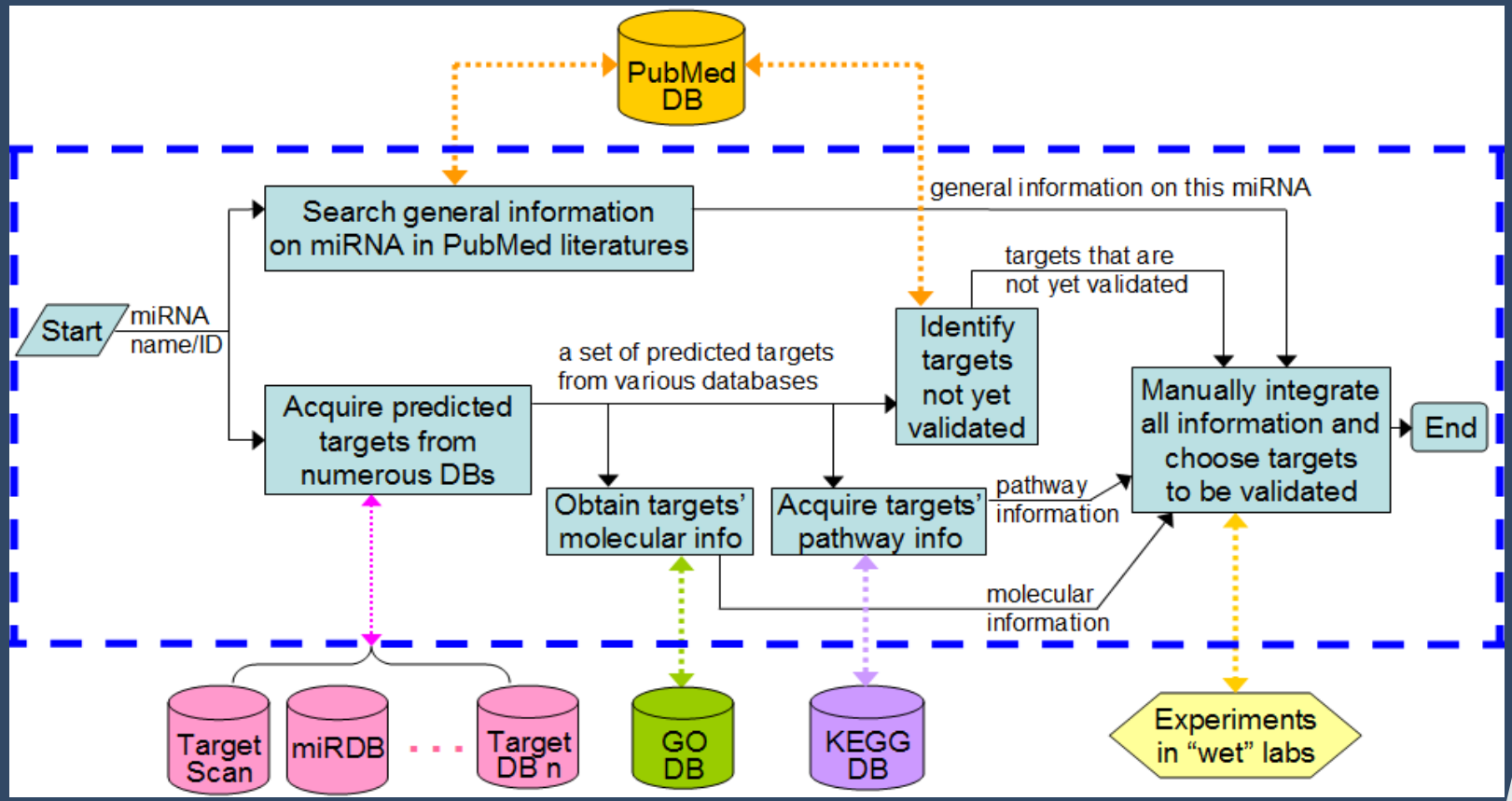
Project Overview

- OmniSearch is a semantic search software specifically designed for cancer biologists
- It will assist biologists and bioinformaticians in unraveling critical roles of microRNAs (a.k.a. miRNAs or miRs) in various human cancers
- OmniSearch can be used to obtain unified knowledge related to miRs and thus derive unique insights for the regulation and control of cancer disease processes

