

RAČUNALNIŠKI VID IN UMETNA INTELIGENCA NA SLOVENSKI SUPERRAČUNALNIŠKI INFRASTRUKTURI

JANEZ PERŠ

Univerza v Ljubljani, Fakulteta za elektrotehniko, Ljubljana, Slovenija
janez.pers@fe.uni-lj.si

Prispevek podaja pregled možnosti za izvajanje računske zahtevnih nalog računalniškega vida in umetne inteligence v Sloveniji na infrastrukturi SLING. Osredotočimo se na dve ključni okolji: petaflopni superračunalnik Vega (240 GPU Nvidia A100 in 960 vozlišč CPU) ter Arnesovo računsko gručo s starejšimi (V100) in novejšimi (H100) GPU za interaktivno in izobraževalno rabo. Opišemo postopke pridobivanja dostopa (testni dostop, razvojni dostop, redni/veliki raziskovalni dostop) in značilne kvote, ki vplivajo na razpoložljivost GPU virov. Pri tem povzamemo tudi vlogo Nacionalnega kompetenčnega centra (EuroCC 2) pri izobraževanjih in podpori uporabnikom. Posebno pozornost namenimo programski združljivosti za računalniški vid (knjižnice PyTorch/TensorFlow, knjižnice CUDA, kontejnerji) ter razlikam med klasičnim paketnim režimom (Slurm) in potrebami sodobnih potekov dela. V zaključku stanje povežemo s projektom Slovenske tovarne umetne inteligence (SLAIF), ki načrtuje nov, za umetno inteligenco optimiziran superračunalnik ter hibridno HPC-oblačno okolje z orkestracijo Kubernetes.

DOI
[https://doi.org/
10.18690/um.feri.4.2026.1](https://doi.org/10.18690/um.feri.4.2026.1)

ISBN
978-961-299-111-1

Ključne besede:
računalniški vid,
umetna inteligenca,
superračunalništvo,
SLING,
Vega,
Arnes,
SLAIF,
Slurm,
Kubernetes

DOI
[https://doi.org/
10.18690/um.feri.4.2026.1](https://doi.org/10.18690/um.feri.4.2026.1)

ISBN
978-961-299-111-1

Keywords:

computer vision,
artificial intelligence,
supercomputing,
SLING,
Vega,
Arnes,
SLAIF,
Slurm,
Kubernetes

COMPUTER VISION AND AI ON SLOVENIAN SUPERCOMPUTING INFRASTRUCTURE

JANEZ PERŠ

University of Ljubljana, Faculty of Electrical Engineering, Ljubljana, Slovenija
janez.pers@fe.uni-lj.si

This paper reviews how computer vision and AI workloads can be executed in Slovenia on the national SLING infrastructure. We focus on two key environments: the EuroHPC petascale system Vega (240 Nvidia A100 GPUs and 960 CPU nodes) and the Arnes HPC cluster, which includes V100 and some H100-based GPU nodes and serves both education and research. We summarise access paths (Test Access, Development Access, Regular/Large Research Access) and the allocation quotas that affect GPU availability for deep learning training. We also note the role of the Slovenian National Competence Centre (EuroCC 2) in training and user support. We discuss software compatibility for CV (PyTorch/TensorFlow, CUDA libraries, containers) and highlight the gap between batch scheduling (Slurm) and modern AI workflows that require interactive development and long-running inference services. Finally, we outline the perspective of the Slovenian AI Factory (SLAIF), which plans an AI-optimised supercomputer and a hybrid HPC-cloud environment with Kubernetes-based orchestration.

1 Uvod

Računalniški vid (RV) je že dolgo eno od osrednjih področij umetne inteligence (UI) – od detekcije objektov in segmentacije do generativnih modelov, ki iz slik, videa ali 3D podatkov izdelujejo nove predstavitve. Skupna značilnost sodobnih pristopov je velika računska zahtevnost: učenje globokih nevronske mreže in transformerjev praviloma zahteva zmogljive grafične pospeševalnike (GPU), dovolj hitro diskovno in omrežno infrastrukturo ter ustrezno programsko okolje.

