



NCC CROATIA – HRVATSKI CENTAR KOMPETENCIJA ZA  
RAČUNARSTVO VISOKIH PERFORMANSI (HPC)

UPITNIK O OBRAZOVNIM POTREBAMA IZ PODRUČJA  
RAČUNARSTVA VISOKIH PERFORMANSI

– ANALIZA REZULTATA –

(5.-15.3.2021.)

EuroCC – National Competence Centres in the framework of EuroHPC

Ožujak 2021.



**EuroHPC**  
Joint Undertaking

“Projekt EuroCC financiran je sredstvima Zajedničkog poduzeća za europsko računarstvo visokih performansi (EuroHPC JU) pod brojem Ugovora 951732. EuroHPC JU prima financijsku potporu iz EU programa Obzor 2020 i Njemačke, Bugarske, Austrije, Hrvatske, Cipra, Češke Republike, Danske, Estonije, Finske, Grčke, Mađarske, Irske, Italije, Litve, Latvije, Poljske, Portugala, Rumunjske, Slovenije, Španjolske, Švedske, Ujedinjenog Kraljevstva, Francuske, Nizozemske, Belgije, Luksemburga, Slovačke, Norveške, Švicarske, Turske, Republike Sjeverne Makedonije, Islanda i Crne Gore.”

## Sadržaj

1. Uvod .....	5
2. Struktura upitnika i metoda provedbe .....	5
2.1 Pregled pitanja .....	6
2.1.1 Grupa pitanja „O vama“ .....	6
2.1.2 Grupa pitanja „HPC resursi“ .....	6
2.1.3 Grupa pitanja „HPC softver“ .....	7
2.1.4 Grupa pitanja „Obrazovni programi“ .....	8
2.1.5 Grupa pitanja „Završni komentar“ .....	9
3. Detaljni rezultati.....	9
3.1 Rezultati grupe pitanja „O vama“ .....	9
3.2 Rezultati grupe pitanja „HPC resursi“ .....	12
3.3 Rezultati grupe pitanja „HPC softver“ .....	18
3.4 Rezultati grupe pitanja „Obrazovni program“ .....	23
3.5 Rezultati grupe pitanja „Završni komentar“ .....	30
4. Zaključak .....	31
5. Reference .....	33

## Popis tablica:

Tablica 1. Broj zaprimljenih odgovora po ustanovi .....	9
Tablica 2. Broj zaprimljenih odgovora po znanstvenom području .....	11
Tablica 3. Broj zaprimljenih odgovora po poziciji ispitanika .....	11
Tablica 4. Broj CPU i/ili GPU koji ispitanici najčešće koriste .....	12
Tablica 5. Broj CPU i/ili GPU sati koje ispitanik treba u periodu od godine dana .....	14
Tablica 6. Drugi resursi koje ispitanici koriste u značajnoj količini .....	15
Tablica 7. Prepreke u korištenju 10 puta većih resursa .....	17
Tablica 8. Susretanje s problemom u istraživačkom radu za čije je rješavanje potreban HPC resurs veći od 1 PFLOPS .....	18
Tablica 9. Način korištenja HPC softvera .....	18
Tablica 10. Besplatni HPC softveri koje ispitanici trenutno koriste .....	19
Tablica 11. Besplatni HPC softveri koje ispitanici planiraju koristiti .....	20
Tablica 12. Komercijalni HPC softveri koje ispitanici trenutačno koriste .....	21
Tablica 13. Komercijalni HPC softveri koje bi ispitanici koristili da su im dostupni .....	21
Tablica 14. Mogućnost izvođenja vlastitog razvijenog softvera na heterogenim računalnim arhitekturama .....	22
Tablica 15. Korištenje HPDA softvera .....	23
Tablica 16. Korištenje softvera za umjetnu inteligenciju .....	23
Tablica 17. Tema obrazovnog programa .....	24
Tablica 18. Vrsta izvođenja obrazovnog programa .....	26
Tablica 19. Trajanje obrazovnog programa .....	28
Tablica 20. Način izvođenja obrazovnog programa .....	29
Tablica 21. Pregled završnih komentara .....	30

## Popis slika:

Slika 1. Korištenje i pristup HPC resursi.....	25
Slika 2. Specifične vrste simulacija/modeliranja.....	25
Slika 3. Specifični HPC softver .....	25
Slika 4. Programski jezici.....	25
Slika 5. Moderni programski jezici visokih performansi (npr. Julia ili Python) .....	25
Slika 6. MPI (Message Passing Interface) .....	25
Slika 7. OpenMP (Open Multi-Processing) .....	25
Slika 8. CUDA (Compute Unified Device Architecture).....	25
Slika 9. OpenACC (Open Accelerators) .....	26
Slika 10. Umjetna inteligencija .....	26
Slika 11. HPDA (High performance Data Analytics).....	26
Slika 12. Izvedbeni inženjering (engl. performance engineering).....	26
Slika 13. Prezentacije sa slajdovima.....	27
Slika 14. Demonstracije uživo .....	27
Slika 15. Vođene vježbe/hands-on sesije .....	27
Slika 16. Grupni rad i rasprave .....	27
Slika 17. Nadgledani projektni rad .....	27
Slika 18. Samostalno učenje .....	27
Slika 19. Pregled odgovora – način izvođenja obrazovnog programa .....	29

## 1. Uvod

Jedna od glavnih aktivnosti projekta EuroCC je utvrđivanje kompetencija i obrazovnih potreba iz područja računarstva visokih performansi (engl. *High Performance Computing* HPC), analize podataka računalnim okolinama i postupcima visokih performansi (engl. *High Performance Data Analytics* HPDA) i umjetne inteligencije (engl. *Artificial Intelligence* AI) u Hrvatskoj. Istraživanje obrazovnih potreba će biti provedeno u nekoliko faza i bit će distribuirano istraživačima i inženjerima u akademiji i industriji u Republici Hrvatskoj.

U sklopu tog istraživanja Hrvatski centar kompetencija za računarstvo visokih performansi proveo je upitnik o obrazovnim potrebama iz područja računarstva visokih performansi u razdoblju od 5. do 15. ožujka 2021. Ovom fazom provedbe, upitnikom su obuhvaćeni Znanstveni centri izvrsnosti te postojeći korisnici HPC infrastrukture iz akademske zajednice RH.

Slične ankete o obrazovnih potrebama Hrvatski centar kompetencija će provoditi svake godine kako bi pratio daljnji razvoj obrazovnih potreba te kako bi se evaluirali rezultati trenutno raspoloživih obrazovnih programa.

## 2. Struktura upitnika i metoda provedbe

Glavni cilj upitnika je razumijevanje obrazovnih potreba iz područja HPC/HPDA/AI. Radi boljeg razumijevanja tih potreba dodana su pitanja koja se odnose na iskustvo ispitanika u HPC/HPDA/AI okolini.