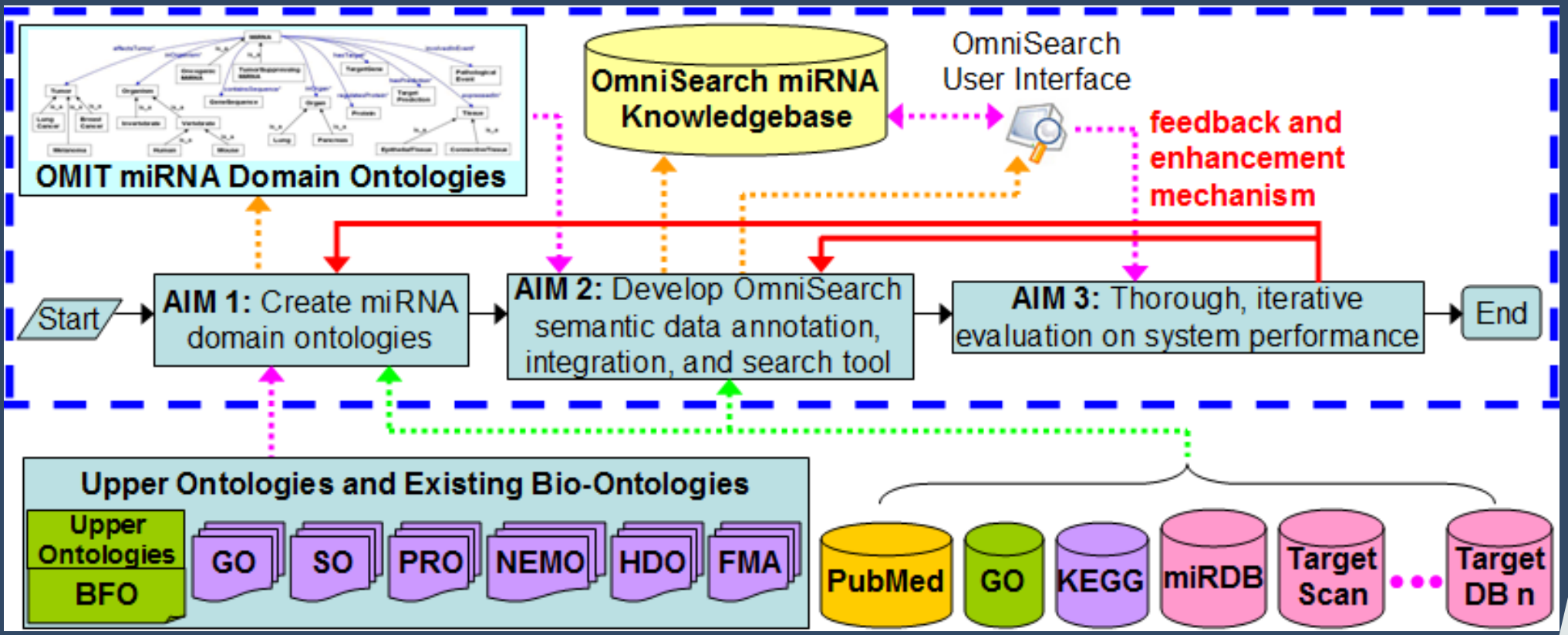
Research Motivation

- Manual integration of information from heterogeneous sources has become labor-intensive and error-prone

A typical workflow demonstrating current search approaches adopted by many cancer biologists



Proposed Solution



Significance
Innovation

1. Handles the urgent need of effective miRNA data sharing, data integration, and knowledge acquisition in everyday human cancer research

Significance
Innovation

1. Handles the urgent need of effective miRNA data sharing, data integration, and knowledge acquisition in everyday human cancer research
2. Significantly accelerates cancer biology research

Significance
Innovation

1. Handles the urgent need of effective miRNA data sharing, data integration, and knowledge acquisition in everyday human cancer research
2. Significantly accelerates cancer biology research
3. Tackles four critical needs recognized by ITCR

Significance
Innovation

1. Handles the urgent need of effective miRNA data sharing, data integration, and knowledge acquisition in everyday human cancer research
2. Significantly accelerates cancer biology research
3. Tackles four critical needs recognized by ITCR
   - establishing data exchange standards and common data elements
   - sustained effort to promote data sharing
   - enhanced support of community-based, research-driven informatics technology development
   - improved mechanisms to support software development

