



Digitalni repozitoriji znanstvenih informacija

Alen Vodopijevec <alen@irb.hr>
Knjižnica Instituta Ruđer Bošković

Sadržaj prezentacije



- Uloga digitalnih repozitorija u procesu prikupljanja, pohrane te obrade i publiciranja rezultata znanstvenog rada
- Primjer: CROSB I i povezivanje s institucijskim repozitorijima

Osnovni pojmovi



Digitalni repozitorij → DR

- Online servis, namijenjen učinkovitom rukovanju pohranjenim sadržajem s mogućnošću povezivanja više instanci u interoperabilni distribuirani sustav koji će osim samog pristupa podacima omogućiti i ekstrakciju znanja na višoj razini.
- baza podataka koja podrazumijeva i web-servis

Institucijski repozitorij → IR

- uži pojam
- odnosi se na digitalni repozitorij čija je glavna svrha arhiviranje outputa znanstvenika i vezan je uz određenu ustanovu ili znanstveno područje.

Pitamo se



- ZAŠTO želimo znanstvene informacije učiti dostupnijima na Web-u?
- ŠTO točno želimo pohraniti u DR?
- KAKO se to radi?



ZAŠTO?

Čemu?

...

Činjenice



- Zasićenost *World Wide Web-a* velikom količinom nestrukturiranih ili djelomično strukturiranih nestandardiziranih informacija;
- očigledna je nemogućnost šireg pristupa rezultatima (javno financiranog) znanstvenog rada;
- nemogućnost povezivanja podataka iz različitih DR-a;
- porast troškova pretplate na časopise;
- potreba za otvaranjem znanosti.



Olakotne okolnosti

“Znanstvena i tehnologijska politika Republike Hrvatske 2006-2010. godine”

Dokument prihvaćen na sjednici Vlade Republike Hrvatske održanoj 5. svibnja 2006. godine.



Citat:

“Znanstveno tehnološki sustav koji se financira iz javnih sredstava mora biti otvoren za javnost. Trebale bi biti jasne razlike između javnih sredstava i onih pribavljenih na tržištu.”

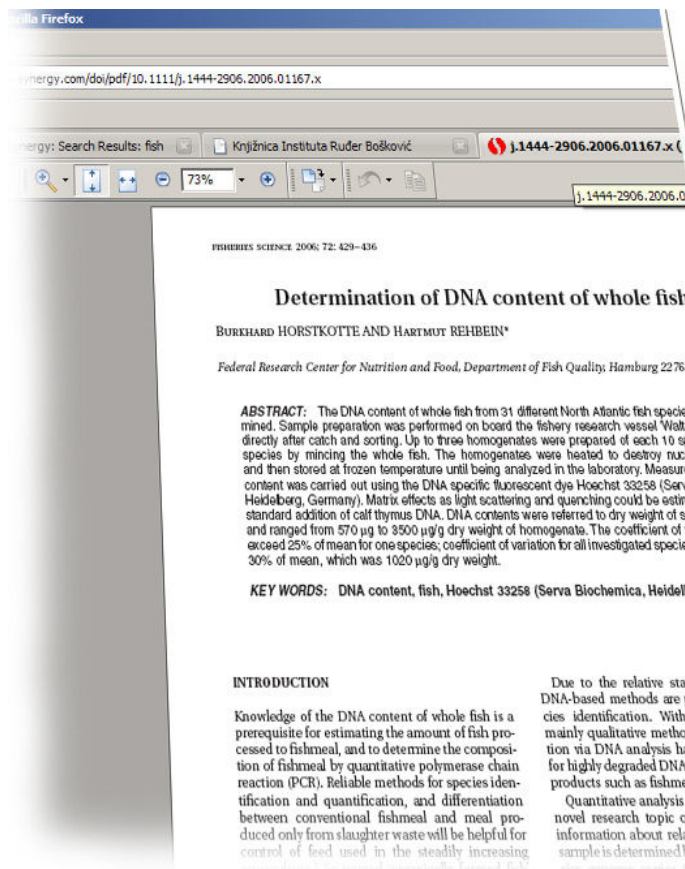
Rezultati istraživanja i razvoja financiranog iz javnih sredstava trebali bi biti dostupni javnosti u obliku otvorenih publikacija ili otvorenog pristupa bazama podataka itd.”



ŠTO?