Istraživanje je provedeno korištenjem alata LimeSurvey [1]. Period provođenja upitnika je bio od 5. ožujka do 15. ožujka 2021.

Definirane su dvije ciljane skupine prema kojima je upućen direktan mail od strane predstavnika projekta EuroCC, kako bi se osigurao najbolji odaziv:

- Znanstveni centri izvrsnosti [2] – 20 kontakata raspoloživih na web stranici,
- Korisnici Računalnog klastera Isabelle [3] – 271 kontakt (vlastiti/fizički kontakti Srca).

Sveukupno je zaprimljeno 112 odgovora od čega 47 potpunih. Najviše zaprimljenih odgovora je bilo u prvom danu provođenja upitnika, a zatim 11. 3. (tjedan dana kasnije) kada je poslan podsjetnik ispitanicima.

Analiza rezultata upitnika se odnosi samo na potpune odgovore.

## 2.1 Pregled pitanja

Upitnik sadrži 22 pitanja raspoređenih u 5 grupa pitanja: O vama, HPC resursi, HPC softver, Obrazovni programi i Završni komentar. Pitanja označena s \* su obavezna.

### 2.1.1 Grupa pitanja „O vama“

1. Odaberite ustanovu s popisa znanstvenih organizacija. Ako ustanova nije na popisu upišite ime odabirom opcije "Ostalo:".

(ponuđena je lista od 175 ustanova)

2. \*Odaberite znanstveno područje ili područje istraživanja.

(ponuđena je lista od 18 znanstvenih područja)

3. \*Navedite trenutnu poziciju.

(moguće odabrati više odgovora)

- a. Student
- b. Doktorand
- c. Postdoktorand
- d. Nastavno osoblje
- e. Znanstvenik/istraživač
- f. Inženjer/programer istraživačkog softvera
- g. Podatkovni istraživač
- h. Istraživač iz područja umjetne inteligencije
- i. Ostalo

4. \*Imate li iskustva u uporabi HPC tehnologija (npr. pristup i uporaba resursa i softvera, programiranje, ...)?

(Da/Ne)

### 2.1.2 Grupa pitanja „HPC resursi“

5. Koliko procesorskih jezgri (CPU) i/ili grafičkih procesora (GPU) najčešće koristite? (slobodan unos broja jezgri/procesora)

- a. Uobičajeni broj procesorskih jezgri
- b. Maksimalni broj procesorskih jezgri

- c. Uobičajeni broj grafičkih procesora
  - d. Maksimalni broj grafičkih procesora
6. Koliko CPU odnosno GPU sati trebate u periodu od godine dana?  
(slobodan unos broja sati)
- a. CPU sati
  - b. GPU sati
7. Navedite druge HPC resurse (npr. radna memorija, brzo spremište, mreža), ako ih koristite u značajnoj količini.  
(slobodan unos teksta)
8. Ako biste dobili pristup 10 puta većim resursima od onih koje trenutno koristite, koje bi bile glavne prepreke u uporabi te mogućnosti?  
(mogućnost odabira više odgovora)
- a. Nedostatak HPC softvera
  - b. Nedostatak vještina programiranja
  - c. Skaliranje problema
  - d. Nema prepreka
  - e. Ostalo
9. \*Jeste li se susreli s problemom u istraživačkom radu za čije učinkovito rješavanje je nužan pristup računalnim resursima većim od 1 PFLOPS?  
(Da/Ne/Ne znam)

### 2.1.3 Grupa pitanja „HPC softver“

10. \*Na koji način najčešće koristite HPC softver?
- a. Koristim postojeći HPC softver
  - b. Razvijam vlastiti HPC softver
  - c. Oboje od prethodno navedenog
  - d. Ništa od prethodno navedenog
  - e. Upišite dodatni komentar ako smatrate potrebnim
11. Navedite besplatne HPC softvere koje trenutačno koristite.
12. Navedite besplatne HPC softvere koje planirate koristiti.

13. Navedite komercijalne HPC softvere koje trenutno koristite.
14. Navedite komercijalne HPC softvere koje biste koristili da su vam dostupni.
15. Ako razvijate vlastiti HPC softver, može li se on izvoditi na heterogenim računalnim arhitekturama? (npr. na procesorskim jezgrama i grafičkim procesorima)
16. \*Koristite li High Performance Data Analytics (HPDA) softver?  
(Da/Ne/Ne znam)
17. \*Koristite li softver za umjetnu inteligenciju (neuronske mreže, dubinsko učenje, itd.)?  
(Da/Ne/Ne znam)

#### 2.1.4 Grupa pitanja „Obrazovni programi“

18. \*Kakvu temu ste zainteresirani pohađati te na kojoj razini? (Uvodnoj, Srednjoj, Naprednoj, Ne zanima me, Ne znam što je to)  
(za svaku ponuđenu temu potreban je odabir razine)
- a. Korištenje i pristup HPC resursima
  - b. Specifične vrste simulacija/modeliranja
  - c. Specifični HPC softver
  - d. Programski jezici
  - e. Moderni programski jezici visokih performansi (npr. Julia ili Python)
  - f. MPI (Message Passing Interface)
  - g. OpenMP (Open Multi-Processing)
  - h. CUDA (Compute Unified Device Architecture)
  - i. OpenACC (Open Accelerators)
  - j. Umjetna inteligencija
  - k. HPDA (High performance Data Analytics)
  - l. Izvedbeni inženjering (engl. performance engineering)
19. \*Ocijenite koliko korisnim smatrate ove vrste obrazovnih programa (Nedovoljno, Dovoljno, Dobro, Vrlo dobro, Odlično)  
(za svaki ponuđeni obrazovni program potreban je odabir razine)
- a. Prezentacije sa slajdovima



- b. Demonstracije uživo
- c. Vođene vježbe/hands-on sesije
- d. Grupni rad i rasprave
- e. Nadgledani projektni rad
- f. Samostalno učenje

20. Koliko dana bi mogli izdvojiti za ove obrazovne programe?

(slobodan unos)

21. \*Preferirate li obrazovne programe uživo ili online?

- a. Uživo
- b. Online
- c. Oboje od prethodno navedenog
- d. Ne znam
- e. Nisam zainteresiran

2.1.5 Grupa pitanja „Završni komentar“

22. Vaši komentari za kraj

### 3. Detaljni rezultati

#### 3.1 Rezultati grupe pitanja „O vama“

Najviše ispitanika koji su potpuno odgovorili na upitnik je iz ustanove Institut „Ruđer Bošković“. Takav rezultat je i očekivano, budući da zaposlenici ove ustanove godinama nose najveći udio u korištenju računalnog klastera Isabella za potrebe istraživanja. Detaljan pregled odgovora dan je u Tablici 1.

