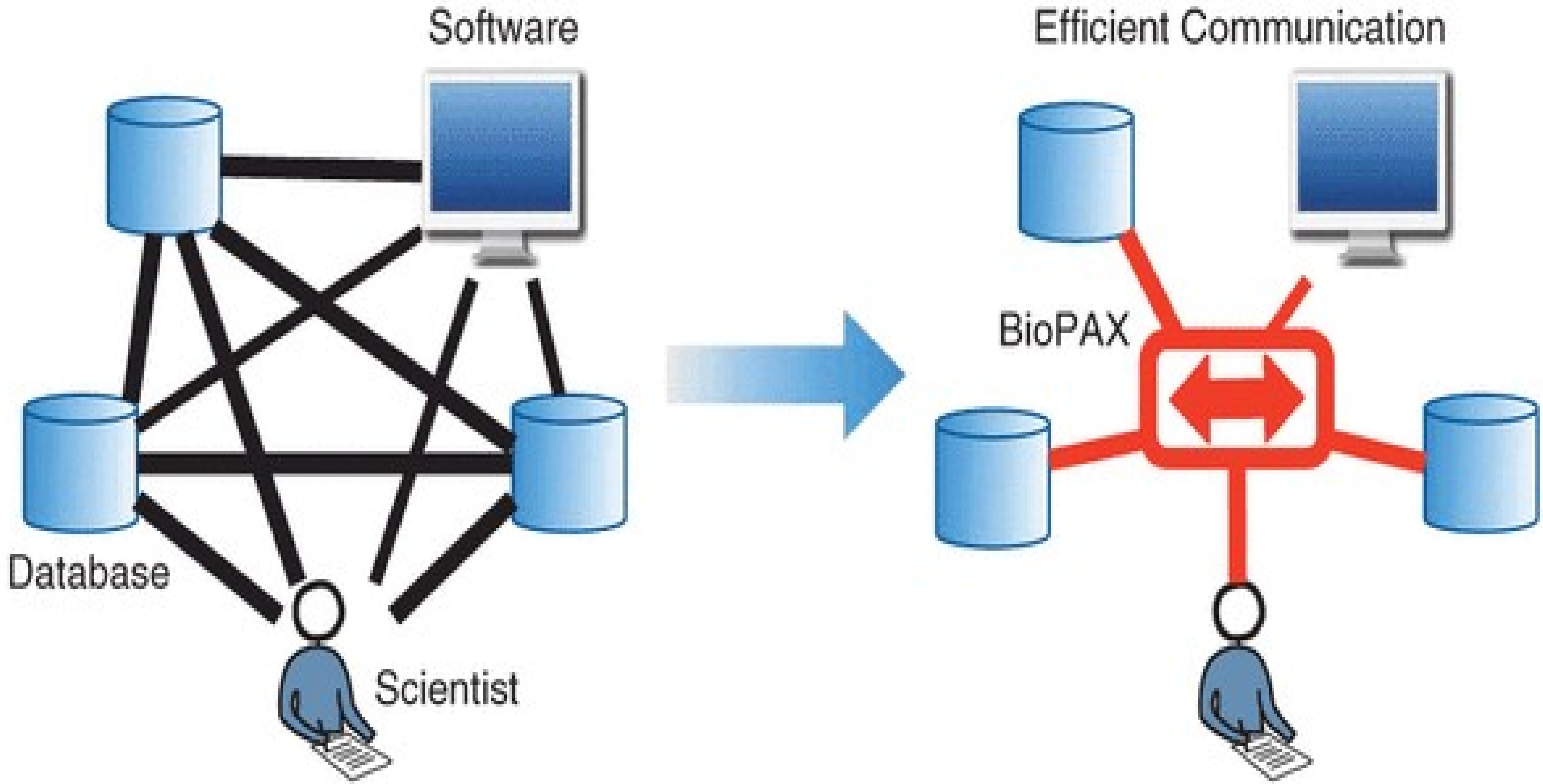