V Sloveniji je temeljni javni okvir za takšne naloge nacionalna superračunalniška mreža SLING, ki vključuje EuroHPC superračunalnik Vega in Arnesovo računsko gručo, za izbrane projekte pa tudi dostop do dodatnih sistemov prek evropskih mehanizmov odprtega dostopa. Razpoložljivost virov in način pridobivanja dostopa neposredno vplivata na izvedljivost (in časovnico) eksperimentov računalniškega vida, zato sta v prispevku obravnavana skupaj z vprašanjem programske združljivosti.

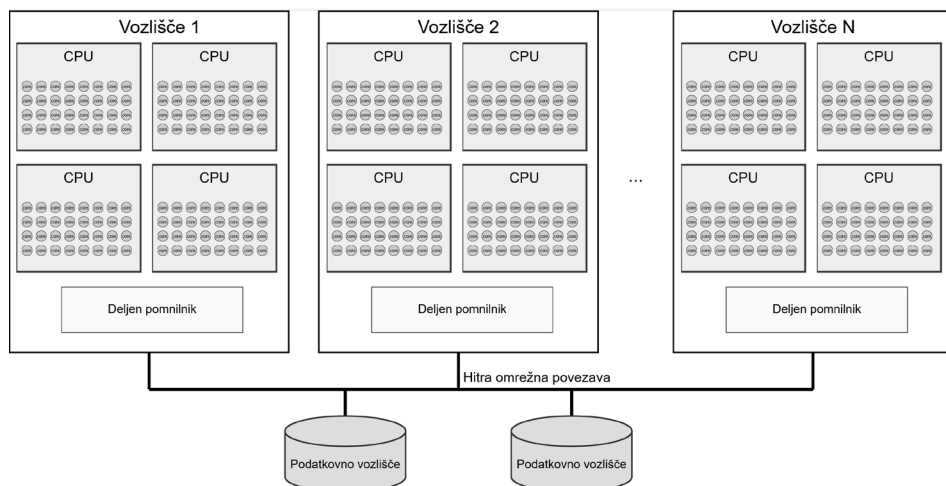
2 Kaj točno je superračunalnik?

Superračunalnik (HPC, angl. high performance computer), prikazan na sliki 1, ni ena “posebna” naprava, temveč velika gruča (angl. cluster) računalniških vozlišč, povezanih z zelo hitrim omrežjem. Vsako vozlišče ima več procesorjev (CPU), pogosto tudi grafične procesne enote (GPU) in deljen pomnilnik. Izjemna zmogljivost ne izhaja iz hitrejših jeder, temveč iz zelo velikega števila jeder, ki lahko računajo vzporedno. Takšna infrastruktura je tehnično zelo podobna tisti pri ponudnikih oblčnih storitev v t.i. podatkovnih centrih (angl. datacenter).

Ključna razlika je v načinu uporabe:

- V oblaku si več uporabnikov hkrati deli dele posameznih vozlišč (virtualizacija).
- Pri superračunalniku pa sistem uporabniku za določen čas dodeli celotna vozlišča ali več vozlišč, medtem ko drugi čakajo v čakalni vrsti.

Zato mora biti problem primeren za masivno vzporedno izvajanje – razdeljen na veliko manjših nalog, ki tečejo hkrati na številnih jedrih ali več GPU.



Slika 1: Arhitektura superračunalnika ni zelo drugačna od arhitekture tipičnega podatkovnega centra, razlika je v načinu dostopa in uporabe.

Vir slike: lasten.

3 Slovenska superračunalniška infrastruktura SLING

Konzorcij SLING (Slovensko nacionalno superračunalniško omrežje) ureja odprti dostop do dela nacionalnih kapacitet in uvaja enoten okvir za prijavo projektov, dodelitev virov in podporo uporabnikom. Ključna prednost takšnega pristopa je, da se lahko uporabnik od začetnega testiranja hitro premakne do večjih dodelitev, hkrati pa se ohranjata sledljivost in transparentnost porabe virov. (SLING, n.d.-b)