*Tablica 1. Broj zaprimljenih odgovora po ustanovi*

Ustanova	Broj zaprimljenih odgovora	%
Institut "Ruđer Bošković", Zagreb	20	42,55%
Farmaceutsko-biokemijski fakultet, Zagreb	4	8,51%
Fakultet kemijskog inženjerstva i tehnologije, Zagreb	3	6,38%

Fakultet elektrotehnike i računarstva, Zagreb	2	4,26%
Medicinski fakultet, Osijek	2	4,26%
Sveučilište u Osijeku – Odjel za fiziku	2	4,26%
Agronomski fakultet, Zagreb	1	2,13%
Ekonomski fakultet, Zagreb	1	2,13%
Fakultet elektrotehnike, računarstva i informacijskih tehnologija Osijek	1	2,13%
Filozofski fakultet, Zagreb	1	2,13%
Geotehnički fakultet, Varaždin	1	2,13%
Institut za fiziku, Zagreb	1	2,13%
Institut za medicinska istraživanja i medicinu rada, Zagreb	1	2,13%
Other	1	2,13%
Prirodoslovno–matematički fakultet, Split	1	2,13%
Prirodoslovno–matematički fakultet, Zagreb	1	2,13%
Sveučilište u Splitu	1	2,13%
Sveučilište u Zagrebu	1	2,13%
Tehnički fakultet, Rijeka	1	2,13%
Učiteljski fakultet, Zagreb	1	2,13%
Ukupno	47	100%

Znanstveno područje prirodnih znanosti, i to znanstvena polja Fizika i Kemija, dominantno koriste HPC resurse. Tablica 2 prikazuje udio ostalih znanstvenih područja i polja u korištenju HPC resursa.

Tablica 2. Broj zaprimljenih odgovora po znanstvenom području

Znanstveno područje, znanstveno polje	Broj zaprimljenih odgovora	%
PRIRODNE ZNANOSTI , Kemija	16	34,04%
PRIRODNE ZNANOSTI , Fizika	11	23,40%
TEHNIČKE ZNANOSTI , Računarstvo	5	10,64%
BIOMEDICINA I ZDRAVSTVO , Temeljne medicinske znanosti	4	8,51%
BIOTEHNIČKE ZNANOSTI , Interdisciplinarne biotehničke znanosti	4	8,51%
PRIRODNE ZNANOSTI , Biologija	2	4,26%
DRUŠTVENE ZNANOSTI	1	2,13%
HUMANISTIČKE ZNANOSTI	1	2,13%
PRIRODNE ZNANOSTI , Interdisciplinarne prirodne znanosti	1	2,13%
PRIRODNE ZNANOSTI , Matematika	1	2,13%
TEHNIČKE ZNANOSTI , Temeljne tehničke znanosti	1	2,13%

Najveći broj ispitanika radi kao znanstvenik/istraživač što je korektno no čini se previše generički. Ovo pitanje ima prostora za poboljšanje u segmentu većeg izbora radnog mjesta/pozicije. Detaljan pregled odgovora dan je u Tablici 3.

Tablica 3. Broj zaprimljenih odgovora po poziciji ispitanika

Pozicija ispitanika	Broj zaprimljenih odgovora	%
Znanstvenik/istraživač	33	70,21%
Nastavno osoblje	16	34,04%
Doktorand	12	25,53%
Istraživač iz područja umjetne inteligencije	5	10,64%

Postdoktorand	4	8,51%
Inženjer/programer istraživačkog softvera	2	4,26%
Podatkovni istraživač	2	4,26%
Student	0	0,00%

Iskustva u uporabi HPC tehnologija ima 43 ispitanika (91,49%) što je očekivani rezultat jer je upitnik poslan korisnicima HPC resursa Isabella.

### 3.2 Rezultati grupe pitanja „HPC resursi“

Ispitanici vrlo dobro znaju svoje potrebe za brojem procesorskih jezgri (CPU) i/ili grafičkih procesora (GPU) koje najčešće koriste. Većina ispitanika vrlo precizno definira svoje uobičajene i maksimalne CPU potrebe. Kod GPU potreba je značajno manji broj ispitanika koji koriste GPU (Tablica 4). Ovi ispitanici također znaju procijeniti i svoje potrebe za brojem CPU i/ili GPU sati u periodu od godine dana (Tablica 5).

Tablica 4. Broj CPU i/ili GPU koji ispitanici najčešće koriste

Uobičajeni broj procesorskih jezgri	Maksimalni broj procesorskih jezgri	Uobičajeni broj grafičkih procesora	Maksimalni broj grafičkih procesora
8-28	do sada 28	nisam koristio	nisam koristio
8, 16, 28	28200	/	/
500-1000	10000	-	-
100	1000	100	200
100	512	16	16
100	512	10	12
64	256	4	10
62	200	4	10
32	200	3	8
28	160	2	4

28	128	2	4
28	100	2	3
28	100	1	3
28	64	1	3
28	48	1	3
20	40	1	2
16	32	1	2
16	32	1	2
16	32	1	2
16	32	1	2
14	32	1	1
14	28	1	1
12	28	1	1
10	28	1	1
8	28	1	1
8	28	1	1
8	28	0	1
8	28	0	1
8	28	0	1
8	28	0	0
8	28	0	0
8	28	0	0
8	24	0	0

4	20	0	0
4	20	0	0
4	16	0	0
4	16	0	0
4	16	0	0
4	16	0	0
4	16	0	0
4	12	0	0
4	10	0	0
4	8	0	0
1	4		
1	4		

Tablica 5. Broj CPU i/ili GPU sati koje ispitanik treba u periodu od godine dana

CPU				
varijabilno, prema potrebi	500000-700000	500000	3000	750
puno, koristit ću koliko mi se dopusti	500 000	500000	3000	720
Ne mogu procijeniti	>1000000	500000	1400	600
cijelu godinu	7E+11	500000	1200	500
	6000000	100000	1000	500
	1000000	100000	1000	200
		10000	1000	200
		10000	1000	150
		8765	1000	100

		5000	1000	
		4000		
GPU				
puno, koristi ću koliko mi se dopusti	1000000	5000	720	0
Ne mogu procijeniti	200000	2000	500	0
200 000	150000	2000	200	0
~20000	100000	2000	200	0
/	100000	1000	100	0
-		1000	24	0
		1000	12	0
				0
				0
				0
				0

Pregled drugih HPC resursa (npr. radna memorija, brzo spremište, mreža) koje ispitanici koriste u značajnoj količini prikazan je u Tablici 6.

