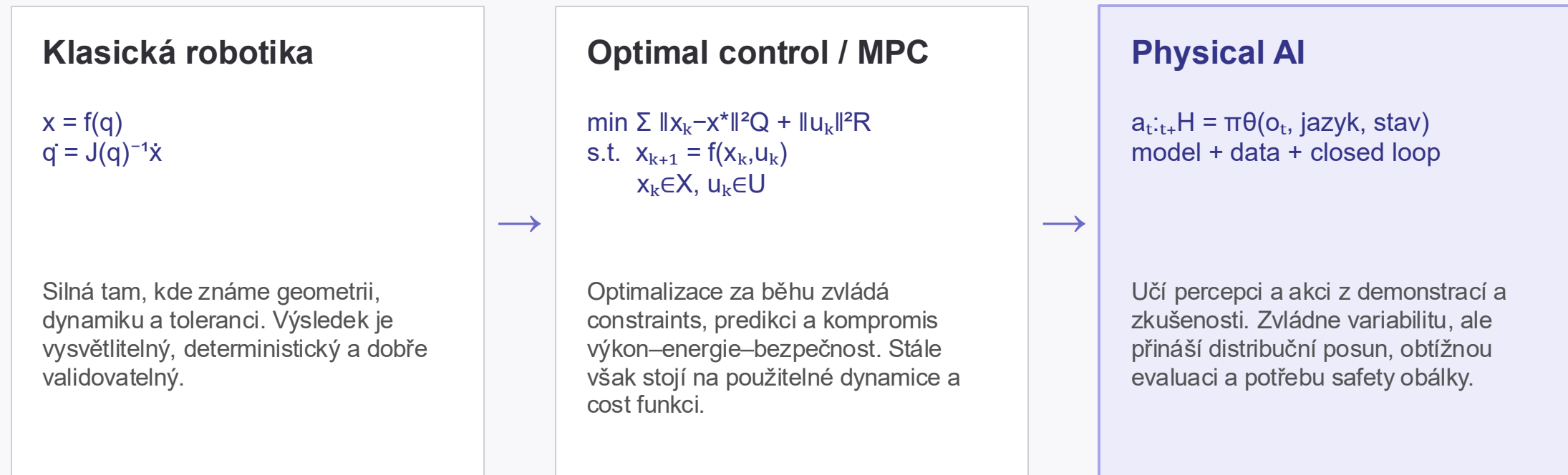
45 min — reálný průmyslový use case

- Jak vybrat úlohu, u které dává Physical AI technický i ekonomický smysl.
- Příklad humanoida pro prohlídky: navigace, interakce, znalostní databáze a LLM.
- Učící pipeline od demonstrace přes evaluaci až po řízený release.
- Safety, CE, odpovědnost, servis a uzavřená datová smyčka.

Od kinematiky přes MPC k Physical AI

I. JAK VYPADÁ DNEŠNÍ PHYSICAL AI

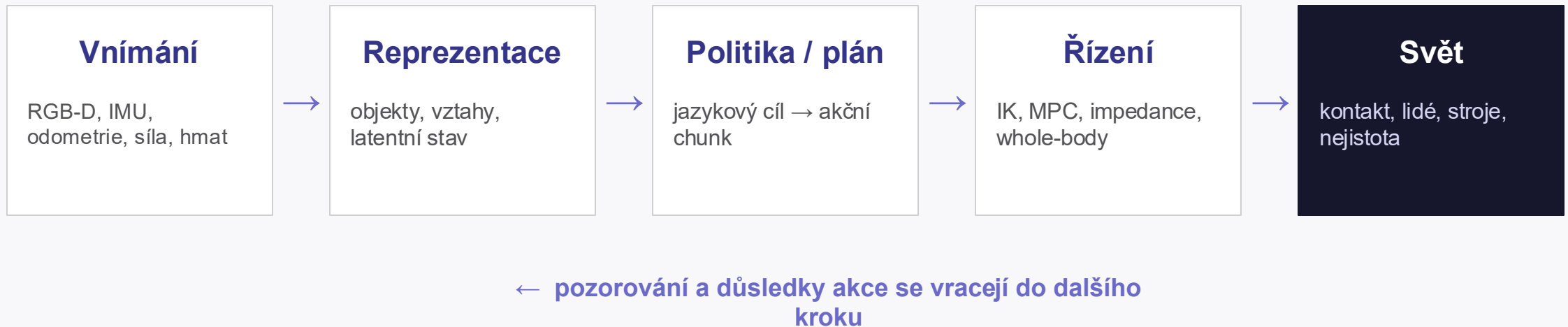
Physical AI nenahrazuje klasickou robotiku. Přidává naučené reprezentace a politiky tam, kde přesný model světa, objektů a kontaktu nelze rozumně napsat ručně.



Physical AI je uzavřená smyčka, ne chatbot s motory

I. JAK VYPADÁ DNEŠNÍ PHYSICAL AI

Model musí vnímat neúplný svět, zvolit akci, provést ji přes řízení a z následků znovu odvodit stav. Chyba se zde neprojeví špatnou větou, ale fyzickým pohybem.



Úskalí: kumulace chyb, latence, nepozorovaný stav, kontakt a změna distribuce. Výhoda: schopnost pracovat s variabilitou, jazykovým zadáním a dlouhým ocasem situací.

Proč se Physical AI zrychluje právě teď

I. JAK VYPADÁ DNEŠNÍ PHYSICAL AI

Nejde o jediný průlom. Současně dozrálo několik vrstev, které se navzájem násobí.

-
- Hardware je dostupnější: bezpečnější manipulátory, levnější humanoidní platformy, RGB-D kamery, haptika a výkonné edge GPU.
 - Robotické datasety překročily řád stovek až tisíců hodin a začínají zachycovat mnoho tasků, prostředí i embodimentů.
 - VLA modely přenášejí vizuálně-jazykové znalosti do akčního prostoru a umožňují efektivnější adaptaci na nový task.
 - Simulace a world modely zlevňují ablace, generování variant, testování regresí a hledání failure modes.
 - Stále však neexistuje zkratka kolem fyzické evaluace, integrace a bezpečnostního inženýrství.

Humanoid je zajímavý hlavně kvůli prostředí navrženému pro lidi

I. JAK VYPADÁ DNEŠNÍ PHYSICAL AI

Dvě ruce, lidský pracovní prostor a mobilita umožňují použít existující dveře, regály, nástroje a pracovní instrukce. To ale neznamená, že humanoid je automaticky nejlepší stroj pro každý task.

- Silná stránka: víceúčelovost, dosah, bimanuální koordinace a interakce s lidmi, učení z lidských dat.
- Slabá stránka: vyšší cena, energetika, stabilita, složité whole-body řízení a velká bezpečnostní obálka.
- Pro pevně definovanou vysokorychlostní operaci bývá klasická buňka stále lepší.
- Humanoid dává smysl tam, kde se úloha mění, prostředí nelze snadno přestavět nebo je důležitá mobilita.