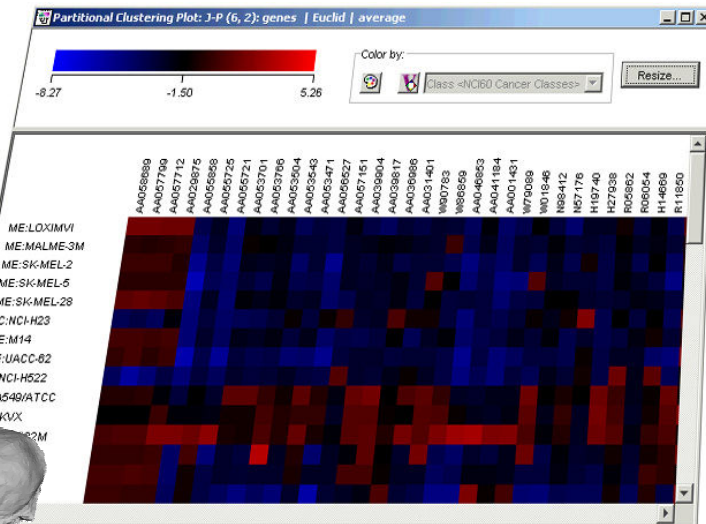
Koje vrste podataka?

Sadržaj DR-a



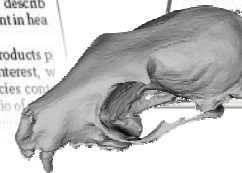
| Variable | Type | Distinct | Missing | Min | Max | Avg | Std |
|----------|---------|----------|---------|-----|------|------|-----|
| R01 | integer | 3000 | 0 | 1 | 3000 | 1500 | 866 |
| R02 | char | 2 | 0 | N | Y | 0 | 0 |
| R03 | char | 2 | 0 | A | B | 0 | 0 |
| R04 | char | 2 | 0 | C | D | 0 | 0 |
| R05 | char | 4 | 0 | au | uk | 0 | 0 |
| R06 | char | 4 | 0 | au | uk | 0 | 0 |
| R07 | char | 3 | 0 | au | uk | 0 | 0 |
| R08 | char | 3 | 0 | au | uk | 0 | 0 |
| R09 | char | 3 | 0 | au | uk | 0 | 0 |
| R10 | char | 1 | 0 | go | go | 0 | 0 |
| R11 | integer | 7 | 0 | 3 | 9 | 7 | 1 |
| R12 | integer | 10 | 0 | 0 | 9 | 8 | 0 |

Dataset – izvorni podaci



Znanstveni članak u časopisu

Vizualizacija obrađenih podataka





KAKO?

Alati, strategija

Alati



- E-prints
[<http://www.eprints.org>]



- DSpace
[<http://www.dspace.org/>]



- CDS Invenio
[<http://cdsware.cern.ch/cdsware/overview.html>]

fedora



- FEDORA
[<http://www.fedora.info>]