Tablica 6. Drugi resursi koje ispitanici koriste u značajnoj količini

Drugi resursi koje ispitanici koriste u značajnoj količini
Mreža visoke propusnosti (100Gb/s) i malog kašnjenja, lokalni NVMe diskovi, učinkovito središnje spremište, cold spremište za čuvanje podataka tijekom trajanja projekta.
radna memorija, spremište
U načelu, poslovi koje zadajem u programskom paketu Gaussian zahtijevaju oko 1-2 GB radne memorije po procesoru, te scratch od 0,5-1 GB po procesoru. Mrežni zahtjevi nisu značajni, obzirom kako koristim Gaussian najčešće na mpisingle redu poslova. "Idealni" poslužitelj koje zamišljam za svoje potrebe bio bi opremljen s dva Xeon Gold 5220R procesora (48 jezgri), 128 GB radne memorije i 2x1 TB SSD u RAID 0.

128GB, a volio bih 256–512GB koristiti
64GB RAM,300GB spremište
80GB RAM per node, 10 TB data storage
Radna memorija: 254 GB, Pohrana podataka: 5 TB
Prosječno koristim između 4 i 5 GB radne memorije po jezgri.
1TB
min 4gb RAM/core/cpu, min 4tb HDD/gpu
Radna memorija – u prosjeku iznad 100 GB
Brzo spremište i mreža.
radna memorija
radna memorija, 1–2 Gb po jezgri
tipičan "veći"posao koji sam koristio je tražio 60.05 GB dinamičkog rama (2.14 GB po procesoru). Više je potrebno za eventualne zahtjevnije poslove.
Glavninu HPC sati trošim na DFT (density functional theory) račune koji često zahtjevaju i 8GM RAM po jezgri. Sustavi koji se proučavaju su takvi da često ne stanu na jedan čvor, a problem nije trivijalno paralelizirati pa brza mreža igra ulogu.
16GB RAM, 50GB scratch
Memorija uvijek dobro dođe :-)
Spremište, treba mi daleko više, u prosjeku mi po radu treba 1 Tb, uvijek istovremeno radim na 3–4 rada....
Rad s trajektorijama koje zahtijevaju dosta skladišnog prostora, red veličine do 100TB.
radna memorija do 50 GB, brzo spremište 200 GB
Brzo spremište – nekoliko stotina TB
Koristim sve HPC resurse u značajnoj količini.



radna memorija, brzo spremište
do 96 GB RAM-a, do 3 TB privremenog diskovnog prostora (dok traje račun), do 500 GB prostora na "home" disku
96 Gb RAM, 2 Tb privremenog prostora na HD, 200 Gb na home
Ništa posebno.
radna memorija, mreža

Ako bi ispitanici dobili pristup 10 puta većim resursima od onih koje trenutno koriste, najveća prepreka u uporabi je nedostatak vještina programiranja. Detaljan pregled odgovora dan je u Tablici 7.

Tablica 7. Prepreke u korištenju 10 puta većih resursa

Prepreke u korištenju 10 puta većih resursa	Broj zaprimljenih odgovora	%
Nema prepreka	28	50,00%
Nedostatak vještina programiranja	13	23,21%
Skaliranje problema	6	10,71%
Nedostatak HPC softvera	5	8,93%
Ostalo: <ul style="list-style-type: none"> <li>• moje neznanje kak to dobro koristiti</li> <li>• Pristup samo iz CARNET mreže</li> <li>• Zagušenost HPC resursa tj. dugoročno čekanje na pokretanje obrade.</li> </ul>	4	7,14%

S problemom u istraživačkom radu za čije učinkovito rješavanje je nužan pristup računalnim resursima većim od 1 PFLOPS susrelo se je 23,4% ispitanika. Detaljan pregled odgovora dan je u Tablici 8.

Tablica 8. Susretanje s problemom u istraživačkom radu za čije je rješavanje potreban HPC resurs veći od 1 PFLOPS

	Broj zaprimljenih odgovora	%
Susretanje s problemom u istraživačkom radu za čije učinkovito rješavanje je nužan pristup računalnim resursima većim od 1 PFLOPS		
Da	11	23,40%
Ne	18	38,29%
Ne znam	18	38,29%

### 3.3 Rezultati grupe pitanja „HPC softver“

Ispitanici najčešće koriste postojeći HPC softver. Detaljan pregled odgovora dan je u Tablici 9.

Tablica 9. Način korištenja HPC softvera

	Broj zaprimljenih odgovora	%
Način korištenja HPC softvera		
Koristim postojeći HPC softver	31	65/96%
Razvijam vlastiti HPC softver	4	8,51%
Oboje od prethodno navedenog	7	14,89%
Ništa od prethodno navedenog Ostalo:		10,64%
<ul style="list-style-type: none"> <li>• Razvoj vlastitog softverskog workflowa koji koristi postojeći software optimiziran za HPC upotrebu.</li> <li>• Bilo bi praktično kad bi se s klastera Isabella moglo direktno pristupati github-u.</li> <li>• Ja sam voditelj projekta a softver koristi kvalificirani suradnik</li> <li>• C, Fortran</li> <li>• Nisam baš vičan korištenju HPC pa uglavnom koristim lokalne PC, no volio bih imati pristup bržim računalima.</li> </ul>	5	

<p>Problem je i spora mreža na IRB-u: ja sad imam daleko brži internet kod kuće (100 Mbps) nego je prosječni throughput na IRB-u a podaci koje mam za obradu su ogromni te bi kod bržih računala vrijeme uploada podataka bilo jedbako ili dulje nego vrijeme obrade. Nadalje, softver kojeg koristim nije paraleliziran a brzina kloka na klasteru je manja nego na mom PCu (4.4 GHz) tako da mi je jedina korist pokretati više procesa istovremeno. Ne znam kako to načiniti na HPC bez da se iskilavim. Nadalje, moj strj počinje s obradom odmah (nemam queue). Uglavnm, milsim da HPC neće nikada biti prilagođen za neke primjene.</p>		
---	--	--

Pregled odgovora ispitanika o besplatnim HPC softverima koje trenutačno koriste dan je u Tablici 10.

Tablica 10. Besplatni HPC softveri koje ispitanici trenutno koriste

Besplatni HPC softveri koje ispitanici trenutno koriste	
GCC, Python	Bowtie2, Stacks, BayeScan, BayeScEnv, BayPass, Bayenv2
QE	Quantum Espresso, Abinit
GROMACS, NAMBD, AutoDock Vina, MMPBSA	pytorch, torchvision, cudnn, cuda
Canu 2.1.1 assembler, Guppy basecaller, Flye assembler	Amber20, Gromacs
Orca	R
cp2k, xtb	Iako to vjerojatno nije ono što pitate, od softvera otvorenog koda koristim LAPACK knjižnicu rutina za invertiranje matrica i traženje svojstvenih vrijednosti.
BEAST 1.10.4	Ne znam mogu li ih nazvati HPC softwareom, ali samo za prijenos podataka i kompajliranje vlastitih modela napisanih u fortranu/c/c++/python.
ORCA	Isabella

Abinit, QuantumEspresso	-
Bagel, CFour	C, Fortran
Orca, Quantum Espresso	GAMESS
Isabella cluster	GAMESS, VMD
trinity, blast, fastqc, trimmomatic, cd-hit, orthofinder, R...	Ništa
Quantum Espresso, CP2K, GPAWT	cuda wrapper
python, tensorflow, pytorch	OpenHPC

Pregled odgovora ispitanika o besplatnim HPC softverima koje planiraju koristiti dan je u Tablici 11.