Autonomie je hierarchie více řídicích vrstev

I. JAK VYPADÁ DNEŠNÍ PHYSICAL AI

Jeden foundation model nemá řídit každý motorický detail. Praktická architektura rozděluje odpovědnost podle časové škály, ověřitelnosti a rizika.

Vrstva	Typická frekvence	Co rozhoduje	Co musí garantovat
Mise / task graph	sekundy–minuty	pořadí úloh, výjimky, cíle	stavový automat, audit, recovery
VLA / learned policy	1–20 Hz	akční chunky, grasp, koordinace	omezený akční prostor, confidence
Motion	20–200 Hz	trajektorie a constraints	kolize, limity, plynulost
Servo / impedance	0,2–2 kHz	moment, rychlost, kontakt	stabilita a hard limity
Safety layer	nezávislá	zóna, rychlost, E-stop, interlock	bezpečný stav bez ohledu na AI

Hlavní bottleneck jsou kvalitní data

I. JAK VYPADÁ DNEŠNÍ PHYSICAL AI

Robotická data jsou drahá, protože musí být časově synchronizovaná, akčně ukotvená a fyzicky bezpečná. Jedna hodina videa není totéž jako jedna hodina použitelných demonstrací.

- Potřebujeme pokrýt nejen úspěšné trajektorie, ale i odchylky, zásahy, částečné úspěchy a recovery.
- Data musejí nést kontext: instrukci, variantu prostředí, objekt, kalibraci, stav nástroje a výsledek.
- Rozmanitost zvyšuje generalizaci; přesnost a opakovatelnost jsou nutné pro specializované průmyslové skilly.
- Nejcennější dataset vzniká z reálných edge cases, ne z dokonale připraveného laboratorního stolu.

Hlavní bottleneck jsou kvalitní data

I. JAK VYPADÁ DNEŠNÍ PHYSICAL AI

Epizoda propojuje cíl, pozorování, akci a výsledek v čase. Bez správné synchronizace a metadat se záznam obtížně používá pro trénink i reprodukci chyby.

Co musí epizoda obsahovat

- task ID, přirozenou instrukci a definici úspěchu
- RGB/RGB-D, joint state, end-effector pose a stav gripperu
- akce v robotickém nebo normalizovaném akčním prostoru
- force/tactile signály, pokud rozhodují o kontaktu
- časové značky, kalibraci a metadata procesu

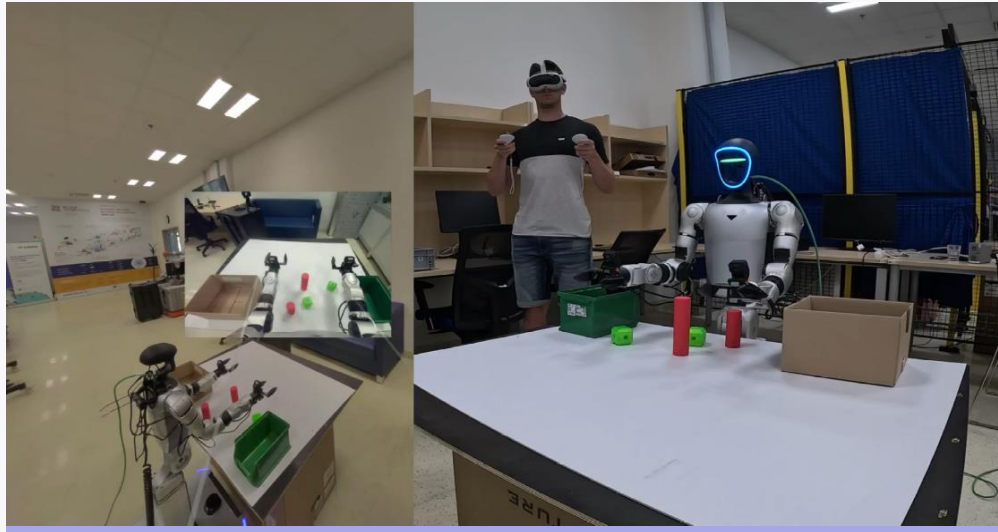
Co přidat pro provoz

- success/failure a důvod zásahu operátora
- pre-roll a post-roll okolo incidentu
- SKU, nástroj, stav linky a PLC/MES event
- segment recovery: odkud se systém dokázal vrátit
- verzi policy, runtime a bezpečnostní konfigurace

VR teleoperace je dobrý začátek, ale špatný konečný stav

I. JAK VYPADÁ DNEŠNÍ PHYSICAL AI

VR/XR headset a hand controllers umožní rychle spustit sběr a využít přirozenou koordinaci člověka. Při stovkách hodin však začnou dominovat ergonomie, kalibrace a nízký throughput.

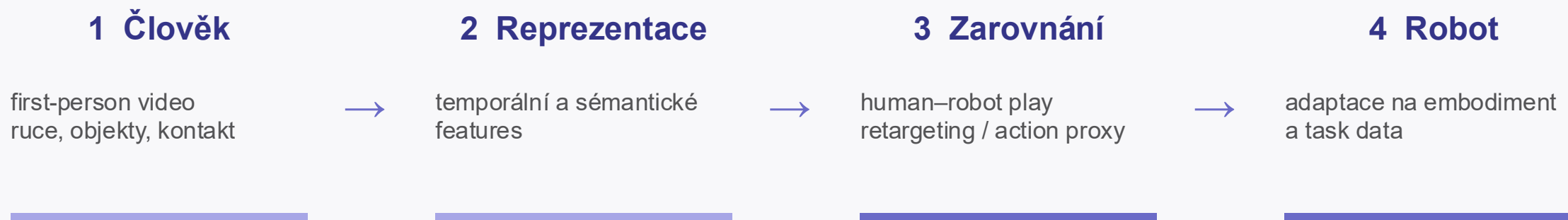


- Výhoda: intuitivní mapování rukou, pohledu a těla; vhodné pro prototyp, demonstraci a remote assist.
- Nevýhoda: operátor je oddělený od přirozené práce, rychle se unaví a musí sledovat latenci i omezené zorné pole.
- Retargeting člověk → robot je netriviální: jiný dosah, počet kloubů, self-collision a kontakt.
- V praxi je VR cenný bootstrap a fallback; pro škálování dat bývají lepší UMI/GELLO nebo pasivní egocentrický sběr.

Egocentrická data škálují zkušenost člověka

I. JAK VYPADÁ DNEŠNÍ PHYSICAL AI

Video z pohledu člověka zachycuje long tail běžných činností bez drahého robota. Nemá však přímo robotické akce ani přesnou propriocepci, proto slouží hlavně pro reprezentaci a mid-training.