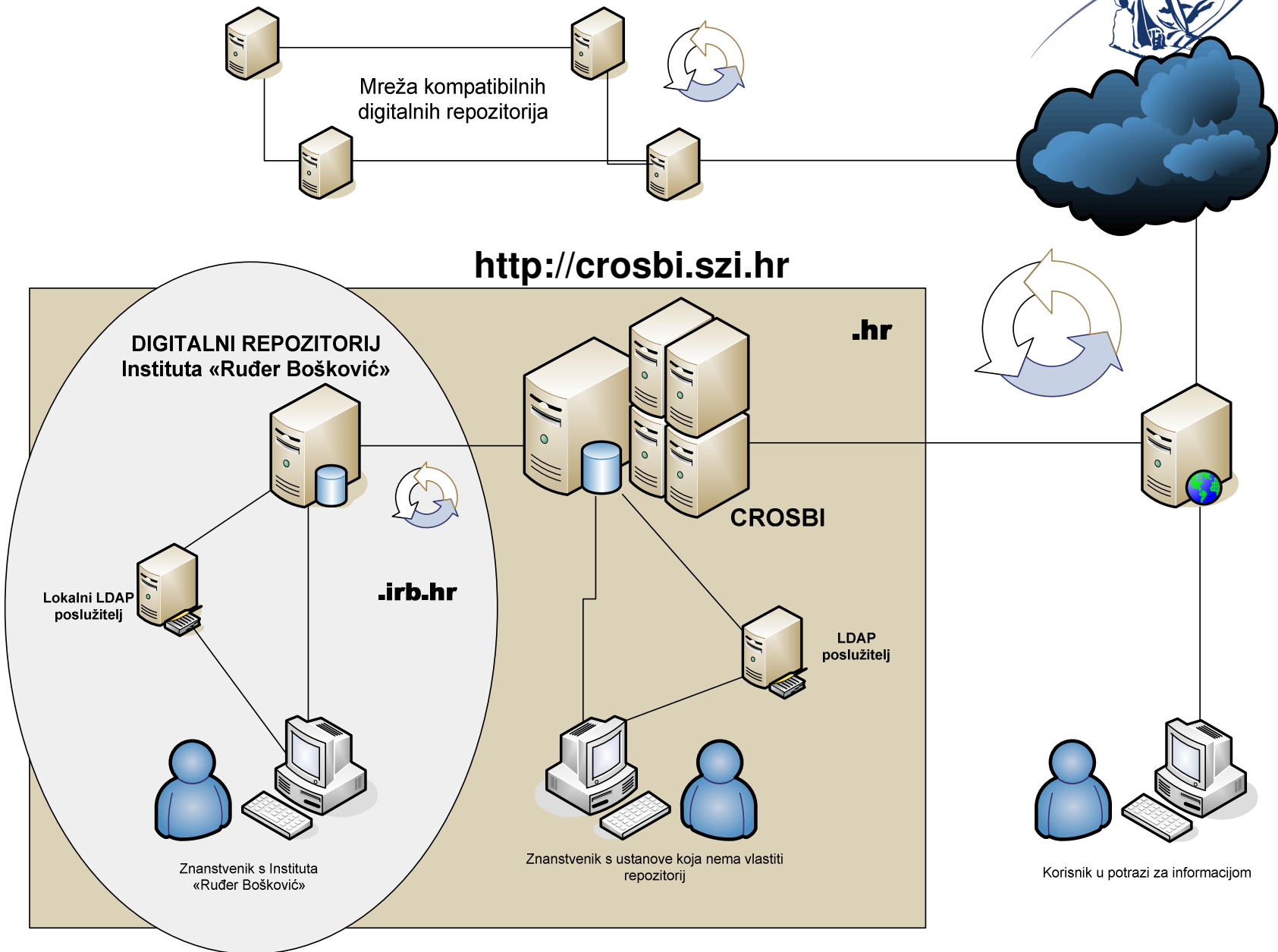
•Virtual Data Center
[<http://www.thedata.org/>]