*Tablica 11. Besplatni HPC softveri koje ispitanici planiraju koristiti*

Pregled odgovora – besplatni HPC softveri koje ispitanici planiraju koristiti	
Julia, R, CUDA	Isabella cluster
Q6 for molecular dynamics	python, tensorflow, pytorch
Gromacs	TensorFlow, Numpy
Orca, Gromacs	QuantumEspresso
XTB, GROMACS	Gromacs, Chimera, AmberTools, OpenBabel, Autodock Vina, MGL Tools, Rosetta, NAMD, HEX
cp2k	Q for molecular dynamics
BigDFT, Gromacs	GAMESS, ORCA
Siesta	GAMESS, VMD
isto kao pod pitanjem 11	Ništa

Gaussian i Amber su komercijalni HPC softveri koje ispitanici trenutno najviše koriste. Detaljan pregled odgovora dan je u Tablici 12.

Tablica 12. Komercijalni HPC softveri koje ispitanici trenutno koriste

Pregled odgovora – komercijalni HPC softveri koje ispitanici trenutno koriste	
Gaussian	/
Gaussian, Amber16	VASP, QuantumATK
Gaussian	Gaussian16
Amber, Gaussian	Gaussian
Gaussian	Gaussian09 i Gaussian16
Gaussian	Nema ih.
Gaussian, Amber	Gaussian, Amber
Gaussian	Gaussian16, Amber, 20 AIMall, NBO6.0,
Turbomole	Gaussian09, Gaussian16
Amber, Gaussian	Ništa
Gaussian 16, Matlab (HPC?)	

Pregled komercijalnih softvera koje bi ispitanici koristili da su im dostupni dan je u Tablici 13.

Tablica 13. Komercijalni HPC softveri koje bi ispitanici koristili da su im dostupni

Pregled odgovora – komercijalni HPC softveri koje bi ispitanici koristili da su im dostupni	
Amber	CLC Genomics Workbench
Terachem	QuantumATK
VASP	ADF
Turbomole, Terachem	Amber
Amber	–
Turbomole, Terachem	Mathematica
Amsterdam Modeling Suite (AMS)	MOLPRO

TeraChem	AIMAll
VASP	Ništa
QChem	Intel® Parallel Studio XE
Turbomole (on je zapravo dostupan, ali ne znam kome točno), Schrodingerov Jaguar (možda)	

Stav o izvođenju softvera na heterogenim računalnim arhitekturama unijelo je 14 ispitanika koji razvijaju vlastiti HPC softver. Rezultati su prikazani u Tablici 14.

Tablica 14. Mogućnost izvođenja vlastitog razvijenog softvera na heterogenim računalnim arhitekturama

Mogućnost izvođenja vlastitog razvijenog softvera na heterogenim računalnim arhitekturama
Može se izvoditi na heterogenim računalnim arhitekturama. Ne može se u potpunosti izvoditi samo na GPU-u ali dijelovi workflowa mogu koristiti GPU node-ove umjesto CPU.
Ne
ne razvijam vlastiti HPC softver
Ne može.
Ne
U principu da
Ne
Naš softver nema smisla koristiti na CPU-ovima a iskoristit će koliko god GPU-ova su dostupni.
Za sada ne, ali potencijalno da (na primjer, za računanje propagatora potrebno je invertiranje matrica, a to bi se sigurno moglo vrlo brzo izvršavati na grafičkim procesorima).
Da
DA
-
Da

-
Nešto razb'vijam ali nemam znanja pisati paralelne programe.
Da, na procesorskim jezgrama.

HPDA softver koristi svega 6,38% ispitanika. Detaljan pregled odgovora dan je u Tablici 15.

*Tablica 15. Korištenje HPDA softvera*

Korištenje HPDA softvera	Broj zaprimljenih odgovora	%
Da	3	6,38%
Ne	40	85,11%
Ne znam	4	8,51%

Softver za umjetnu inteligenciju (neuronske mreže, dubinsko učenje, itd.) koristi 25,53% ispitanika. Detaljan pregled odgovora dan je u Tablici 16.

*Tablica 16. Korištenje softvera za umjetnu inteligenciju*

Korištenje softvera za umjetnu inteligenciju	Broj zaprimljenih odgovora	%
Da	12	25,53%
Ne	34	72,34%
Ne znam	1	2,13%

### 3.4 Rezultati grupe pitanja „Obrazovni program“

Na ponuđenih 12 tema obrazovnog programa ispitanici su se izjasnili o interesu za pohađanjem ponuđenih tema s određenom razinom obrazovanja. Najveći interes ispitanici su pokazali za temu Korištenje HPC resursa i to na naprednoj razini. Specifične vrste simulacije su na drugom mjestu interesa također na naprednoj razini. Na trećem mjestu najinteresantnijih područja i to na uvodnoj razini su Moderni programski jezici i Umjetna

inteligencija. Istovremeno za teme Umjetna inteligencija i Izvedbeni inženjering 38,3% ispitanika nisu zainteresirani. Detaljniji pregled rezultata dan je u Tablici 17.

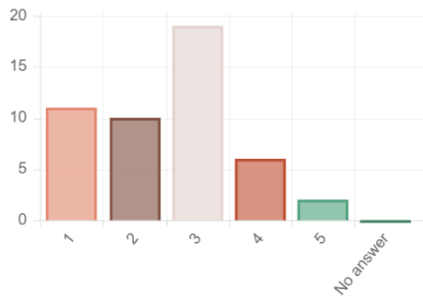
Tablica 17. Tema obrazovnog programa

Tema	Razina	Top %
Korištenje i pristup HPC resursima	Napredna	38,30
Specifične vrste simulacija/modeliranja	Napredna	34,04
Specifični HPC softver	Srednja	29,79
Programski jezici	Uvodna	29,79
Moderni programski jezici visokih performansi (npr. Julia ili Python) (Uvodna i Napredna razina imaju isti % odabira)	Uvodna, Napredna	31,91
MPI (Message Passing Interface)	Uvodna	23,40
OpenMP (Open Multi-Processing)	Uvodna	27,66
CUDA (Compute Unified Device Architecture)	Uvodna	25,53
OpenACC (Open Accelerators)	Ne znam što je to	31,91
Umjetna inteligencija *(drugi sljedeći postotak odabira 31,91% se odnosi i na polje Uvodna razina)	Ne zanima me	38,30
HPDA (High performance Data Analytics)	Ne zanima me	31,91
Izvedbeni inženjering (engl. performance engineering)	Ne zanima me	38,30