3.1 Modeli dostopa in kvote (od testiranja do produkcije)

Razpisi SLING ločijo več tipov dostopa, ki se razlikujejo po namenu in obsegu dodeljenih virov: testni dostop (TA) je namenjen spoznavanju okolja in preverjanju oziroma validaciji programske opreme, razvojni dostop (DA) razvoju in optimizaciji aplikacij, medtem ko sta redni raziskovalni dostop (RRA) in veliki raziskovalni dostop (LRA) predvidena za produkcijsko rabo ter obsežnejše raziskovalne projekte z večjimi dodelitvami virov. Za GPU-intenzivne naloge so pomembne kvote na posamezno dodelitev – npr. v okviru RRA je za Vega-GPU predvidena dodelitev do 25.000 vozliščnih, za Arnes-GPU pa do 10.000 vozliščnih ur na dostop. (SLING, n.d.-b)

V praksi to pomeni, da je smiselno projekt zasnovati iterativno: najprej izvesti manjše teste (npr. preizkus učnih skript, preverjanje odvisnosti, merjenje porabe pomnilnika GPU), nato pa na podlagi rezultatov utemeljiti zahtevo za večjo dodelitev. Tak postopek je pri računalniškem vidu posebej uporaben, ker je strojni učinek (npr. izbira velikosti paketov (batch), uporaba mešane natančnosti, I/O ozka grla pri branju slik ali videa) pogosto odločilen. (SLING, n.d.-b)

3.2 HPC Vega

Vega je petaflopni sistem, ki ga gosti IZUM v Mariboru. Po javno objavljenih specifikacijah ima 960 vozlišč CPU (skupaj 1.920 procesorjev AMD EPYC 7H12 in približno 122.000 jeder) ter 60 vozlišč GPU s skupno 240 grafičnimi pospeševalniki Nvidia A100 s 40 GB VRAM. Zmogljivost je navedena kot 6,9 PFLOPS (trajno) oziroma 10,1 PFLOPS (špice). (IZUM, n.d.-a)

Za računalniški vid je ključna tudi programska oprema: Vega kot delovno okolje navaja podporo za tipične okvirje in knjižnice, potrebne za UI, kot so TensorFlow in PyTorch, ter ekosistem NVIDIA (CUDA, cuDNN, NCCL, TensorRT, DALI, DeepStream). To poenostavi postavitev eksperimentov (zlasti pri uporabi GPU) in zmanjša potrebo po ročnem prevajanju odvisnosti. (IZUM, n.d.-a)

Za reproducibilnost in prenosljivost okolij so pomembni kontejnerji. Dokumentacija sistema navaja, da je na Vegi nameščena in prilagojena platforma Singularity, ki omogoča zaganjanje kontejnerjev v okolju HPC. (HPC Vega documentation, n.d.-b)

Za delitev GPU na manjše enote je pogosto uporabna funkcija Multi-Instance GPU (MIG), zlasti pri manjših modelih RV ali pri inferenci. Dokumentacija Vege navaja, da MIG trenutno ni omogočen, zato je delitev GPU tipično organizirana na ravni dodeljevanja celotnih vozlišč GPU ali GPU-jev znotraj vozlišča. (HPC Vega documentation, n.d.-a)

Uporabniški vmesnik je tradicionalno ukazna vrstica (Slurm), vendar dokumentacija opisuje tudi storitev Open OnDemand, ki omogoča oddajo in spremljanje opravil ter upravljanje datotek prek spletnega brskalnika. (HPC Vega documentation, n.d.-c)

3.3 Arnesova gruča HPC

Arnesova računska gruča je po opisu SLING drugi najzmogljivejši sistem v Sloveniji in je namenjena tako izobraževalnim kot raziskovalnim uporabnikom, pri čemer je med raziskovalnimi področji posebej izpostavljena uporaba v umetni inteligenci in obdelavi velikih podatkov. (SLING, n.d.-a)

Objavljene specifikacije izpostavljajo heterogen nabor vozlišč GPU: 24 vozlišč GPU z $2 \times$ Nvidia V100S, 7 vozlišč GPU z $2 \times$ Nvidia H100 PCIe in 3 vozlišča GPU s $4 \times$ Nvidia H100 SXM5. Sistem uporablja Slurm (navedeni so tudi 'tagi' Slurm za posamezne tipe vozlišč), dostop pa zahteva dvostopenjsko avtentikacijo (2FA). (SLING, n.d.-a)

Tabela 1: Primerjava ključnih okolij za UI/RV v Sloveniji in v projektu SLAIF.