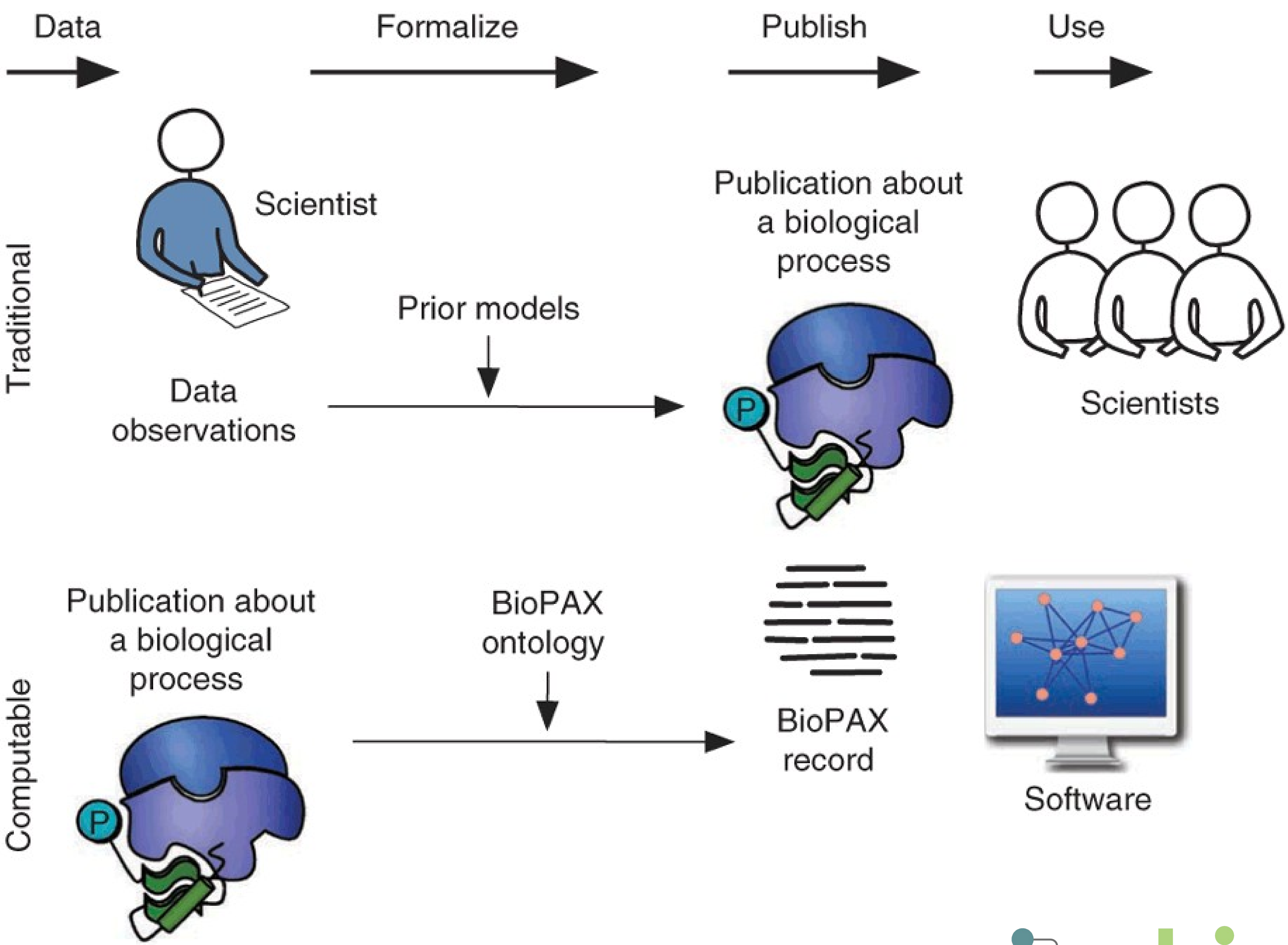
BioPAX: Biological Pathway Exchange