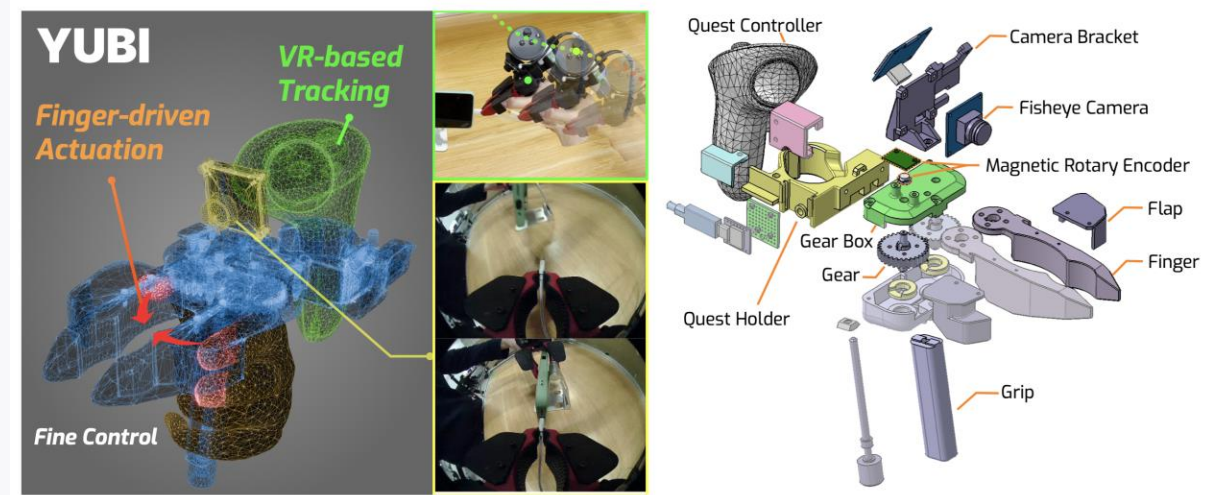
EgoScale (NVIDIA GEAR, 2026) ukazuje praktickou myšlenku: lehký mid-training na zarovnaných human-robot play datech může zlepšit přenos reprezentací. Egocentrická data nenahrazují robotické epizody; rozšiřují jejich pokrytí.

UMI sbírá demonstrace v reálném světě bez robota

I. JAK VYPADÁ DNEŠNÍ PHYSICAL AI

Universal Manipulation Interface používá ruční gripper se snímáním pohybu a obrazu. Operátor pracuje přímo v cílovém prostředí; záznam se následně převádí do akčního prostoru robota.

- Rychlejší a mobilnější sběr než s plným robotickým pracovištěm.
- Přirozený kontakt a reálná variabilita objektů, osvětlení i pozadí.
- Nutnost přesné kalibrace, lokalizace gripperu, synchronizace kamer a retargetingu.
- Nezachytí všechny limity cílového embodimentu: dosah, self-collision, compliance ani dynamiku pohonu.
- Největší hodnota: škálování task-specific dat ještě před instalací finálního robota.



GELLO a ALOHA: levný hardware, vysoká datová hodnota

I. JAK VYPADÁ DNEŠNÍ PHYSICAL AI

Pro přesné robot-aligned demonstrace je výhodné ovládat followera kinematicky podobným leaderem. Operátor cítí limity robotického prostoru a data obsahují přímo realizovatelné akce.



GELLO

Nízká cena, kinematická podobnost a jednoduché logování. Vhodné pro rychlý task-specific sběr a různé robotické paže.



ALOHA + ACT

Bimanuální low-cost platforma a action chunking pro jemné dlouhé trajektorie. Silná reference pro koordinaci, ne hotový průmyslový produkt.

Datasety se liší rozsahem i blízkostí cílovému robotu

I. JAK VYPADÁ DNEŠNÍ PHYSICAL AI

Velký dataset přináší rozmanitost; lokální task data přinášejí přesnost. Úspěšná politika obvykle kombinuje obojí.

Zdroj	Měřítko / charakter	Co učí	Omezení
DROID	76k trajektorií, 350 h, 564 scén	in-the-wild rozmanitost	jedna hlavní platforma, nákladný setup
Open X-Embodiment	>1M trajektorií, 22 embodimentů	cross-embodiment přenos	heterogenní akce a kvalita
ABC-130K	134 806 epizod, 3 553 h, 195 tasků	bimanuální BC a dexterita	YAM setup; stále laboratorní distribuce
ALOHA / lokální data	desítky–stovky epizod na task	přesný skill a kalibraci	úzké pokrytí a riziko overfitu

Metody Behavior cloning

I. JAK VYPADÁ DNEŠNÍ PHYSICAL AI

Z demonstrací se učíme napodobit akci operátora. Architektura může být transformer, diffusion policy nebo flow model; rozhodující je datová distribuce a closed-loop chování.

Co optimalizujeme

- $\pi_{\theta}(a_{t:t+H} | o_t, \text{instrukce})$: predikce krátkého akčního chunku
- supervised loss nad trajektoriemi demonstrátora
- action chunking redukuje horizon a vyhlazuje řízení
- sekvenční model drží kontext více kamer a propriocepce

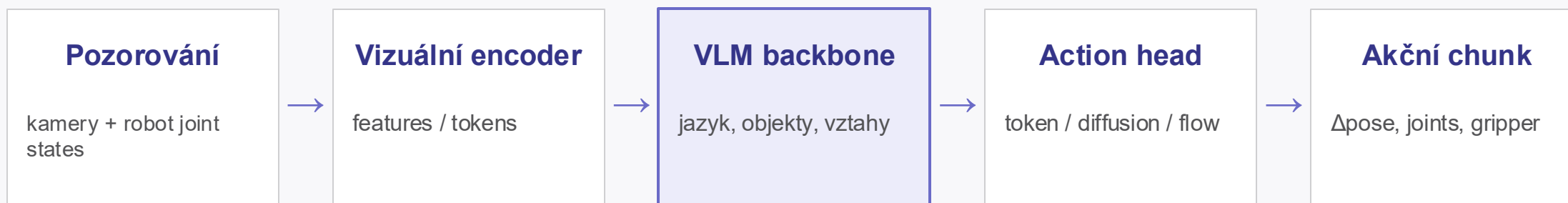
Proč se policy rozpadne

- covariate shift: vlastní malá chyba vede do stavu, který v datech nebyl
- multimodalita: pro stejnou scénu existuje více správných trajektorií
- nevyvážená data: úspěchy převažují nad recovery a výjimkami
- řešení: DAgger-like sběr, interence, failure mining a opakovaná evaluace

VLA propojuje význam instrukce s časovou sekvencí akcí

I. JAK VYPADÁ DNEŠNÍ PHYSICAL AI

Vision–Language–Action model přijímá obraz, jazyk a robotický stav. Výstupem není text, ale akce: často jako tokeny, diffusion/flow trajectory nebo krátký akční chunk.