Okolje	Strojna oprema (izbor)	Programski sloj za UI/RV (izbor)	Dostopnost / opombe (izbor)
HPC Vega (IZUM)	60 vozlišč GPU (skupaj $240 \times$ Nvidia A100); 960 vozlišč CPU; 6,9 PFLOPS (10,1 PFLOPS peak).	Slurm; TensorFlow, PyTorch; CUDA, cuDNN, NCCL, TensorRT; DALI, DeepStream; Singularity; MIG ni omogočen.	SLING + EuroHPC; RRA do 25.000 GPU-vozliščnih ur; spletni dostop tudi prek Open OnDemand.
HPC Arnes	GPU: $24 \times (2 \times V100S)$, $7 \times (2 \times H100 \text{ PCIe})$, $3 \times (4 \times H100 \text{ SXM5})$; CPU: heterogen nabor vozlišč EPYC/Xeon.	Slurm; heterogeni profili GPU; primeren za učenje, razvoj in eksperimente; 2FA zahtevan.	razpisi SLING; RRA do 10.000 GPU-vozliščnih ur; pogosto uporabljen v izobraževanju in raziskavah UI.
HPC »Vega 2« (načrtovano v okviru SLAIF)	≈ 100 PFLOPS (HPC) in ≈ 10 exaFLOPS (mešana natančnost za UI); ≈ 2000 pospeševalnikov; 10 PB hitrega diska + 100 PB shrambe Ceph; »cloud« particija.	Hibridni HPC-cloud; poudarek na storitvah in API; orkestracija kontejnerjev s Kubernetes; namenjeno tudi velikim modelom UI in inferenci.	Novi Arnesov podatkovni center pri HE Mariborski otok; ciljna uvedba 2027 (vir navajajo konec leta 2027).

Viri: IZUM (n.d.-a), HPC Vega documentation (n.d.-a; n.d.-b; n.d.-c), SLING (n.d.-a; n.d.-b), IZUM (2025), EuroHPC JU (2025), SLAIF (n.d.), Univerza v Ljubljani (2025).

4 Dostopnost virov in praktični vidik razpoložljivosti

Za uporabnika računalniškega vida je razpoložljivost v praksi kombinacija (i) dodeljene kvote (npr. vozliščne ure na particiji GPU), (ii) politike prioritet in omejitev (časovne omejitve opravil, število vozlišč, število sočasnih opravil), ter (iii)

trenutne zasedenosti sistema. Ker sta Vega in Arnes deljena sistema, je tipičen način dela oddaja opravil v paketni obliki in čakanje v vrsti, kar je bistveno drugače kot pri komercialnih oblakih.

Razpisi SLING so zato zasnovani tako, da podpirajo postopno uvajanje uporabnika: TA in DA sta stalno odprta razpisa z relativno manjšimi dodelitvami, namenjenimi testiranju in razvoju, medtem ko sta RRA in LRA časovno vezana na razpisne roke in omogočata večje dodelitve za produkcijske eksperimente. (SLING, n.d.-b)

Pomemben del dostopnosti je tudi organizacijska in izobraževalna podpora. Slovenski Nacionalni kompetenčni center (NCC) deluje znotraj SLING in je v okviru projekta EuroCC 2 organiziran kot konzorcij partnerjev (Arnes kot vodilni, IJS, univerze in drugi). NCC organizira brezplačna izobraževanja na osnovni, napredni in višji ravni ter pomaga uporabnikom pri uvajanju v okolje HPC. Kot primer specializiranih vsebin sta bila v letu 2025 izvedena »bootcampa« za multi-GPU programiranje in profiliranje programske opreme za UI. (SLING, n.d.-c; EuroCC ACCESS, n.d.; Univerza v Ljubljani FMF, n.d.; IJS Indico, 2025a; IJS Indico, 2025b)

Pri načrtovanju eksperimentov RV je priporočljivo, da se dodelitve vežejo na merljive cilje (npr. število epoh, število poskusov pri optimizaciji hiperparametrov, velikost učne množice) in da se že v fazi TA/DA pripravita avtomatizacija zagona (npr. opravila Slurm) ter zbiranje metrik. S tem se zmanjša tveganje, da bi večja dodelitev RRA/LRA ostala neizkoriščena zaradi programskih ali podatkovnih ozkih grl.