A community standard for representing pathways



Pathway information processing



Data → Formalize → Publish → Use

Traditional

Scientist

Data observations

Prior models

Publication about a biological process

Scientists

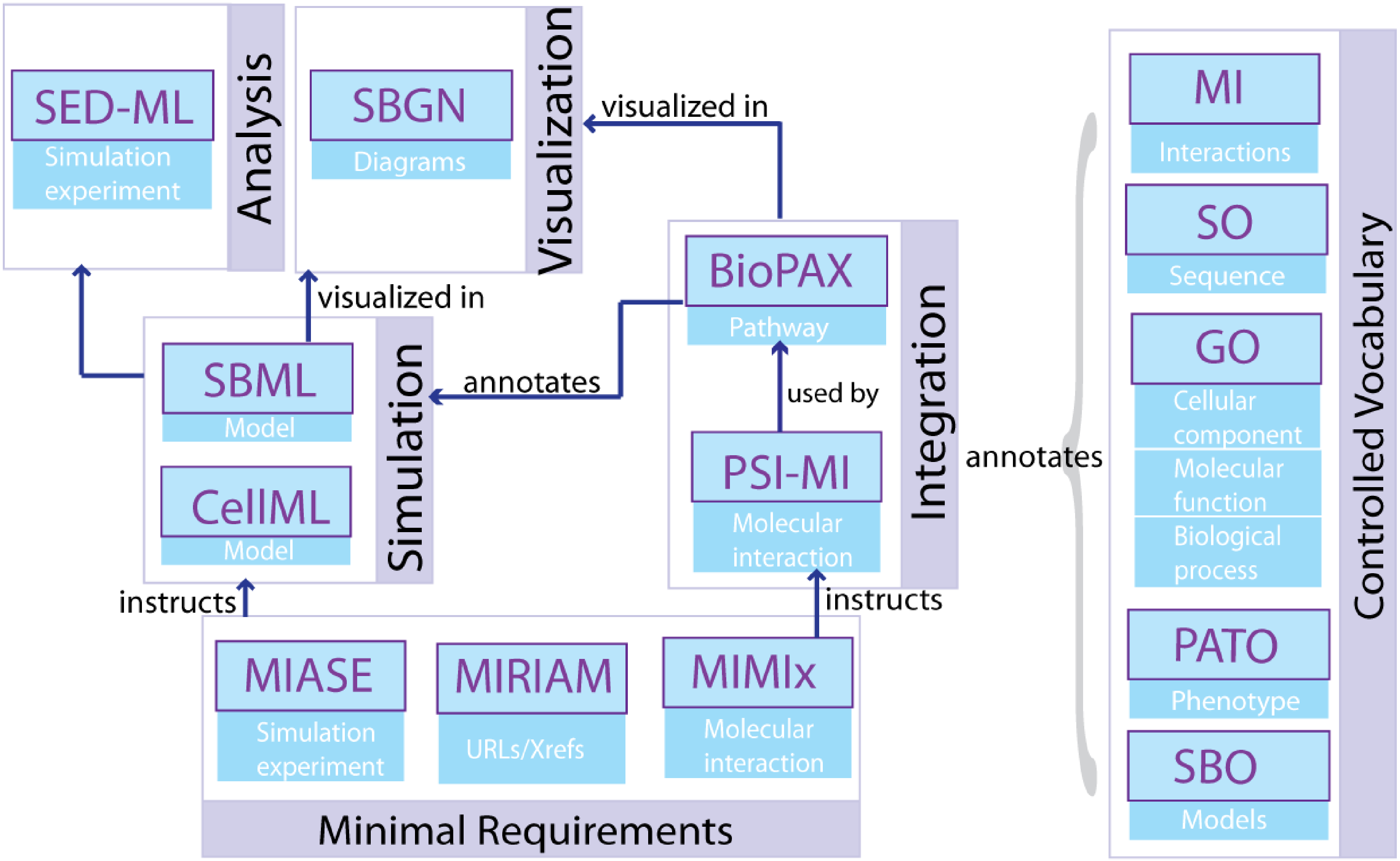
Computable

Publication about a biological process

BioPAX ontology

BioPAX record

Software



Analysis

SED-ML
Simulation experiment

Visualization

SBGN
Diagrams

visualized in

Simulation

SBML
Model

CellML
Model

visualized in

annotates

Integration

BioPAX
Pathway

used by

PSI-MI
Molecular interaction

instructs

instructs

Minimal Requirements

MIASE