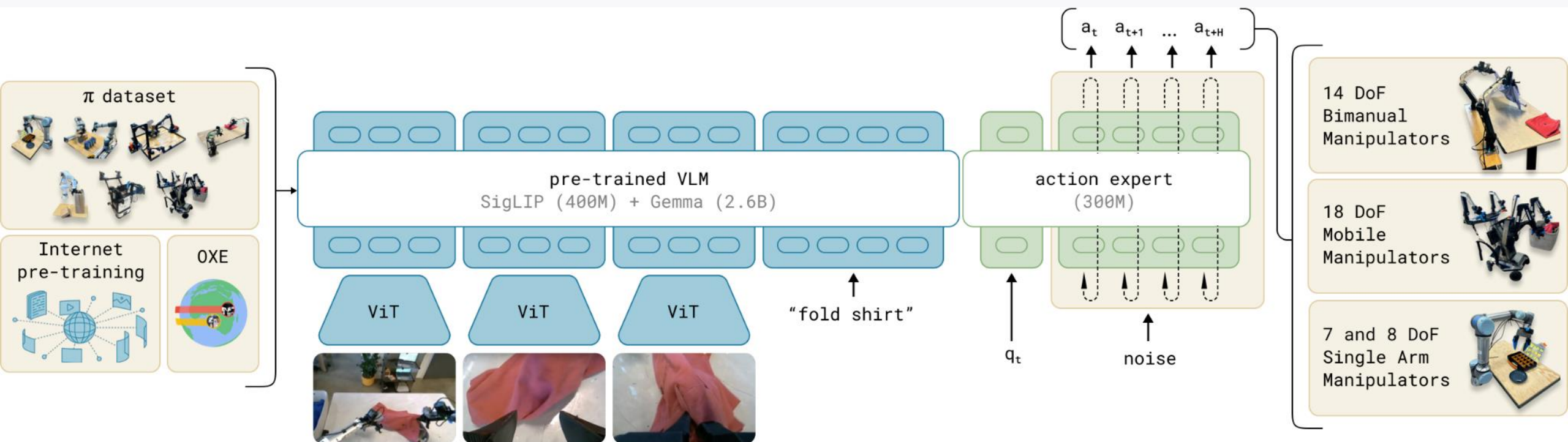
Trénink web/VLM pretraining + cross-embodiment robot data + task-specific fine-tuning

Runtime VLA navrhuje záměr; motion control, constraints a safety layer rozhodují, co se smí skutečně vykonat.

VLA

I. JAK VYPADÁ DNEŠNÍ PHYSICAL AI

Vision–Language–Action model přijímá obraz, jazyk a robotický stav. Výstupem není text, ale akce: často jako tokeny, diffusion/flow trajectory nebo krátký akční chunk.



Generalistický model vs. specialista

I. JAK VYPADÁ DNEŠNÍ PHYSICAL AI

Dnešní modely poskytují dobrý prior pro vnímání, jazyk a širokou paletu manipulací. Průmyslový task ale stále vyžaduje lokální data, evaluaci, runtime a odpovědnost.

Otevřenější stack

- OpenVLA, Octo a SmolVLA: výzkumné baselines a adaptace.
- $\pi 0$ / $\pi 0.5$: flow-based action model a open-world generalization.
- ABC stack: otevřená data, trénink, simulace a evaluace behavior cloningu.
- Výhoda: kontrola nad daty a experimenty; nevýhoda: engineering a compute.

Průmyslové foundation modely

- NVIDIA Isaac GR00T N1.x: humanoidní VLA modely a navazující Isaac stack.
- Gemini Robotics: embodied reasoning a modely pro manipulaci a celé tělo.
- Cosmos: world foundation models pro generování a porozumění fyzickému světu.
- Výhoda: silný ekosystém; otázka: licence, runtime, data rights a vendor lock-in.

Isaac Sim, MuJoCo a Warp řeší různé části simulace

I. JAK VYPADÁ DNEŠNÍ PHYSICAL AI

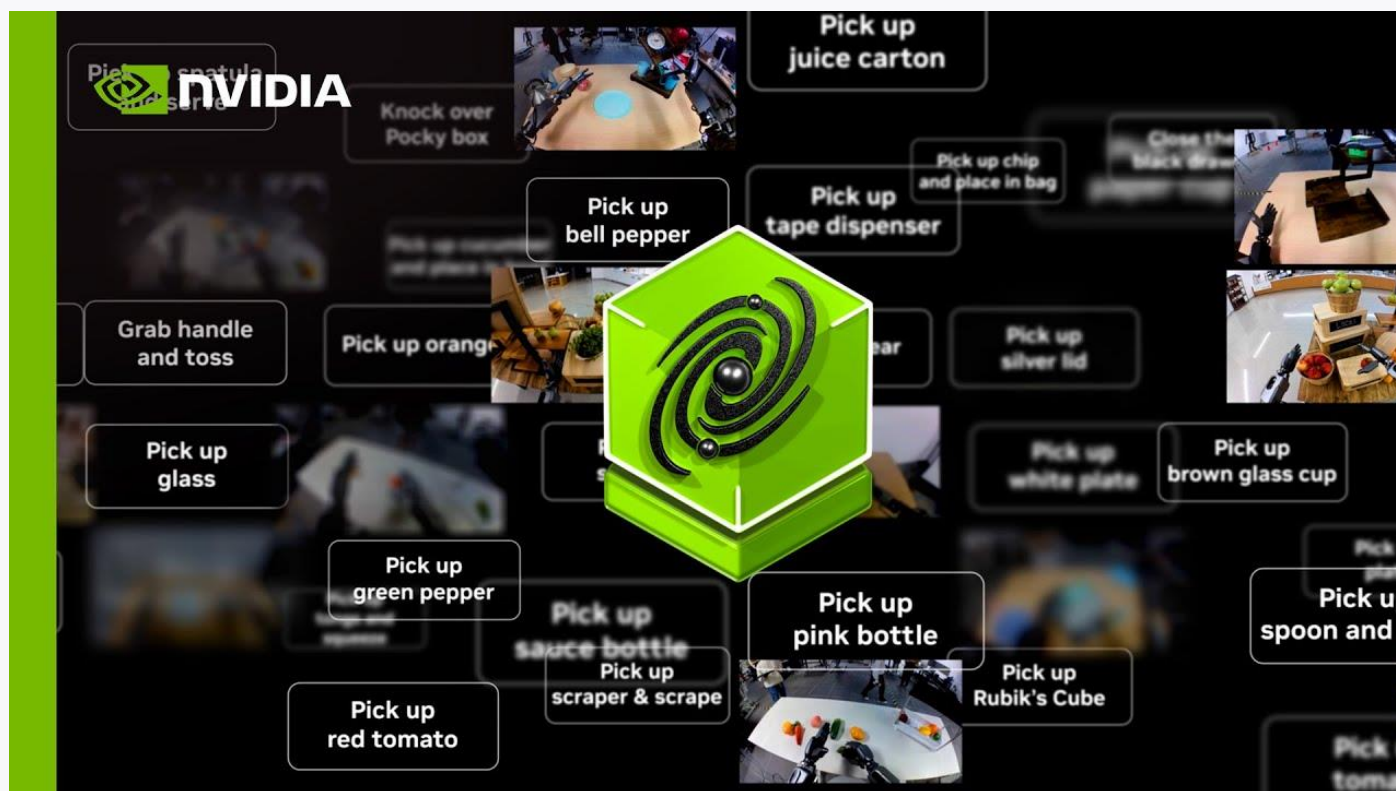
Simulátor volíme podle toho, zda potřebujeme fotorealismus, rychlou dynamiku, differentiable kernels, RL throughput nebo integraci s digitálním dvojčetem.