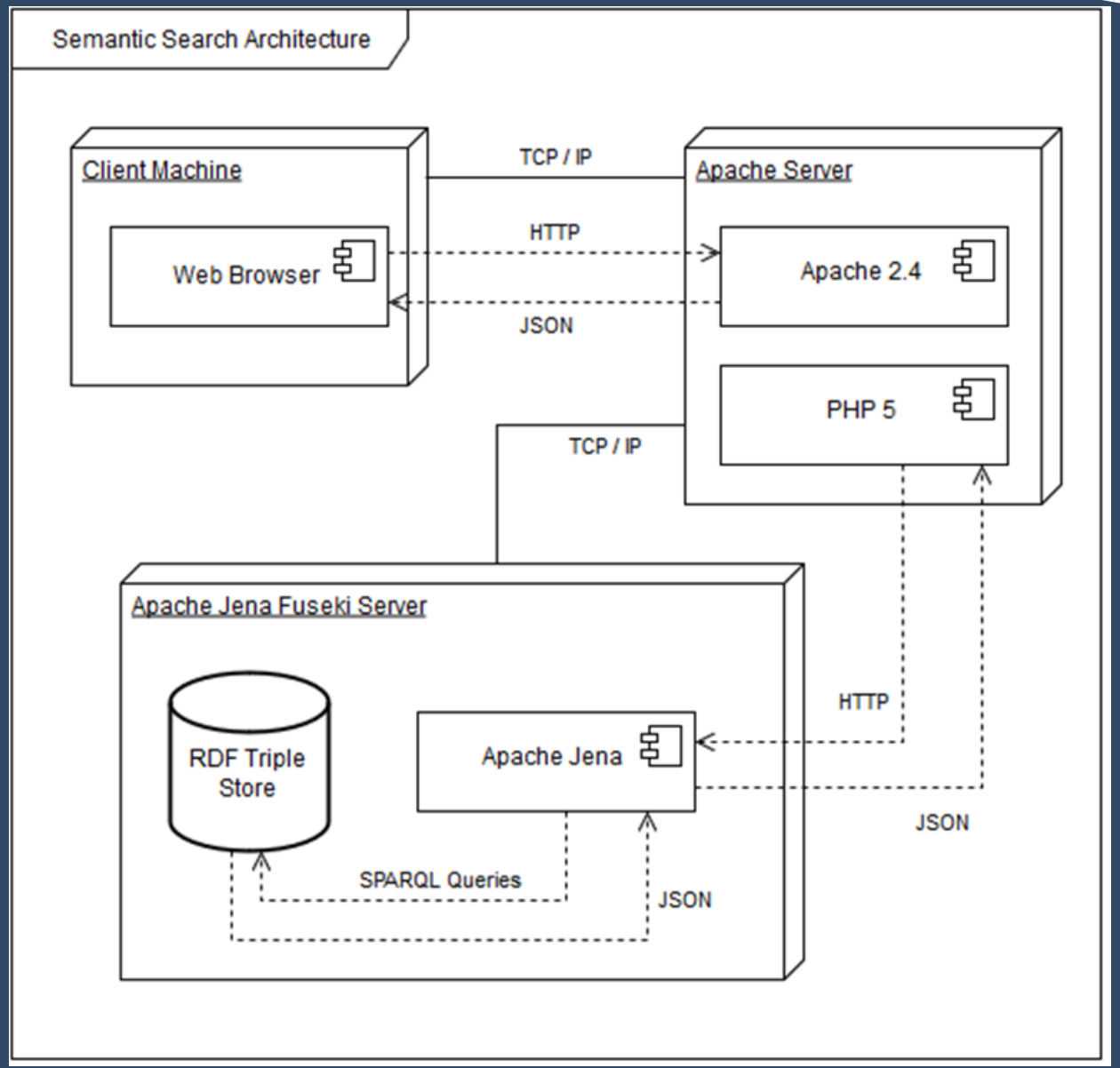
Significance
Innovation

1. Handles the urgent need of effective miRNA data sharing, data integration, and knowledge acquisition in everyday human cancer research
2. Significantly accelerates cancer biology research
3. Tackles four critical needs recognized by ITCR
   - establishing data exchange standards and common data elements
   - sustained effort to promote data sharing
   - enhanced support of community-based, research-driven informatics technology development
   - improved mechanisms to support software development
4. OMIT is the first ontology of its kind that formally encodes miRNA knowledge, providing the community with a systematically structured, precisely defined controlled vocabulary, along with relationships among concepts, thus rendering precise semantics for the domain