Simulation experiment

MIRIAM
URLs/Xrefs

MIMIx
Molecular interaction

annotates

Controlled Vocabulary

MI
Interactions

SO
Sequence

GO
Cellular component
Molecular function
Biological process

PATO
Phenotype

SBO
Models

combine

When to use BioPAX?

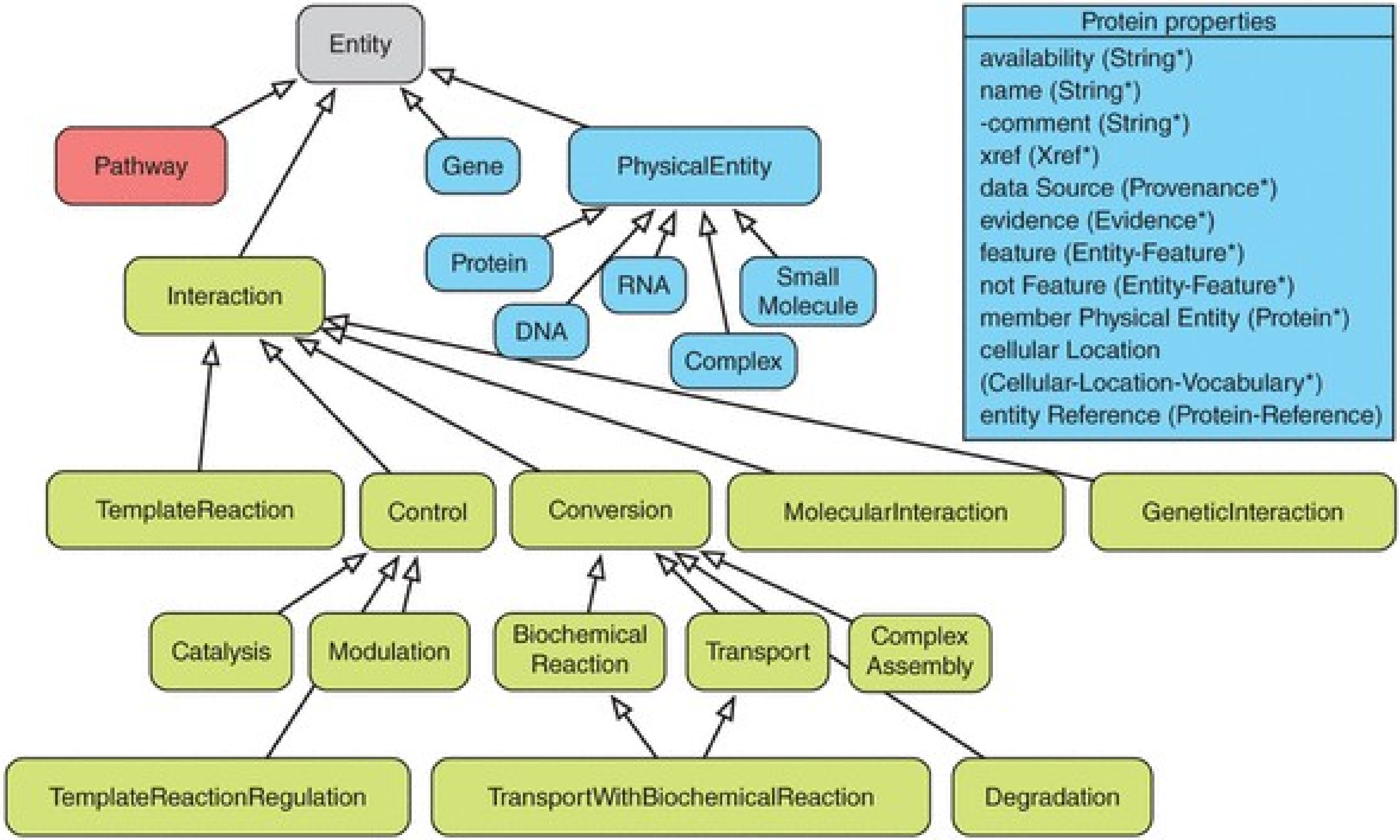
- Information Extraction/Retrieval
  - Pathway Curation, Literature Mining, Finding relevant pathways from high throughput data
- Reporting High Throughput Experiments.
  - ChipSeq, Genetic and Protein-Protein Interactions, high throughput PTM, drug-target interactions.
- Qualitative Network Algorithms : Active Pathways, Paradigm, GEM, MINDY