Nástroj	Silná stránka	Typické použití	Pozor na
NVIDIA Isaac Sim / Lab	USD, senzory, rendering, robotický ekosystém	syntetická data, RL, digital twin	GPU nároky, složitost scény
MuJoCo	rychlá a stabilní kontaktní dynamika	control, RL, reprodukovatelné benchmarky	menší důraz na fotorealismus
NVIDIA Warp	GPU kernel framework, differentiable simulation	vlastní fyzika, masivní paralelizace	není to hotový robotický workflow
Kombinace	rychlé ablace + věrná validační scéna	train fast, validate realistic	musí existovat real-to-sim kontrola

Simulace a Cosmos rozšiřují varianty

I. JAK VYPADÁ DNEŠNÍ PHYSICAL AI

Fyzikální simulace dává controllable ground truth; generativní world model dává rozmanitost a vizuální prior. Nejlepší pipeline používá každý nástroj k jiné otázce.



randomization: geometrie, materiály, osvětlení, tření, sensor noise.

raw data: detekce, segmentace, pose estimation a generování policy.

generování a porozumění video-world trajectories; pro augmentaci a reasoning.

real gap se měří, ne předpokládá. Každá důležitá metrika korelaci s fyzickým robotem.

simulace je levný gate; finální důkaz je vždy closed-loop v cílovém prostředí.

JEPA a world modely učí predikovat relevantní budoucnost

I. JAK VYPADÁ DNEŠNÍ PHYSICAL AI

Místo generování každého pixelu lze predikovat budoucí reprezentaci. To podporuje temporální porozumění, plánování a efektivnější učení z velkého množství videa.

JEPA princip

- encoder mapuje pozorování do latentního prostoru
- predictor odhaduje reprezentaci skrytého nebo budoucího obsahu
- loss se počítá v reprezentaci, ne nutně v pixelech
- výhoda: model se soustředí na stabilní sémantiku a dynamiku

Robotický world model

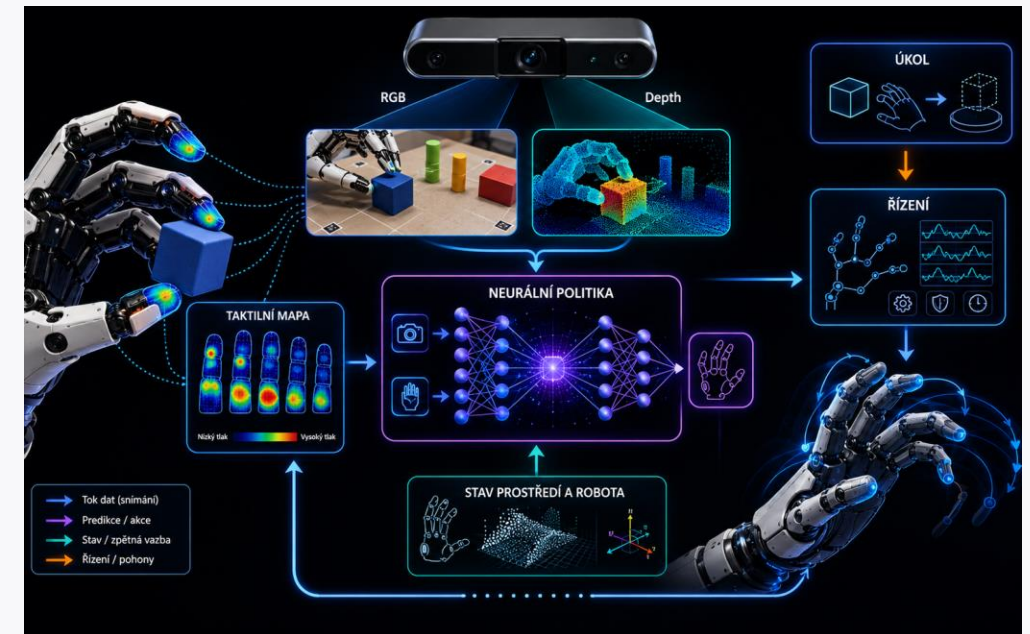
- je action-conditioned: budoucnost závisí na zvolené akci
- umožňuje candidate rollout, plánování a failure prediction
- hlavní problém je drift při dlouhých autoregresivních rolloutech
- PersistWorld z CIIRC stabilizuje multi-step rollout RL post-trainingem

Hmat přidává informaci, kterou kamera během kontaktu ztrácí

I. JAK VYPADÁ DNEŠNÍ PHYSICAL AI

Při zasouvání, skluzu, deformaci nebo manipulaci s průhledným objektem nestačí vizuální obraz. Tactile sensor měří lokální kontakt, tlak, smyk nebo deformaci povrchu.

- Vision-based tactile sensors převádějí deformaci elastomeru na obraz; jiné senzory měří tlakové pole, sílu, vibrace nebo zvuk.
- Hmat umožňuje slip detection, jemné dávkování síly, odhad materiálu a closed-loop korekci úchopu.
- Datový problém: různé senzory, malá veřejná data, drift a obtížná kalibrace mezi kusy.
- Řídicí problém: taktika je rychlá a lokální; nestačí ji připojit jako další obrázek k pomalému VLA modelu.