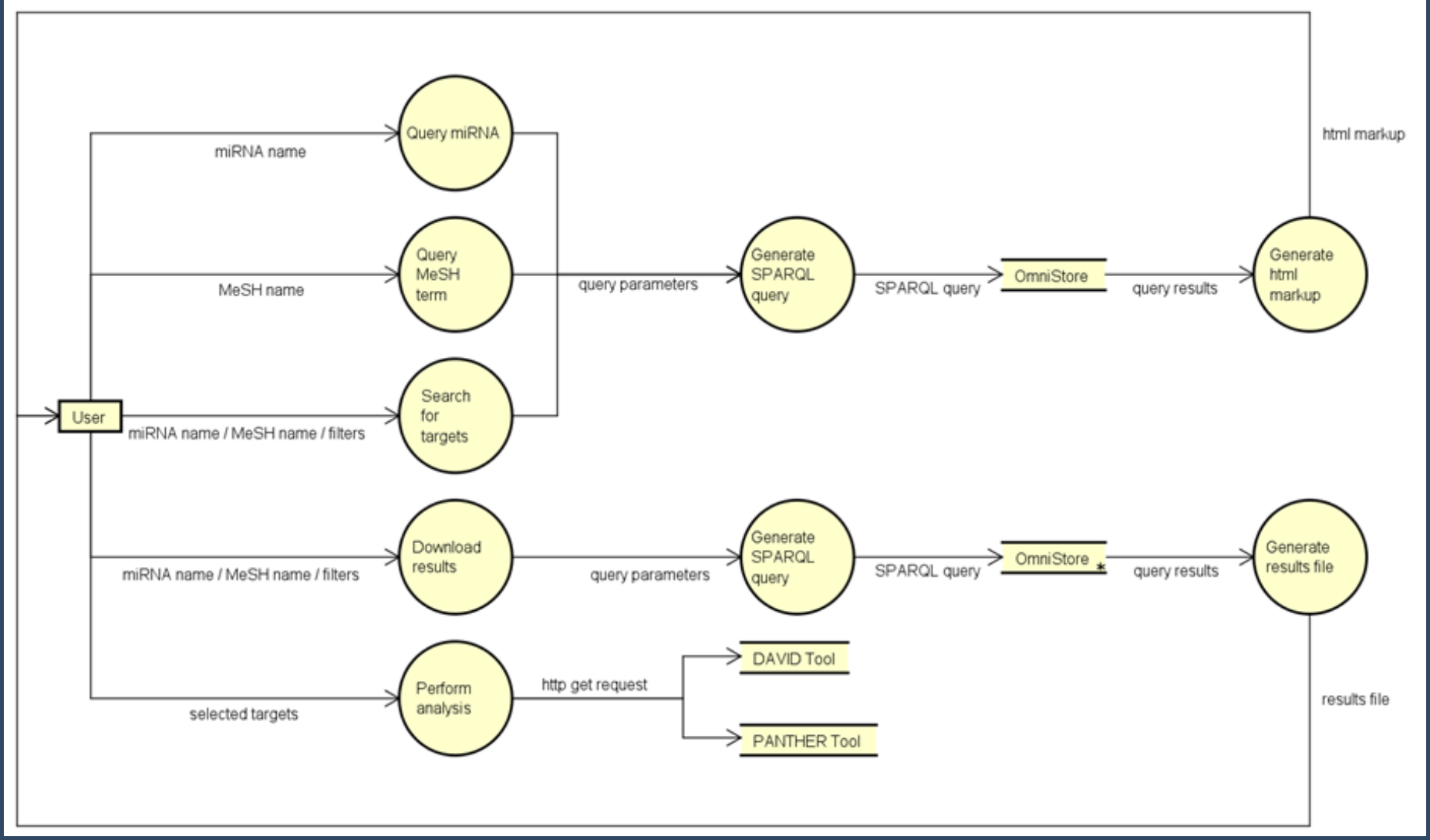
Significance
Innovation

1. Handles the urgent need of effective miRNA data sharing, data integration, and knowledge acquisition in everyday human cancer research
2. Significantly accelerates cancer biology research
3. Tackles four critical needs recognized by ITCR
   - establishing data exchange standards and common data elements
   - sustained effort to promote data sharing
   - enhanced support of community-based, research-driven informatics technology development
   - improved mechanisms to support software development
4. OMIT is the first ontology of its kind that formally encodes miRNA knowledge, providing the community with a systematically structured, precisely defined controlled vocabulary, along with relationships among concepts, thus rendering precise semantics for the domain
5. Contributes to bio-ontology community by term reuse and ontology cross-referencing

Software Architecture



User Interface Flowchart



Project Timeline

	Project Activities	Milestones by the End of Each Year
Y1	• Start ontology development • Start data annotation & integration	• Initial version of OMIT ontologies • “Key Phrase Extraction” & “Ontology Mapping” modules
Y2	• Continue ontology development • Continue data annotation & integration	• Stable version of OMIT ontologies • New terminology contributed to bio-ontology community • Other software modules; initial OmniSearch package
Y3	• Design a friendly semantic search GUI • Develop use cases and evaluating queries • Conduct iterative system evaluation • Integrate feedback from the community	• Final version of OMIT ontologies and OmniSearch package • A comprehensive miRNA KB unified from numerous sources • A collection of use cases on miRNA knowledge acquisition • A set of semantic search evaluating queries • A collection of formative evaluation reports • A final, summative evaluation report