Zadaci za rješavanje

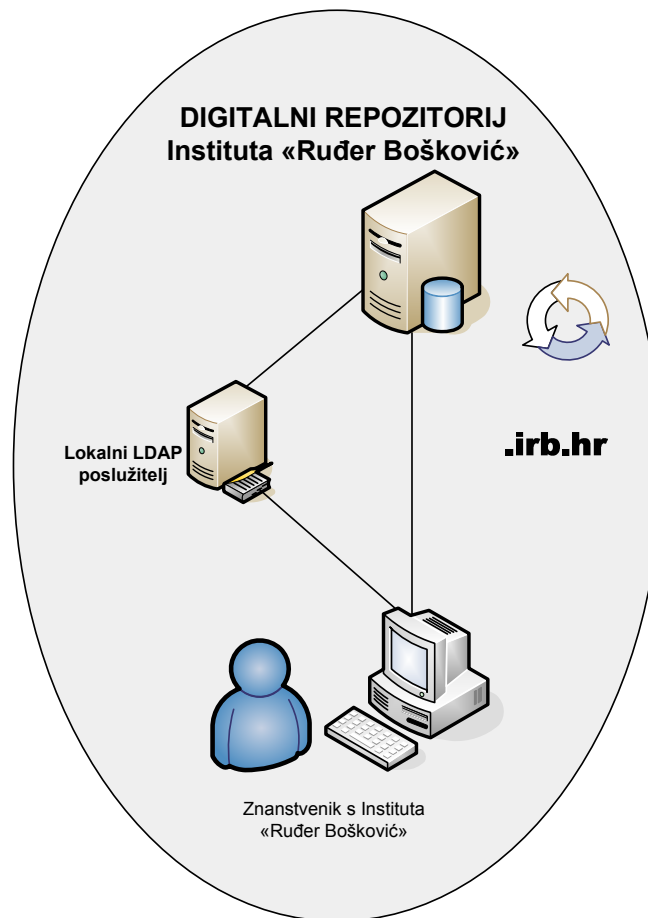


- **Prikupljanje podataka**
 - obaveza samoarhiviranja rezultata znanstvenih istraživanja financiranih javnim sredstvima, problem autorskih prava;
- **Pohrana podataka**
 - ujednačavanje podataka prikupljenih različitim metodama
 - kako smanjiti rizik od gubitka podataka te minimizirati eventualne štete;
- **Reguliranje pristupa**
 - za znanstvenu i istraživačku zajednicu (npr., pristup računalnim resursima za procesiranje podataka temeljem elektroničkog identiteta u sustavu AAI@EduHr);
- **Interoperabilnost**
 - standardi metapodataka i format pohranjenih datoteka;
- **Servisi**
 - softverske aplikacije koje bi omogućile kvalitetnu manipulaciju tako velikim setom dostupnih informacija u svrhu ekstrakcije znanja.

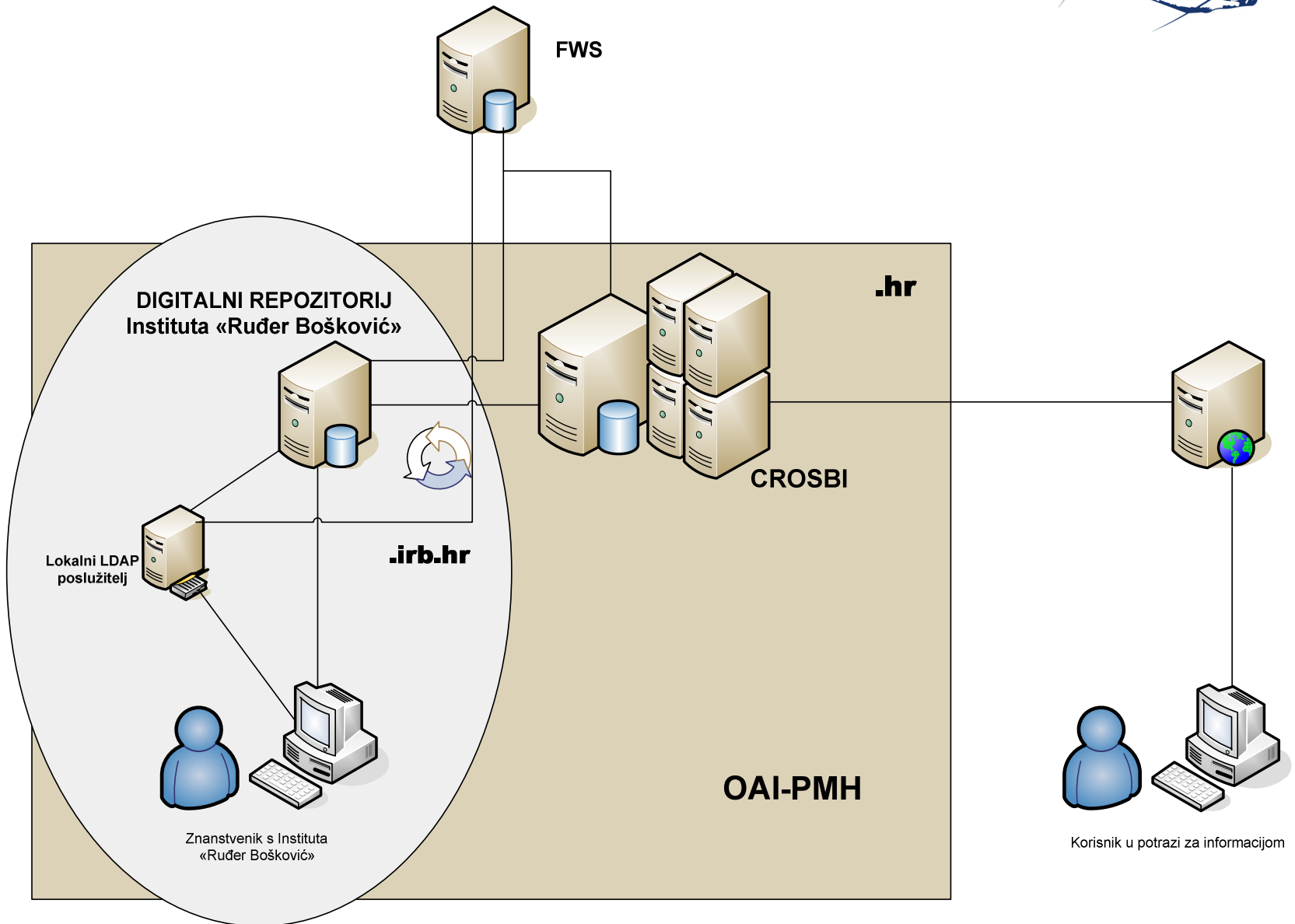
Integracija IR-a s CROSBi-jem (1)