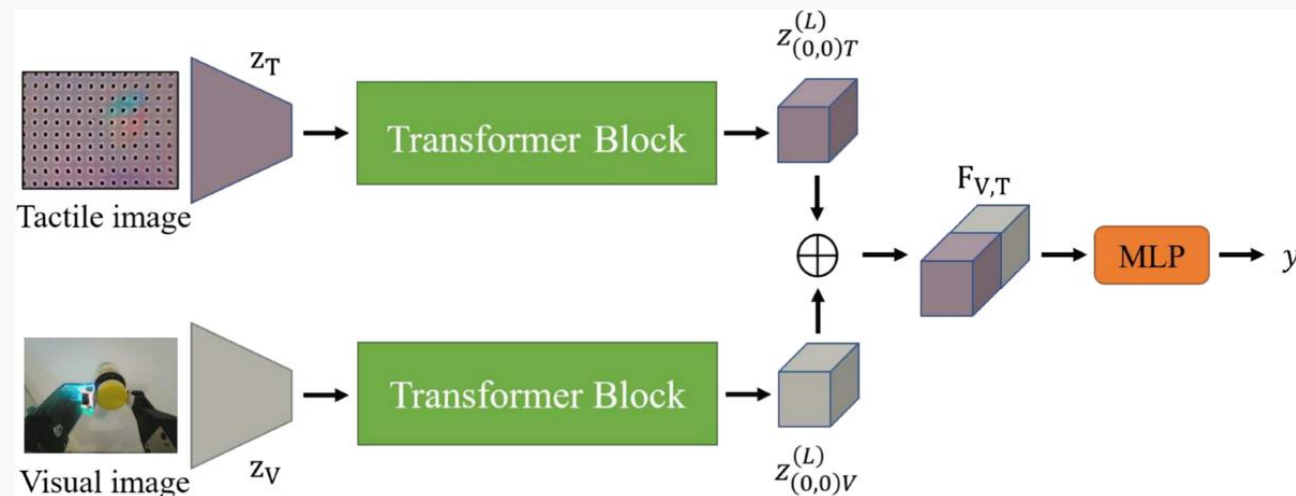


Visuo-tactile transformer musí spojit dvě časové škály

I. JAK VYPADÁ DNEŠNÍ PHYSICAL AI

Vizuální větev říká, co a kde manipulujeme; hmatová větev říká, co se děje v místě kontaktu. Fúze může proběhnout v tokenovém prostoru, přes cross-attention nebo v odděleném rychlém controlleru.

- Early fusion je jednoduchá, ale citlivá na synchronizaci a rozdílné frekvence.
- Cross-attention umožní, aby vizuální kontext vybral relevantní hmatové tokeny.
- Hierarchická architektura drží VLA pro plán a tactile/force controller pro rychlou korekci.
- VLA-Touch a novější práce ukazují, že hmat zlepšuje contact-rich tasky i bez přetrénování celého backbone.



Na CIIRC se propojuje percepce, učení a fyzická evaluace

I. JAK VYPADÁ DNEŠNÍ PHYSICAL AI

Práce skupin Josefa Šivice a Vladimíra Petříka je relevantní právě tím, že nekončí na jednom benchmarku: kombinuje reprezentace, robotická data, manipulaci, simulaci a real-world ověření.

Josef Šivic / IMPACT

- visual learning a reprezentace pro porozumění scéně
- velké robotické modely a efektivní adaptace
- evaluace generalizace a učení bez zapomínání
- world modely pro predikci a plánování

Vladimír Petřík / AAG

- geometrie, optimalizace a manipulace
- robotické učení propojené s kinematikou a fyzikou
- contact-rich a visuo-tactile manipulace
- sim-to-real a experimentální validace na reálných robotech

REALM, RIVET a PersistWorld: tři problémy po prvním demu

I. JAK VYPADÁ DNEŠNÍ PHYSICAL AI

Jakmile policy jednou funguje, začínají těžší otázky: věříme evaluaci, nezapomněli jsme staré skilly a vydrží predikce dost dlouho pro plánování?



REALM

Real-to-sim validovaný benchmark: simulace je užitečná jen tehdy, když její pořadí a failure modes korelují s realitou.

RIVET

Učení nového tasku bez katastrofického zapomínání: specializace nesmí zničit dříve získané schopnosti.

PersistWorld

Stabilnější multi-step rollouts robotického world modelu pomocí RL post-trainingu na vlastních predikcích.



II. REÁLNÝ PRŮMYSLOVÝ USE CASE

Od modelu k odpovědnému provozu

Průmysl nekupuje VLA model. Kupuje bezpečně provedený proces s měřitelným výsledkem, jasným vlastníkem a plánem pro výjimky.

Firmy nejvíc netrápí robot, ale riziko nehotového systému

II. REÁLNÝ PRŮMYSLVÝ USE CASE

Technická fascinace humanoidem rychle narazí na otázky provozu. Úspěšný discovery workshop proto začíná procesem, metrikami a odpovědností.

- ROI a kapacita: kolik stojí jedna dokončená operace včetně integrace, dozoru a prostojů.
- Variabilita: kolik SKU, variant, orientací, světelných podmínek a nečekaných stavů se vyskytuje.
- Spolehlivost: co se stane po 99 úspěších, když stý pokus selže u rozpracovaného dílu.
- Integrace: PLC, MES/WMS, bezpečnostní zóny, pracovní instrukce, přístupová práva a audit log.
- Odpovědnost: kdo řeší incident, model update, změnu procesu, servis a formální převzetí.

Dnes se vyplatí úlohy s variabilitou, ale měřitelným koncem

II. REÁLNÝ PRŮMYSLOVÝ USE CASE

Nejlepší první use case není nejtěžší demo. Je to úloha, kde learned policy řeší variabilitu a zbytek systému dokáže omezit riziko.

Use case	Proč dává smysl	Hlavní data	První KPI
Inspekce / průvodce	mobilita, HRI, nízká fyzická síla	mapa, trasa, RAG, interakce	pokrytí, zásahy, kvalita odpovědi
Sorting / kitting	proměnlivé objekty, jasný výsledek	teleop + obraz + task metadata	success, cycle time, recovery
KLT / intralogistika	mobilita + manipulace	bimanuál, PLC/WMS event	přepravka/h, poškození
Machine tending	opakovatelný stroj, variabilní manipulace	pose, interlock, force	availability, bezpečný stav
Kabely / insertion	contact-rich, obtížná klasická automatizace	teleop + tactile/force	úspěch, síla, čas

Jak vybrat dobrý use case

II. REÁLNÝ PRŮMYSLOVÝ USE CASE

Před PoC musí být jasné, co znamená úspěch a zda lze proces bezpečně ohraničit. Pokud nelze vytvořit eval protokol, neumíme ani dokázat zlepšení.

Dobré signály

- patná ergonomie, nebezpečné úkony s provozními daty
- výsledek lze automaticky nebo rychle manuálně ověřit
- variabilita je velká pro skript, ale stále omezená doménou
- lze zavést bezpečný fallback a dočasně akceptovat teleop assist
- zákazník povolí sběr edge cases a řízené model updates

Varovné signály

- požadavek na univerzálního humanoida bez konkrétního procesu
- nulová tolerance k zásahu, ale žádná historická data o výjimkách
- extrémní takt a síly, kde je účelovější pevná automatizace
- nejasné právo k videu, výrobním datům nebo zákaznickému know-how
- safety a CE se mají řešit až po hotovém demu

Humanoid pro prohlídky a kontakt s lidmi

II. REÁLNÝ PRŮMYSLOVÝ USE CASE

Záměrně jsme začali úlohou, která kombinuje mobilitu, sociální interakci a práci se znalostmi, ale nevyžaduje zatím těžkou silovou manipulaci.

- Robot se lokalizuje v budově, naviguje mezi waypointy a průběžně se vyhýbá překážkám a lidem.
- Mission layer řídí trasu, zastávky, prezentaci a návrat do výchozího bodu nebo k nabíjení.
- Návštěvník se ptá hlasem; ASR převede dotaz a LLM odpovídá z řízené znalostní databáze přes RAG.
- Robot využívá gesta, orientaci těla a obrazovku/hlas, aby bylo zřejmé, komu a o čem odpovídá.
- Každá nejistota má fallback: upřesňující otázka, přivolání člověka nebo bezpečné ukončení mise.