Project Timeline

	Project Activities	Milestones by the End of Each Year
Y1	• Start ontology development • Start data annotation & integration	• Initial version of OMIT ontologies • “Key Phrase Extraction” & “Ontology Mapping” modules
Y2	• Continue ontology development • Continue data annotation & integration	• Stable version of OMIT ontologies • New terminology contributed to bio-ontology community • Other software modules; initial OmniSearch package
Y3	• Design a friendly semantic search GUI • Develop use cases and evaluating queries • Conduct iterative system evaluation • Integrate feedback from the community	• Final version of OMIT ontologies and OmniSearch package • A comprehensive miRNA KB unified from numerous sources • A collection of use cases on miRNA knowledge acquisition • A set of semantic search evaluating queries • A collection of formative evaluation reports • A final, summative evaluation report

Project Dissemination

- **Collaborating with a wide range of bio-ontology community:**
  Gene Ontology (GO), Sequence Ontology (SO), PRotein Ontology (PRO), Chemical Entities of Biological Interest Ontology (CHEBI), Ontology for Biomedical Investigations (OBI), and Uber Anatomy Ontology (UBERON)

Project Dissemination

- **Collaborating with a wide range of bio-ontology community:**
  Gene Ontology (GO), Sequence Ontology (SO), PRotein Ontology (PRO), Chemical Entities of Biological Interest Ontology (CHEBI), Ontology for Biomedical Investigations (OBI), and Uber Anatomy Ontology (UBERON)

- **OBO Foundry:** http://www.obofoundry.org/ontology/omit.html
- **NCBO BioPortal:** https://bioportal.bioontology.org/ontologies/OMIT
- **Project wiki site:** http://omnisearch.soc.southalabama.edu/
- **GitHub:** https://github.com/OmniSearch