BioPAX Syntax and Semantics

- Specification defined as a set of OWL classes and properties.
  - http://www.biopax.org/specification.php
- Semantics occasionally differ from the OWL semantics.
  - e.g. Closed world reactions
- Additional restrictions are checked by the BioPAX validator.

What is covered?

- Metabolic Pathways
  - Reactions, reactants, EC numbers, enyzmes, equilibrium constants, co-factors and modulators
- Molecular Interactions (PSI-MI)
- Signaling Pathways
  - PTMs and other features, protein families, cellular location, transport, degradation, complex assembly.
- Gene Regulation
  - Transcription and Translation, DNA “regions” (promoter, enhancer etc.), TF binding, operon.
- Genetic Interactions
  - Interacting partners and resulting phenotype


Entity

Pathway

Gene

PhysicalEntity

Protein

DNA

RNA

Complex

Small Molecule

Interaction

TemplateReaction

Control

Conversion

MolecularInteraction

GeneticInteraction

Catalysis

Modulation

Biochemical Reaction

Transport

Complex Assembly

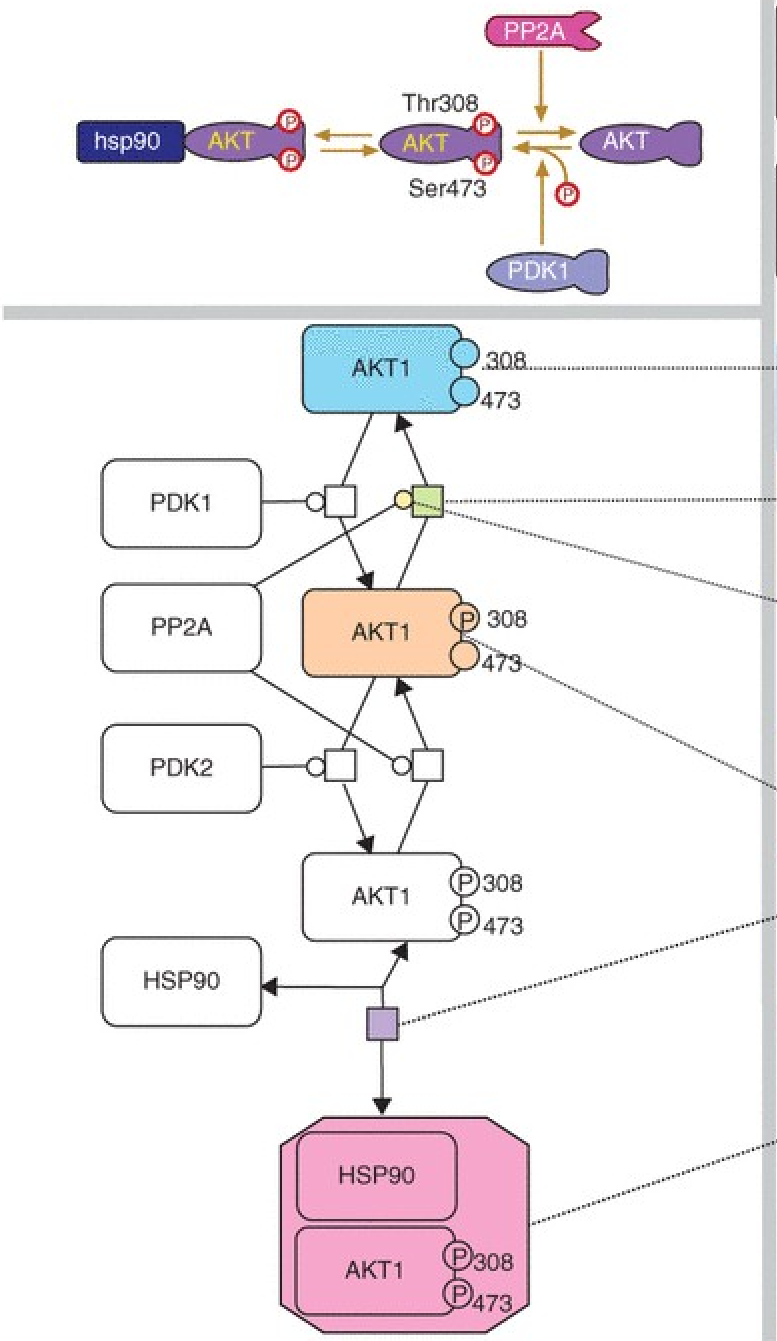
TemplateReactionRegulation

TransportWithBiochemicalReaction

Degradation

Protein properties

availability (String*)
name (String*)
-comment (String*)
xref (Xref*)
data Source (Provenance*)
evidence (Evidence*)
feature (Entity-Feature*)
not Feature (Entity-Feature*)
member Physical Entity (Protein*)
cellular Location (Cellular-Location-Vocabulary*)
entity Reference (Protein-Reference)



rAKT1 is a *ProteinReference*
has *standard-name* “AKT1”
has *name* “PKB”
has *xref* Uniprot-P31749

p@308 is a *ModificationFeature*