Průvodce je systém šesti služeb, ne jedna konverzační aplikace

II. REÁLNÝ PRŮMYSLOVÝ USE CASE

Rozdělení služeb umožňuje samostatně testovat navigaci, znalosti, dialog i bezpečný stav. LLM nikdy přímo neposílá motorické příkazy.

- Dialog: ASR/TTS, správa kontextu, filtrování a potvrzení záměru.
- Knowledge layer: kurátorovaný obsah, zdroje, verze, přístupová práva a RAG.
- Mission layer: stavový automat pro trasu, zastávky, časování a výjimky.
- Navigation: mapa, lokalizace, Nav2, human-aware plánování a docking.
- Safety/operations: rychlostní limity, zóny, E-stop, monitoring a incident replay.

Průvodce

II. REÁLNÝ PRŮMYSLOVÝ USE CASE

Výzkumně zajímavé komponenty mohou fungovat jednotlivě, ale kvalitu zkušenosti určuje přechod mezi nimi: ztráta lokalizace, přerušení dialogu, člověk v cestě nebo neaktuální znalost.

- State management musí přežít restart, ztrátu sítě i změnu trasy.
- RAG potřebuje vlastnictví obsahu, datum platnosti a odpověď „nevím“ — ne pouze větší LLM.
- Navigace v davu je sociální problém: rychlost, odstup, gesta a předvídatelnost pohybu.
- Telemetrie musí spojit dialog, pozici, mapu a systémové události do jednoho incident replay.
- Remote assist je součástí provozního návrhu, nikoli přiznání neúspěchu autonomie.



Další vlna je průmyslovější: manipulace, logistika a kontakt

II. REÁLNÝ PRŮMYSLOVÝ USE CASE

Z prohlídek přenášíme provozní vrstvy — mise, monitoring, fallback a integraci. Nově přidáváme robotické epizody, precise pose, force/tactile a přísnější akceptační metriky.

- Sorting a kitting proměnlivých objektů do definovaných pozic.
- KLT manipulace, intralogistika a předávání materiálu mezi pracovišti.
- Machine tending se stavem stroje a bezpečnostním interlockem.
- Kably, insertion, nalévání a laboratorní úlohy s hmatem a silou.
- Inspekce strojů a prostředí s automatickým reportem a eskalací odchylky.

Sorting na Unitree G1-D: důkaz celé učící cesty

II. REÁLNÝ PRŮMYSLOVÝ USE CASE

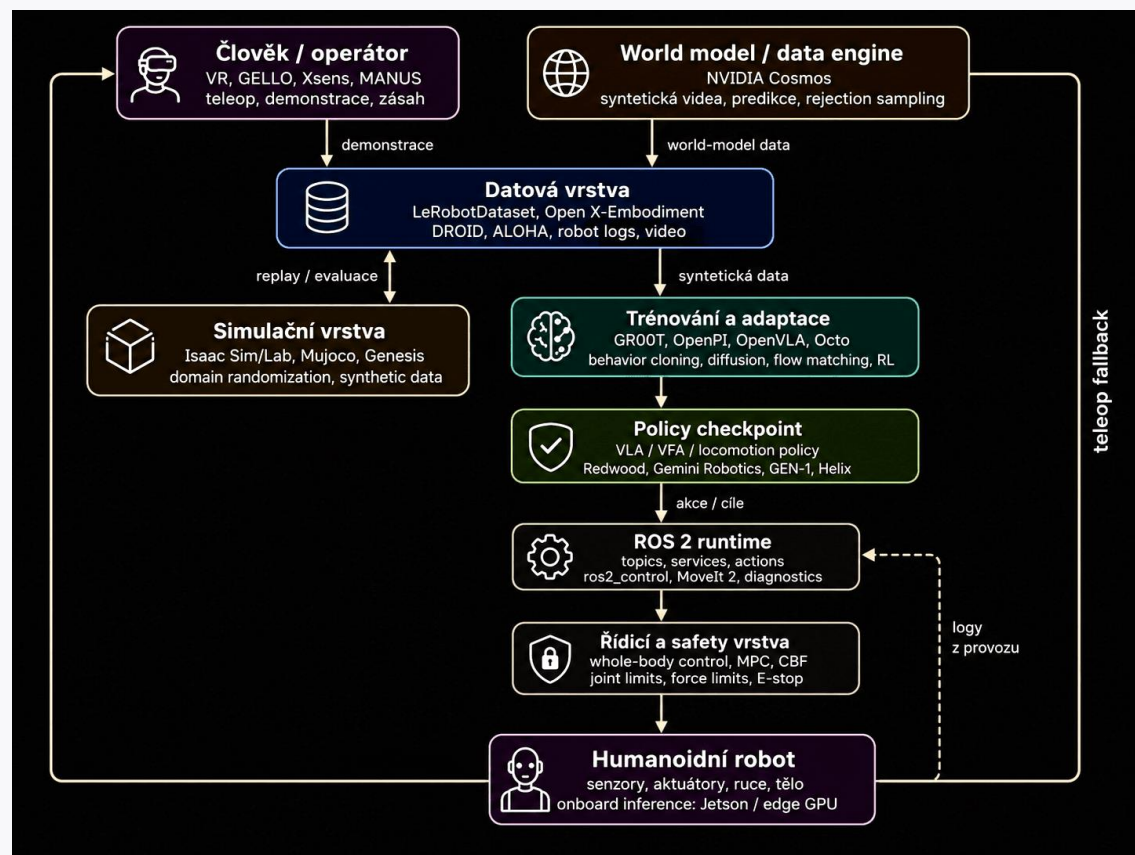
Referenční task není cenný jen tím, že robot přemístí předmět. Cenné je, že lze reprodukovat sběr, trénink, evaluaci a nasazení jako jeden workflow.

- Operátor předvede varianty pick–place tasku a systém uloží synchronizované epizody.
- Data projdou segmentací, QA a kontrolou pokrytí objektů, pozic a pořadí.
- $\pi 0.5$ / VLA policy se adaptuje na cílový setup a akční prostor.
- Closed-loop test měří success, cycle time, generalizaci, zásahy a recovery.
- Další krok: robustní automatická evaluace v simulaci a evidence korelace s reálným robotem.

Referenční architektura odděluje data, skill a deployment

II. REÁLNÝ PRŮMYSLOVÝ USE CASE

Největší technologická hodnota není jeden robot ani jedna VLA architektura. Je to opakovatelná schopnost rychle vytvořit, ověřit, nasadit a dále zlepšovat specialistu.



- Data engine: operátor, GELLO/UMI/VR, metadata, episode store, QA a coverage.
- Skill factory: adaptace VLA/policy, simulace, evaluace, failure mining a recovery.
- Deployment stack: hardware abstraction, ROS 2, navigace, motion control a safety envelope.
- Factory integration: PLC, MES/WMS, digitální dvojče, identity a audit.
- Operations: monitoring, remote assist, incident replay a řízený release modelu.