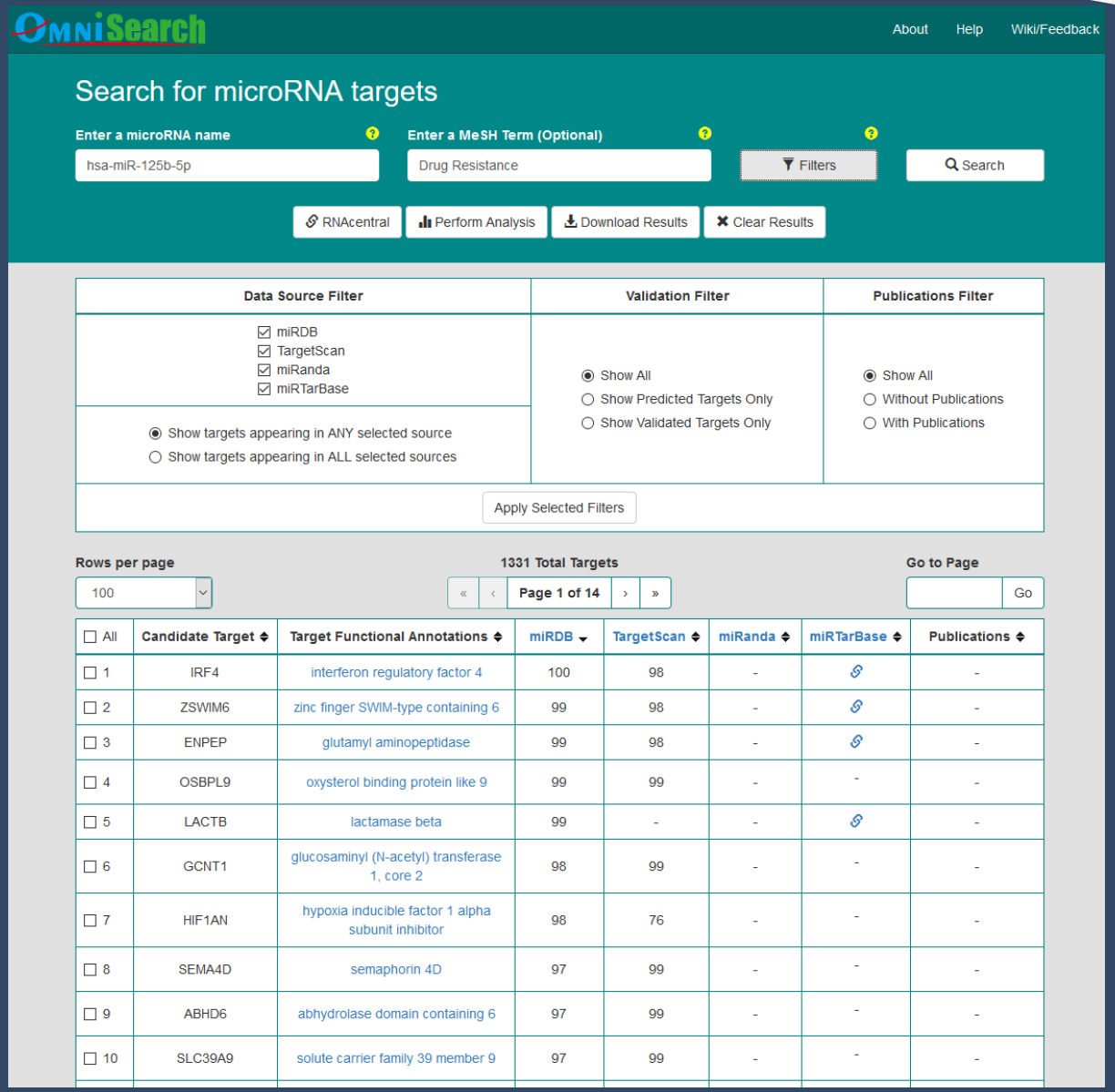
Project Dissemination

- **Collaborating with a wide range of bio-ontology community:**
  Gene Ontology (GO), Sequence Ontology (SO), PRotein Ontology (PRO), Chemical Entities of Biological Interest Ontology (CHEBI), Ontology for Biomedical Investigations (OBI), and Uber Anatomy Ontology (UBERON)

- **OBO Foundry:** http://www.obofoundry.org/ontology/omit.html
- **NCBO BioPortal:** https://bioportal.bioontology.org/ontologies/OMIT
- **Project wiki site:** http://omnisearch.soc.southalabama.edu/
- **GitHub:** https://github.com/OmniSearch