Integracija IR-a s CROSBi-jem (2)



Integracija IR-a s CROSBi-jem (4)

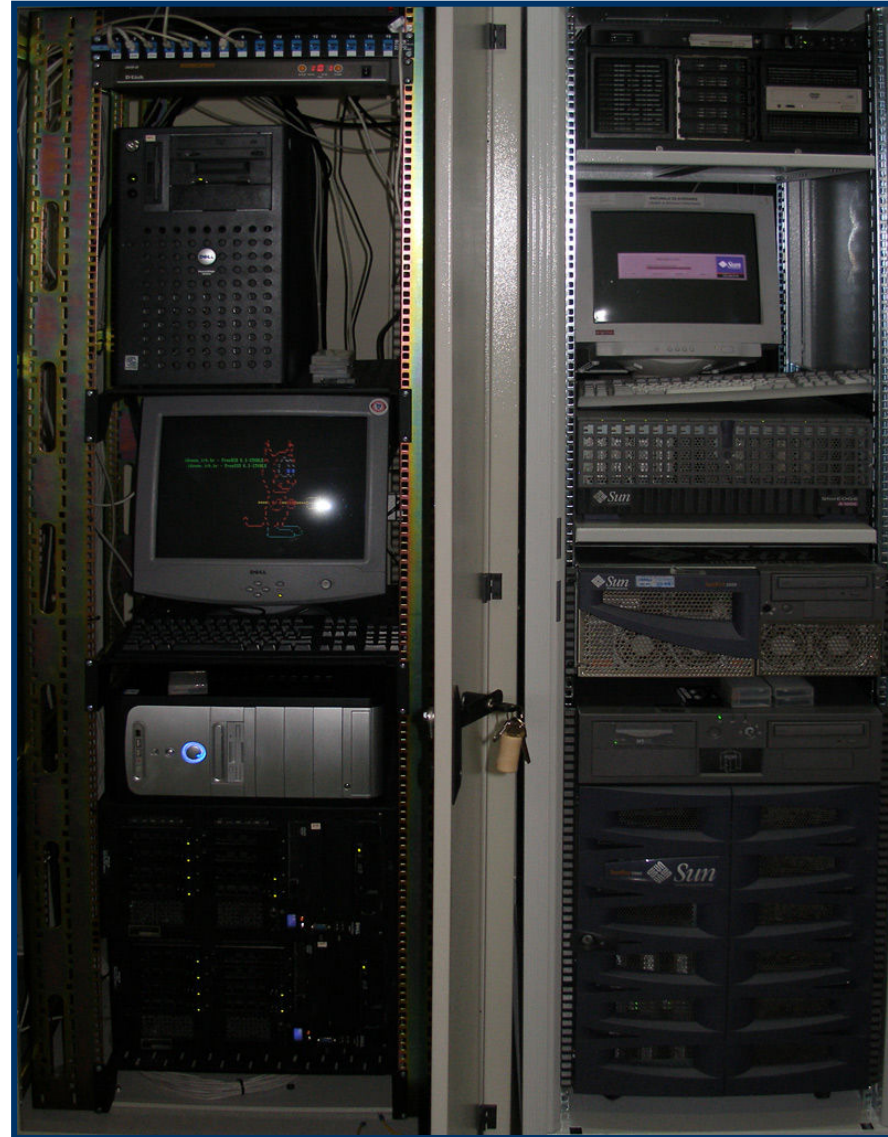


Svrha implementacije IR-a



- pohrana i očuvanje digitalnih sadržaja,
- jednostavan pristup, pretraživanje i pregled vlastitih dokumenata,
- stvaranje uvjeta znanstvenicima za samostalno arhiviranje svojih radova,
- publiciranje rezultata znanstvenog rada i istraživanja na temeljima “otvorenog pristupa” i u skladu s time
- povećanje vidljivosti i utjecaja tih radova u znanstvenoj zajednici na globalnoj razini.

Jesam, jesam, vidio sam ga





Digitalni repozitoriji znanstvenih informacija

Alen Vodopijevec <alen@irb.hr>
Knjižnica Instituta Ruđer Bošković