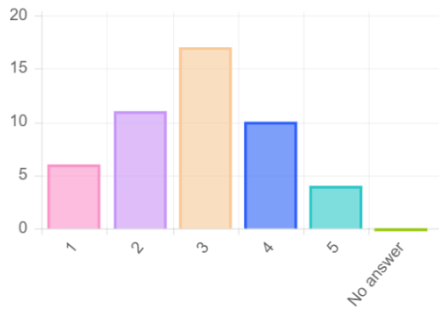
Grafički pregled distribucije odgovora ispitanika po svakoj temi prikazan je na grafovima dolje, pri čemu su šifre:

- 1 – Uvodna
- 2 – Srednja
- 3 – Napredna
- 4 – Ne zanima me
- 5 – Ne znam što je to.

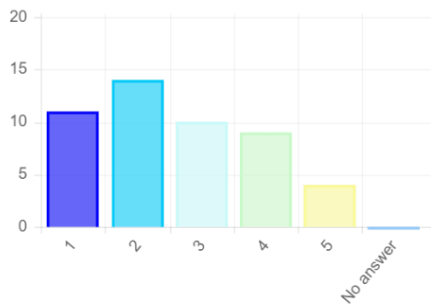




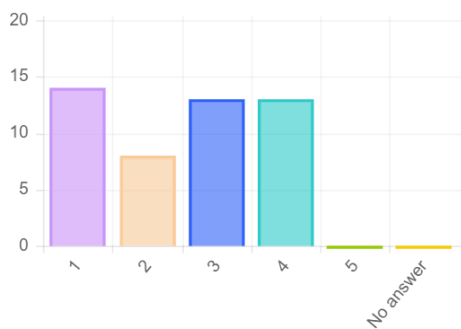
Slika 1. Korištenje i pristup HPC resursi



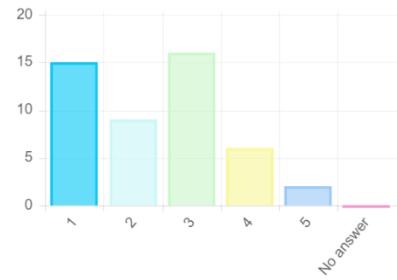
Slika 2. Specifične vrste simulacija/modeliranja



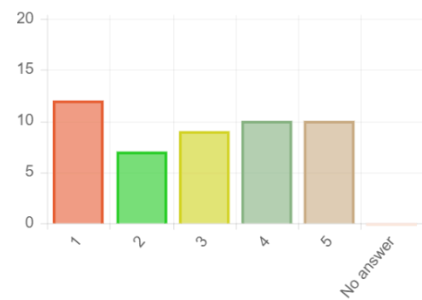
Slika 3. Specifični HPC softver



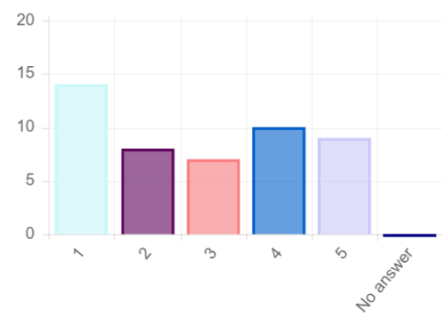
Slika 4. Programski jezici



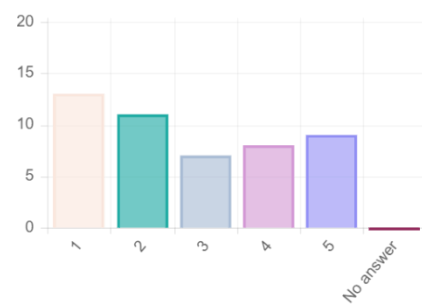
Slika 5. Moderni programski jezici visokih performansi (npr. Julia ili Python)



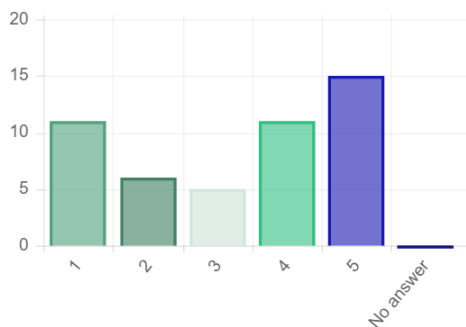
Slika 6. MPI (Message Passing Interface)



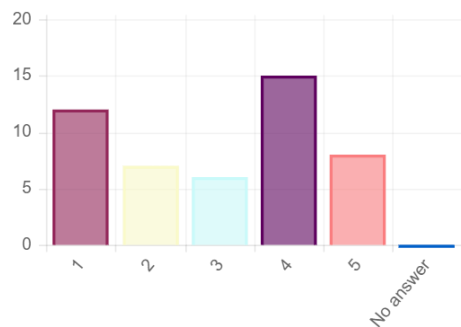
Slika 7. OpenMP (Open Multi-Processing)



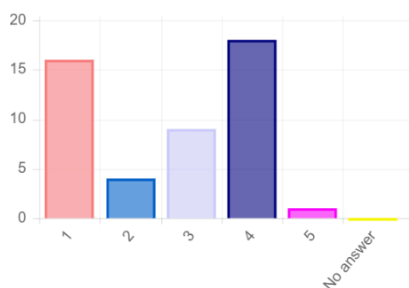
Slika 8. CUDA (Compute Unified Device Architecture)



Slika 9. OpenACC (Open Accelerators)



Slika 11. HPDA (High performance Data Analytics)



Slika 10. Umjetna inteligencija



Slika 12. Izvedbeni inženjering (eng. performance engineering)

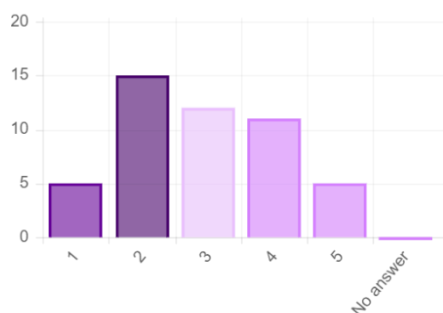
Na ponuđenih 6 vrsta izvođenja obrazovnog programa ispitanici su se izjasnili o interesu s određenom razinom koristi. Najkorisniji način izvođenja obrazovnog programa su vođene vježbe/hands-on sesije ocijenjene s odličnim od strane 57,45% ispitanika. Detaljniji pregled rezultata dan je u Tablici 18.