- **GUI Tutorial on YouTube:** https://www.youtube.com/watch?v=kCFm4YkNvEg

OmniSearch Software Interface



All	Candidate Target	Target Functional Annotations	miRDB	TargetScan	miRanda	miRTarBase	Publications
1	IRF4	interferon regulatory factor 4	100	98	-		-
2	ZSWIM6	zinc finger SWIM-type containing 6	99	98	-		-
3	ENPEP	glutamyl aminopeptidase	99	98	-		-
4	OSBPL9	oxysterol binding protein like 9	99	99	-	-	-
5	LACTB	lactamase beta	99	-	-		-
6	GCNT1	glucosaminyl (N-acetyl) transferase 1, core 2	98	99	-	-	-
7	HIF1AN	hypoxia inducible factor 1 alpha subunit inhibitor	98	76	-	-	-
8	SEMA4D	semaphorin 4D	97	99	-	-	-
9	ABHD6	abhydrolase domain containing 6	97	99	-	-	-
10	SLC39A9	solute carrier family 39 member 9	97	99	-	-	-

Recent Publications

6) J. Huang, G. Borchert, D. Dou, J. Huan, W. Lan, M. Tan, and B. Wu, editors, **Bioinformatics in microRNA research: computational methods in exploring microRNAs' functions**, (in-progress) Springer series of Methods in Molecular Biology, 2016.

5) J. Huang, F. Gutierrez, H. Strachan, D. Dou, W. Huang, B. Smith, J.A. Blake, K. Eilbeck, D.A. Natale, Y. Lin, B. Wu, N. de Silva, X. Wang, Z. Liu, G.M. Borchert, M. Tan, and A. Ruttenberg, "**OmniSearch: A semantic search system based on the Ontology for MIcroRNA Target (OMIT) for microRNA-target gene interaction data**," *J Biomed Semantics*. 2016 May 10;7:25. doi: 10.1186/s13326-016-0064-2. eCollection 2016. PubMed PMID: 27175225; PubMed Central PMCID: PMC4863347.

4) J. Huang, K. Eilbeck, B. Smith, J.A. Blake, D. Dou, W. Huang, D.A. Natale, A. Ruttenberg, J. Huan, M.T. Zimmermann, G. Jiang, Y. Lin, B. Wu, H. Strachan, Y. He, S. Zhang, X. Wang, Z. Liu, G.M. Borchert, and M. Tan, "**The Non-Coding RNA Ontology (NCRO): A comprehensive resource for the unification of non-coding RNA biology**," *J Biomed Semantics*. 2016 May 4;7:24. doi: 10.1186/s13326-016-0066-0. eCollection 2016. PubMed PMID: 27152146; PubMed Central PMCID: PMC4857245.

3) J. Huang, K. Eilbeck, B. Smith, J.A. Blake, D. Dou, W. Huang, D.A. Natale, A. Ruttenberg, J. Huan, M.T., Zimmermann, G. Jiang, Y. Lin, B. Wu, H. Strachan, N. de Silva, M. Kasukurthi, V. Jha, Y. He, S. Zhang, X. Wang, Z. Liu, G. Borchert, and M. Tan, "**The Development of Non-Coding RNA Ontology**," *Int. J. Data Mining and Bioinformatics*, 15(3):214-232, June 2016.

2) J. Huang, K. Eilbeck, J.A. Blake, D. Dou, D.A. Natale, A. Ruttenberg, B. Smith, M.T. Zimmermann, G. Jiang, Y. Lin, B. Wu, Y. He, S. Zhang, X. Wang, H. Zhang, Z. Liu, and M. Tan, "**A domain ontology for the non-coding RNA field**," Proc. 2015 IEEE International Conference on Bioinformatics and Biomedicine (BIBM-15), pp. 621-624, IEEE, Washington D.C., Nov. 2015

1) J. Huang, F. Gutierrez, D. Dou, J.A. Blake, K. Eilbeck, D.A. Natale, B. Smith, Y. Lin, X. Wang, Z. Liu, M. Tan, and A. Ruttenberg, "**A semantic approach for knowledge capture of microRNA-target gene interactions**," Proc. BHI Workshop at 2015 IEEE International Conference on Bioinformatics and Biomedicine (BIBM-15), pp. 975-982, IEEE, Washington D.C., Nov. 2015

Acknowledgements

- NIH NCI ITCR Initiative
- University of South Alabama Team
- University of Oregon Team
- University of Buffalo – SUNY Team
- Gene Ontology
- Sequence Ontology
- PRotein Ontology

OMNiSearch

Questions?

THANKS!

YouTube Tutorial Link

https://www.youtube.com/watch?v=kCFm4YkNvEg