Tipični primeri uporabe superračunalniških virov v računalniškem vidu vključujejo:

- učenje in doučevanje (angl. fine-tuning) globokih modelov na velikih slikovnih/video zbirkah (GPU),
- razširjene eksperimente z več modeli in različnimi hiperparametri (več vzporednih opravil),
- porazdeljeno učenje (multi-GPU) pri večjih modelih ali višjih ločljivostih,
- obsežno inferenco (batch) ali pripravo rezultatov za nadaljnje analize.

5 Programska združljivost za računalniški vid

Združljivost v okolju HPC pomeni predvsem: (i) razpoložljivost ustreznih gonilnikov in okolja CUDA za GPU, (ii) dostop do knjižnic za globoko učenje in računalniški vid, (iii) možnost reprodukcije okolja, ter (iv) učinkovito upravljanje podatkov (lokacija, prepustnost, paralelni dostop).

Vega kot primer navaja podporo za PyTorch in TensorFlow ter nabor knjižnic NVIDIA, ki so pomembne za poteke dela v RV. (IZUM, n.d.-a)

Dokumentacija Vege izrecno opisuje uporabo kontejnerjev Singularity, kar omogoča, da si uporabnik pripravi lastno reproducibilno okolje (npr. natančne verzije PyTorch, CUDA, OpenCV, mmcv itd.) in ga nato zaganja v opravilih Slurm brez poseganja v sistemske module. (HPC Vega documentation, n.d.-b)

Za uporabnike, ki preferirajo grafični dostop, je na Vegi na voljo Open OnDemand, ki omogoča oddajo opravil in upravljanje datotek prek brskalnika. To je uporabno pri učnih delavnicah in pri začetnih fazah projektov, ko je cilj hitra validacija okolja. (HPC Vega documentation, n.d.-c)

6 SLURM kot standard HPC in omejitve pri potekih dela UI

Večina superračunalniških okolij uporablja specializiran razporejevalnik opravil (workload manager), ki skrbi za dodeljevanje virov in čakalne vrste. Slurm je eden najbolj razširjenih: poročila navajajo, da ga uporablja približno polovica lokacij HPC v vzorcu več tisoč sistemov (Hyperion Research, 2023), CNCF pa navaja oceno, da Slurm poganja okoli 65 % sistemov s seznama TOP500 (CNCF, 2024).

Slurm je zasnovan kot upravljalnik virov in razporejevalnik opravil na gručah Linux: dodeli računska vozlišča za določeno časovno obdobje, zažene/spremlja izvajanje in upravlja čakalno vrsto ter politiko deljenja virov. (SchedMD, n.d.)

Takšen model je zelo učinkovit za klasične obremenitve HPC (paketne simulacije, opravila MPI), vendar pri UI pogosto nastopijo dodatne zahteve: interaktivno eksperimentiranje (npr. notebooki), cevovodi, orkestracija več korakov (priprava podatkov – učenje – evalvacija – inferenca) ter zagon dolgoročnih storitev (npr. API

za inferenco). V praksi se zato vse pogosteje kombinira klasični HPC pristop s kontejnerskimi in »cloud-native« pristopi.

Kubernetes je odprtokodni sistem za avtomatizacijo nameščanja, skaliranja in upravljanja »kontejneriziranih« aplikacij in je zaradi svoje zasnove primeren za izvajanje dolgoročnih storitev in platformnih komponent (npr. orodja MLOps). (Kubernetes, n.d.), vendar ga trenutni superračunalniki v Sloveniji ne podpirajo.

7 Perspektiva projekta Slovenske tovarne umetne inteligence (SLAIF)

Marca 2025 je EuroHPC JU objavil izbor dodatnih evropskih tovarn UI, med katerimi je tudi slovenska iniciativa SLAIF. Po objavi bo novi superračunalniški sistem razvijal in upravljal IZUM v sodelovanju z IJS in ARNES, nameščen pa bo v novem podatkovnem centru, ki ga ARNES gradi v bližini hidroelektrarne Mariborski otok. (EuroHPC JU, 2025)

Javni opis SLAIF navaja, da bo novi superračunalnik bistveno zmogljivejši od Vege: okvirno 100 PFLOPS HPC zmogljivosti (okoli $16\times$ Vega) ter za potrebe umetne inteligence okoli 10 exaFLOPS v mešani natančnosti, približno 2000 grafičnih pospeševalnikov, 10 PB hitrega diskovnega sistema za podatkovno intenzivne aplikacije in 100 PB shrambe na tehnologiji Ceph. (IZUM, 2025)