Učící pipeline

II. REÁLNÝ PRŮMYSLOVÝ USE CASE

Pipeline je současně technický proces a řízení rizika. V každé fázi vzniká artefakt, který umožní rozhodnout: pokračovat, doplnit data, změnit návrh, nebo zastavit.

Fáze	Hlavní výstup	Primární kontrola
1 Task specification	stav, akce, tolerance, výjimky, business KPI	jednoznačný eval protokol
2 Sběr dat	demonstrace + procesní metadata	epizody/h, zátěž operátora
3 Data QA	synchronizace, segmentace, coverage	% použitelných epizod
4 Trénink	checkpoint, config, dataset version	time-to-first-usable policy
5 Evaluace	varianty, failure modes, recovery	success, cycle time, interence
6 Nasazení	runtime, safety envelope, integrace	go / re-collect / redesign
7 Improve	incident → eval case → nový release	regression suite + release gate

Sběr dat musí co nejméně brzdit operátora

II. REÁLNÝ PRŮMYSL OVÝ USE CASE

Data engine je dobrý tehdy, když sbírá kontext v přirozeném rytmu práce a přitom dovolí rychle označit zásah, chybu nebo nový edge case.

Automatizovat při sběru

- one-touch start/stop, automatický pre/post-roll
- pasivní PLC/MES metadata: zakázka, SKU, stav stroje, nástroj
- kontrola synchronizace, kalibrace a health stavu senzorů
- automatická segmentace dlouhého záznamu na tasky a subtasky
- okamžitá kontrola, zda máme nové pokrytí, nebo jen duplicitu

Nechat člověku

- jedním gestem označit success/failure a důvod zásahu
- potvrdit neobvyklou situaci a začátek recovery segmentu
- zastavit sběr při změně procesu nebo bezpečnostním riziku
- doplnit sémantický kontext, který nelze vyčíst ze senzorů
- schválit, co se smí použít pro další učení a sdílení

Evaluace musí předvídat provozní selhání, ne jen průměrný success

II. REÁLNÝ PRŮMYSLOVÝ USE CASE

Eval suite kombinuje systematické varianty, reálné chyby, recovery a srovnání se skriptem, člověkem nebo teleoperací.

-
- Varianty: objekty, pozice, orientace, osvětlení, clutter, pořadí a stav nástroje.
 - Metriky: success, cycle time, peak force, počet intervencí, poškození a bezpečný stav.
 - Failure taxonomy: percepce, plán, grasp, kontakt, koordinace, integrace, latence.
 - Každý incident se musí stát reprodukovatelným eval casem; jinak se servis nepřeklápí do know-how.
 - Simulace slouží jako levný regression gate pouze po real-to-sim validaci.

Pilot má předem definované gate i důvod pro stop

II. REÁLNÝ PRŮMYSLOVÝ USE CASE

Bez go/no-go kritérií se pilot snadno změní v nekonečnou demonstraci. Každá fáze musí snížit konkrétní nejistotu.

Gate	Otázka	Důkaz	Možné rozhodnutí
Discovery	Je ROI reálné?	procesní data + shortlist	assessment / stop
Assessment	Lze task ohraničit, měřit a zabezpečit?	feasibility + data/safety plán	PoC / redesign / stop
Lab PoC	existuje použitelný skill?	model + eval report + demo	pilot / doplnit data
Pilot	funguje v provozní distribuci?	metriky, incidenty, integrace	nasazení / doučení
Akceptace	přebírá zákazník provoz?	SLA, dokumentace, safe state	rollout / omezený provoz

Safety je vlastnost celého systému, ne confidence score modelu

II. REÁLNÝ PRŮMYSLOVÝ USE CASE

Naučená policy je pouze jedna část. Bezpečnostní argument musí zahrnout mechaniku, rychlost, sílu, nástroj, prostor, člověka, software, integraci a provozní postup.

- AI policy běží uvnitř omezeného akčního prostoru: limity kloubů, rychlosti, pracovního prostoru a povolených kontaktů.
- Nezávislá safety vrstva sleduje zóny, interlocky, E-stop a safe state; nesmí být závislá na správném výstupu VLA.
- Risk assessment pracuje s konkrétní aplikací: stejný humanoid s kamerou a s ostrým nástrojem je jiný systém.
- Safety validation musí zahrnout normální provoz, foreseeable misuse, ztrátu senzorů, sítě, lokalizace i model runtime.
- Každý update modelu nebo procesu může měnit riziko — proto potřebuje regression test a release approval.

CE a certifikace

II. REÁLNÝ PRŮMYSLOVÝ USE CASE

Konkrétní právní a normativní režim závisí na určeném použití a integraci. Prakticky je nutné udržovat trasovatelnou dokumentaci rizik, požadavků, testů a změn.

Oblast	Co řešit v projektu	Typický důkaz
Strojní bezpečnost	hazards, protective measures, safe state, validace	risk assessment, test protocols, technical file
Robotická aplikace	buňka vs. spolupráce, síla/rychlost, zóny	application-specific measurements and validation
Software a AI	verze modelu, dataset, OOD, update, logování	model card, eval suite, release record
Cybersecurity	identity, update chain, remote access, incident response	threat model, access log, patch process
Data a soukromí	video lidí, účel učení, retention, práva	DPIA/records, access policy, anonymization
EU změny	Machinery Regulation od 20. 1. 2027; AI Act dle role/use case	právní posouzení a mapping povinností

Provoz potřebuje jasné L1/L2/L3 vlastnictví

II. REÁLNÝ PRŮMYSLOVÝ USE CASE

Bez servisního modelu se každá chyba vrací k výzkumníkovi. Rozdělení odpovědnosti chrání zákazníka, vývojový tým i rychlost zlepšování.

Úroveň	Owner	Typická odpovědnost	Výstup
L1 — on-site	lokální integrátor / OEM	restart, fyzická inspekce, díly, safe state	obnovení nebo eskalace
L2 — deployment	solution/deployment tým	logy, konfigurace, mise, kalibrace, release	incident report + workaround
L3 — Physical AI	core model/data tým	policy failure, data, retraining, eval, nový skill	ověřený nový checkpoint
Release owner	odpovědná osoba	regression, safety gate, rollback a komunikace	schválený deployment packet

Každý edge case má zlepšit další release

II. REÁLNÝ PRŮMYSLOVÝ USE CASE

Dlouhodobá výhoda vzniká až uzavřenou datovou smyčkou. Incident nesmí skončit restartem; musí vést k reprodukci, novému testu a řízenému zlepšení.



Release se nasazuje pouze s identitou datasetu, modelu, runtime, safety konfigurace a rollback plánem. Učení v produkci neznamena nekontrolované online měnění chování.

Otázky?

KONTAKT

info@ichr.cz · ichr.cz



Václav Kubáček

ICHR · CIIRC ČVUT
vaclav.kubacek@cvut.cz



Lukáš Navara

ICHR · CIIRC ČVUT
Lukas.navara@cvut.cz