Tablica 18. Vrsta izvođenja obrazovnog programa

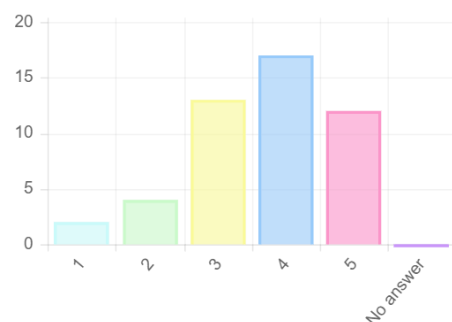
Vrsta izvođenja obrazovnog programa	Razina koristi	Top %
Prezentacija sa slajdovima	Dovoljno	31,91
Demonstracije uživo	Odlično	34,04
Vođene vježbe/hands-on sesije	Odlično	57,45
Grupni rad i rasprave	Vrlo dobro	36,17
Nadgledani projektni rad	Odlično	38,30
Samostalno učenje	Odlično	36,17

Grafički pregled distribucije odgovora ispitanika po vrsti izvođenja obrazovnog programa prikazan je na grafovima dolje, pri čemu su šifre:

- 1 – Nedovoljno
- 2 – Dovoljno
- 3 – Dobro
- 4 – Vrlo dobro
- 5 – Odlično.



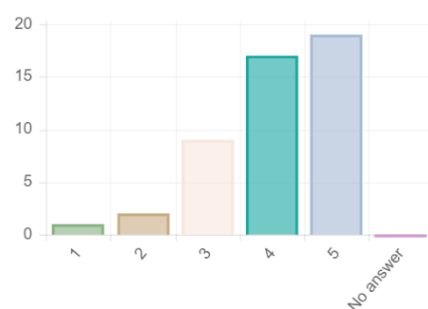
Slika 13. Prezentacije sa slajdovima



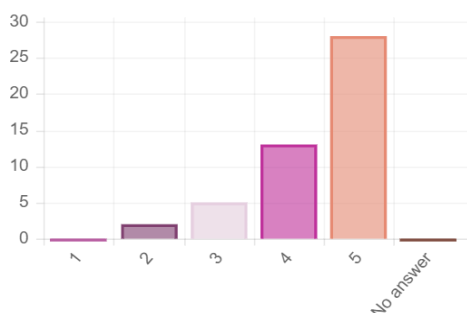
Slika 16. Grupni rad i rasprave



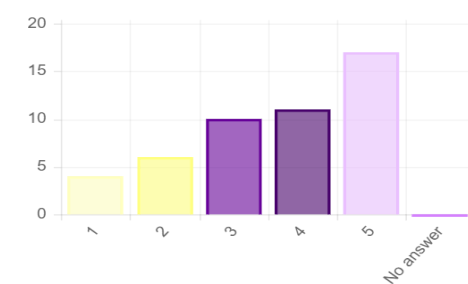
Slika 14. Demonstracije uživo



Slika 17. Nadgledani projektni rad



Slika 15. Vođene vježbe/hands-on sesije



Slika 18. Samostalno učenje

Većina ispitanika se je izjasnila za približno trajanje edukacije do tjedan dana. Detaljan pregled odgovora dan je u Tablici 19.

Tablica 19. Trajanje obrazovnog programa

Trajanje obrazovnog programa/dani	Broj ispitanika s istim odgovorom
1	0
2	1
3	2
5	3
7	5
10	7
14	1
20	2
21	1
30	3
1 tjedan po godini	1
10 godišnje.	1
10-tak	1
1-2	1
1-2 puta tjedno u trajanju od 45-90 min	1
2-3	1
3 - 5	1
5-7	1
do 5 dana po edukaciji	1
Koliko god je potrebno	1
Ne znam	1

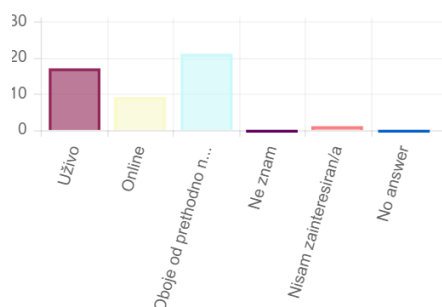
ovisi o čemu se radi i u kojem vremenskom intervalu, recimo 2–3 dana godišnje	1
Ovisi o važnosti. Za korištenje HPC resursa – 2 dana. Ostalo – samo izvan svog radnog vremena.	1
Ovisno o tome što pokriva obrazovni program i ovisno o dijelu godine (radi nastave). Voljan sam izdvojiti dana koliko je god potrebno.	1

Većina ispitanika praktički ne pravi razliku u načinu izvođenja obrazovnog programa uživo ili online.

Detaljniji pregled rezultata dan je u Tablici 20, a isto tako dan je i grafički pregled distribucije odgovora ispitanika.

Tablica 20. Način izvođenja obrazovnog programa

Način izvođenja obrazovnog programa	Broj zaprimljenih odgovora	%
Uživo	16	34,04%
Online	9	19,15%
Oboje od prethodno navedenog	21	44,68%
Ne znam	0	0,00%
Nisam zainteresiran/a	1	2,13%



Slika 19. Pregled odgovora – način izvođenja obrazovnog programa

### 3.5 Rezultati grupe pitanja „Završni komentar“

Završni komentar je unijelo 17 ispitanika. Detaljan pregled komentara dan je u Tablici 21.