Pomemben konceptualni premik v primerjavi s tradicionalnim HPC je načrtovana delitev na več logičnih particij, vključno s 'cloud services partition', ki je namenjena poganjanju aplikacij UI (inferenca) in drugih dolgoročnih uporabniških storitev. (IZUM, 2025)

V delu, ki se nanaša na sistemsko programsko opremo, je izrecno zapisano, da bodo številni uporabniki potrebovali orkestracijo kontejnerjev, kot je Kubernetes, za skalabilne in upravljive namestitve aplikacij; Kubernetes je tako predviden kot pomemben gradnik pri povezovanju HPC in oblačnih storitev. (IZUM, 2025)

Glede časovnice viri navajajo leto 2027 kot ciljno obdobje prehoda: uradna stran SLAIF navaja, da naj bi bil nov UI-optimiziran sistem v produkciji v drugi polovici leta 2027, medtem ko nekatere partnerske objave omenjajo začetek 2027. V vsakem primeru se Vega predvideva kot ključni premostitveni vir do uvedbe novega sistema. (SLAIF, n.d.; Univerza v Ljubljani, 2025)

8 Zaključek

Slovenska infrastruktura SLING danes omogoča izvajanje nalog računalniškega vida in umetne inteligence na dveh komplementarnih okoljih: Vega nudi petaflopno zmogljivost z A100 GPU in bogatim naborom knjižnic za UI, Arnesova gruča pa predstavlja dostopnejše (tudi izobraževalno) okolje s heterogenimi GPU, vključno z novjšimi H100. Razpoložljivost je določena s kvotami in razpisnimi mehanizmi, zato sta pri večjih eksperimentih ključno načrtovanje in iterativni pristop od TA/DA do RRA/LRA.

SLAIF prinaša jasno usmeritev v UI-optimizirano infrastrukturo, kjer bo poleg klasične HPC zasnove poudarek tudi na storitvah, ki so ključne za UI (dostop prek API, orkestracija kontejnerjev, dolgoročne storitve). Če bodo načrtovane funkcionalnosti (npr. Kubernetes in cloud particija) uresničene, se bo uporabniški model približal sodobnim UI platformam, hkrati pa ohranil prednosti superračunalniškega okolja.

Opomba

Vse navedene tehnične specifikacije in kvote so povzete iz javno objavljenih virov, citiranih v poglavju "Viri in literatura". Ker se sistemi posodablajo, je pri praktični uporabi priporočljivo preveriti zadnjo različico dokumentacije. Prispevek temelji na javno dostopni dokumentaciji in objavah organizacij IZUM, SLING, EuroHPC JU ter na javno dostopnih informacijah projekta SLAIF. Pri virih smo dajali prednost tistim v angleškem jeziku.

Izjava o uporabi orodij umetne inteligence

Ta prispevek je bil pripravljen s pomočjo orodij umetne inteligence, razvitih pri podjetju OpenAI, in sicer ChatGPT (vključno z zmogljivostjo Deep Research ter modelom GPT-5.2 Pro). Za točnost, celovitost in izvirnost prispevka v celoti odgovarja avtor.

Zahvala

Članek je nastal v okviru projekta Slovenska tovarna umetne inteligence (SLAIF – Slovenian AI Factory). Zahvaljujemo se za podporo Evropske komisije / EuroHPC JU ter Ministrstva za visoko šolstvo, znanost in inovacije Republike Slovenije (MVZI) v okviru projekta SLAIF (številka pogodbe o sofinanciranju 101254461).

Literatura

CNCF. (2024, 8 July). Slurm: An HPC workload manager. Cloud Native Computing Foundation. Dostopno na: <https://www.cncf.io/blog/2024/07/08/slurm-an-hpc-workload-manager/> (pridobljeno 14. 2. 2026).

- EuroCC ACCESS. (n.d.). NCC Slovenia. EuroCC ACCESS. Dostopno na: <https://www.eurocc-access.eu/about-us/meet-the-nccs/ncc-slovenia/> (pridobljeno 14. 2. 2026).
- EuroHPC Joint Undertaking. (2025, 12 March). The EuroHPC JU selects additional AI Factories to strengthen Europe's AI leadership. EuroHPC JU. Dostopno na: https://www.eurohpc-ju.europa.eu/eurohpc-ju-selects-additional-ai-factories-strengthen-europes-ai-leadership-2025-03-12_en (pridobljeno 14. 2. 2026).
- HPC Vega documentation. (n.d.-a). Architecture. HPC Vega – IZUM, Maribor, Slovenia. Dostopno na: <https://doc.vega.izum.si/architecture> (pridobljeno 14. 2. 2026).
- HPC Vega documentation. (n.d.-b). Vsebniki (Singularity). HPC Vega – IZUM, Maribor, Slovenia. Dostopno na: <https://si-doc.vega.izum.si/vsebniki/> (pridobljeno 14. 2. 2026).
- HPC Vega documentation. (n.d.-c). Vega Open OnDemand. HPC Vega – IZUM, Maribor, Slovenia. Dostopno na: <https://doc.vega.izum.si/ondemand/> (pridobljeno 14. 2. 2026).
- Hyperion Research. (2023, June). Slurm Remains Top Resource Manager (Special Study Results). Hyperion Research. Dostopno na: <https://hyperionresearch.com/wp-content/uploads/2023/06/Hyperion-Research-MCS-Highlights-Slurm-June-2023.pdf> (pridobljeno 14. 2. 2026).
- IJS Indico. (2025a, 17–18 June). Workshop: EuroCC2 Multi-GPU Programming Bootcamp. IJS Indico. Dostopno na: <https://indico.ijs.si/event/2914/> (pridobljeno 14. 2. 2026).
- IJS Indico. (2025b, 10 July). Bootcamp: Profiling AI Software. IJS Indico. Dostopno na: <https://indico.ijs.si/event/2969/> (pridobljeno 14. 2. 2026).
- Institute of Information Science (IZUM). (2025). Slovenian Artificial Intelligence Factory – SLAIF. IZUM. Dostopno na: https://izum.si/en/ai_factory/ (pridobljeno 14. 2. 2026).
- Institute of Information Science (IZUM). (n.d.-a). Vega. IZUM. Dostopno na: <https://izum.si/en/vega-en/> (pridobljeno 14. 2. 2026).
- Kubernetes. (n.d.). Kubernetes Documentation. Kubernetes. Dostopno na: <https://kubernetes.io/docs/home/> (pridobljeno 14. 2. 2026).
- SchedMD. (n.d.). Overview: Slurm Workload Manager. Slurm Workload Manager Documentation. Dostopno na: <https://slurm.schedmd.com/overview.html> (pridobljeno 14. 2. 2026).
- SLAIF. (n.d.). SLAIF – Slovenian AI Factory (Frequently Asked Questions). SLAIF. Dostopno na: <https://www.slaif.si/> (pridobljeno 14. 2. 2026).
- SLING. (n.d.-a). Arnes HPC cluster. SLING. Dostopno na: <https://www.sling.si/en/arnes-hpc-cluster/> (pridobljeno 14. 2. 2026).
- SLING. (n.d.-b). National calls for open access and use of resources. SLING. Dostopno na: <https://www.sling.si/en/calls-for-open-access-and-use-of-resources/> (pridobljeno 14. 2. 2026).
- SLING. (n.d.-c). National Competence Centre. SLING. Dostopno na: <https://www.sling.si/en/competence-centre/national-competence-centre/> (pridobljeno 14. 2. 2026).
- Univerza v Ljubljani. (2025, 14 March). The University of Ljubljana will play a key role in establishing the Slovenian Artificial Intelligence Factory. Univerza v Ljubljani. Dostopno na: <https://www.uni-lj.si/en/news/2025-03-14-the-university-of-ljubljana-will-play-a-key-role-in-establishing-the-slovenian-artificial-intelligence-factory> (pridobljeno 14. 2. 2026).
- Univerza v Ljubljani FMF. (n.d.). EUROCC2 - National Competence Centres in the framework of EuroHPC. University of Ljubljana, Faculty of Mathematics and Physics. Dostopno na: <https://www.fmf.uni-lj.si/sl/raziskave/mednarodni/eurocc2/> (pridobljeno 14. 2. 2026).