Tablica 21. Pregled završnih komentara

Završni komentari
Pozdrav ekipa, najbolji ste (:
Radionica o korištenju Isabelle je odlična, svakako bi volio ponovno sudjelovati na njoj.
<p>Volio bih pristupiti detaljnoj edukaciji i osposobljavanju u radu s HPC resursima usmjerenima prema računalnoj kemiji i modeliranju fizikalno–kemijskih sustava čija svrha i osmišljenost je takva da kao rezultat proizvede osobu sposobnu za samostalan stručni rad, te samostalno usavršavanje i razvoj vještina.</p> <p>Ovo bi zahtjevalo barem dvogodišnju neprekidnu edukaciju; Ali bi takva, vjerujem, bila komplementarna s post–doc i doktorskim studijem.</p>
Most of the times the SRCE cluster is always busy and it takes days for the task in the queue to get into execution.
Isabella je neupotrebljiva s redovima čekanja koji traju po nekoliko dana. Potom, kada se konačno pokrene proces i nastupi problem (utvrdi se bug koji se ranije na manjem datasetu nije mogao utvrditi možda i zbog drugačije konfiguracije testnog računala i Isabella klastera), opet se vraćamo u red čekanja od nekoliko dana. Glupost. To neopisivo usporava istraživanje, do totalnog zastoja. Pored toga, zadnji moji eksperimenti prekidal su se bez ikakvih naznaka razloga, a na nekoliko drugih računala i klastera uredno rade. Stoga sam potpuno odustao od korištenja Isabelle.
Bilo bi praktično kada bi bilo jasno navedeno kome se obratiti za instalaciju besplatnog software–a na isabellu i kada bi se povećala količina radne memorije na najčešće korištenim node–ovima.
uvodjenje nastave kvantne kemije na dodiplomsku razinu je moguće samo uz HPC sustave.
Obrazovni sadržaji su super, ali problem je kad zbog toga moram putovati na drugi kraj grada, kad nema mjesta za parkiranje. Programi uzivo omogućavaju bolju interakciju, ali potroše puno više vremena. Najbolje je on–line, s mogućnošću da se i kasnije pogleda snimka ili prezentacija i da se postavi pitanje predavacu ili grupi tijekom predavanja ili poslije njega.
Broj procesora, memorija, diskovni prostor i ostali resursi koji su mi potrebni ovise o programima koje kanim koristiti za razne znanstvene probleme (uglavnom programe pod GNU licencom). Kako često takve programe "isprobavam" po prvi puta, unaprijed nemogu znati točne brojeve za nevedene resurse, pa su oni u upitniku navedeni vrlo proizvoljno.

Ja sam nabavio, instalirao i održavam računalni klaster Zavoda za teorijsku fiziku IRB-a, te imam veliko iskustvo s korištenjem, kompajliranjem, i razvijanjem HPC kodova, pogotvo DFT (density functional theory) pa također mogu održati radionicu po potrebi.
Zahvaljujem na trudu!
Nastavite s odličnim radom!
Iako sam prije ispunio da nemam direktnu potrebu za korištenjem GPU procesora, smatram da treba nabavljati i osuvremenjivati i CPU i GPU resurse, te ponuditi radionice koje bi korisnicima omogućile i potaknule ih da koriste GPU resurse.
Ovakvi upitnici su dobrodošli i nadam se brznoj realizaciji na temelju njihovih rezultata.
Osobno ne koristim ali postoje potrebe unutar projekta i posao odrađuje kvalificirani suradnik.
Samo naprijed!
Sretno! Ja možda nisam najbolja mušterija, ali podržavam sve što radite.

#### 4. Zaključak

Obzirom na broj direktno adresiranih ispitanika (291), 47 zaprimljenih potpunih odgovora je jako dobar rezultat provedbe ankete. Ispitanici pripadaju znanstveno-istraživačkoj zajednici koja na dnevnoj bazi koristi HPC resurse za potrebe istraživanja. Ispitanici vrlo dobro znaju svoje potrebe za HPC resursima iskazanim kroz broj procesorskih jezgri ili broj grafičkih procesora, svjesni su svojih prosječnih i maksimalnih potreba te mogu komunicirati svoje godišnje potrebe za HPC resursima kroz CPU/GPU sate. Prirodne znanosti, konkretno znanstvena polja Fizika i Kemija, su područja iz kojih dominantno dolaze istraživači. Njihova mišljenja kao i mišljenja istraživača iz ostalih znanstvenih područja (računarstvo, temeljne medicinske znanosti, interdisciplinarne biotehničke znanosti, itd.) su izuzetno važna za stvaranje boljih uvjeta korištenja HPC resursa za postojeće i nove korisnike tih sustava.

Značajan broj ispitanika (23,4%) se je susreo s problemom u istraživačkom radu za čije učinkovito rješavanje je nužan pristup računalnim resursima većim od 1 PFLOPS. Ova informacija je značajna za dizajn budućih HPC sustava u Hrvatskoj te pružanje podrške u korištenju EuroHPC resursa.

Najveća prepreka kod korištenja većih resursa od onih koje trenutno koriste vide u nedostatku vještina programiranja.

Nekoliko glavnih točaka se može sažeti kako bi poslužile kao smjernica za razvoj budućih treninga u HPC ekosustavu Republike Hrvatske. Sljedeće preporuke temelje se na djelomičnoj usporedbi rezultata ove faze ankete s raspoloživim obrazovnim programima koje trenutno nude pružatelji usluga obrazovanja u Republici Hrvatskoj.

Na ponuđenih 12 tema obuhvaćenih unutar ovog prijedloga obrazovnog programa ispitanici se izjasnili o interesu za pohađanjem sljedećih ponuđenih tema s određenom razinom edukacije pri čemu su izdvojene teme najvećim postotkom odabira za pojedinu razinu:

- Korištenje i pristup HPC resursima – napredna razina
- Specifične vrste simulacija/modeliranja – napredna razina
- Specifični HPC softver – srednja razina
- Programski jezici – uvodna razina
- Moderni programski jezici visokih performansi (npr. Julia ili Python) – uvodna i napredna razina
- MPI (Message Passing Interface) – uvodna razina
- OpenMP (Open Multi-Processing)
- Veliku većinu ispitanika ne zanimaju područja obrazovanja u području Umjetne inteligencije (38,30%), HPDA (High Performance Data Analytics) (31,90%) i Izvedbeni inženjering (38,30%). Ipak, za područje obrazovanja Umjetna inteligencija 31,94% ispitanika ima potrebu obrazovanja na uvodnoj razini. Iako prepoznati fakulteti nude ovu vrstu obrazovanja usmjerenu studentskoj populaciji te je i udruga CroAI [4] lansirala besplatni on-line trening umjetne inteligencije krajem 2020. i dalje postoji velika potreba za pružanjem ovog obrazovnog sadržaja i to na uvodnoj razini
- Najkorisniji način izvođenja obrazovnog programa je vođene vježbe/hands-on sesije. Prezentacije sa slajdovima treba izbjegavati
- Većina ispitanika (44,68%) ne pravi razliku između on-line ili on-site izvođenja treninga, te je prepoznata prednost dosega on-line treninga
- Očekivano trajanje treninga je do tjedan dana.

Kako je unutar svake od 12 ponuđenih tema treninga mala razlika između razine interesa u provođenju treninga na određenoj razini, sljedeća faza ankete će obuhvatiti više tema podijeljenih u tri područja HPC, HPDA i AI, te će obuhvatiti i druge ciljane skupine u znanstveno-istraživačkoj zajednici te industriju.



## 5. Reference

[1] "LimeSurvey," [Online]. Available: <https://limesurvey.srce.hr>.

[2] "Znanstveni centri izvrsnosti," [Online]. Available: <https://zci.hr/hr/>.

[2] "Računalni klaster Isabella," [Online]. Available:  
<https://www.srce.unizg.hr/en/isabella-cluster>.

[3] "CroAI," [Online]. Available: <https://www.croai.org/>.