has *featureLocation* AKT1-308
has *modificationType*
phosphorylation

AKT1.1 is a *Protein*
has *proteinReference* rAKT1
has *notFeature* p@308
has *notFeature* p@473

reaction1 is a *BiochemicalReaction*
has *left* AKT1.2
has *right* AKT1.1
is *left-to-right.*

catalysis1 is a *Catalysis*
has *controller* PP2A.1
has *controlled* reaction1
has *direction irr-left-to-right*

AKT1.2 is a *Protein*
has *proteinReference* rAKT1
has *feature* p@308
has *notFeature* p@473

assembly1 is a *ComplexAssembly*
has *left* HSP90.1
has *left* AKT1.3
has *right* complex1
is *reversible*

complex1 is a *Complex*
has *component* AKT1.4
has *component* HSP90.2

HSP90.2 is a *Protein*
has *proteinReference* rHSP90
is *boundTo* AKT1.4

AKT1.4 is a *Protein*
has *proteinReference* rAKT1
has *feature* p@308
has *feature* p@473
is *boundTo* HSP90.2

Common features

- Links to external resources
- Evidence
- Provenance
- Integration with controlled vocabularies
  - GO, PSI-MOD, SO etc.

Who produces BioPAX?

Database	**Type, Download format**	**Number of curated records***
HumanCyc	Metabolic, BioPAX 3	260 small to medium predicted pathways
MSKCC Cancer Cell Map	Signaling, BioPAX 2	10 large curated pathways
Reactome	Metabolic and signaling, BioPAX 3	1,113 small to medium curated pathways
NCI/Nature PID	Signaling, BioPAX 3	186 medium curated pathways
IntAct	Protein interactions, PSI-MI 2.5	157,000 interactions
MINT	Protein interactions, PSI-MI 2.5	121,800 interactions
HPRD	Protein interactions, PSI-MI 2.5	40,600 interactions
MetaCyc	Metabolic, BioPAX 3	1,867 medium curated pathways

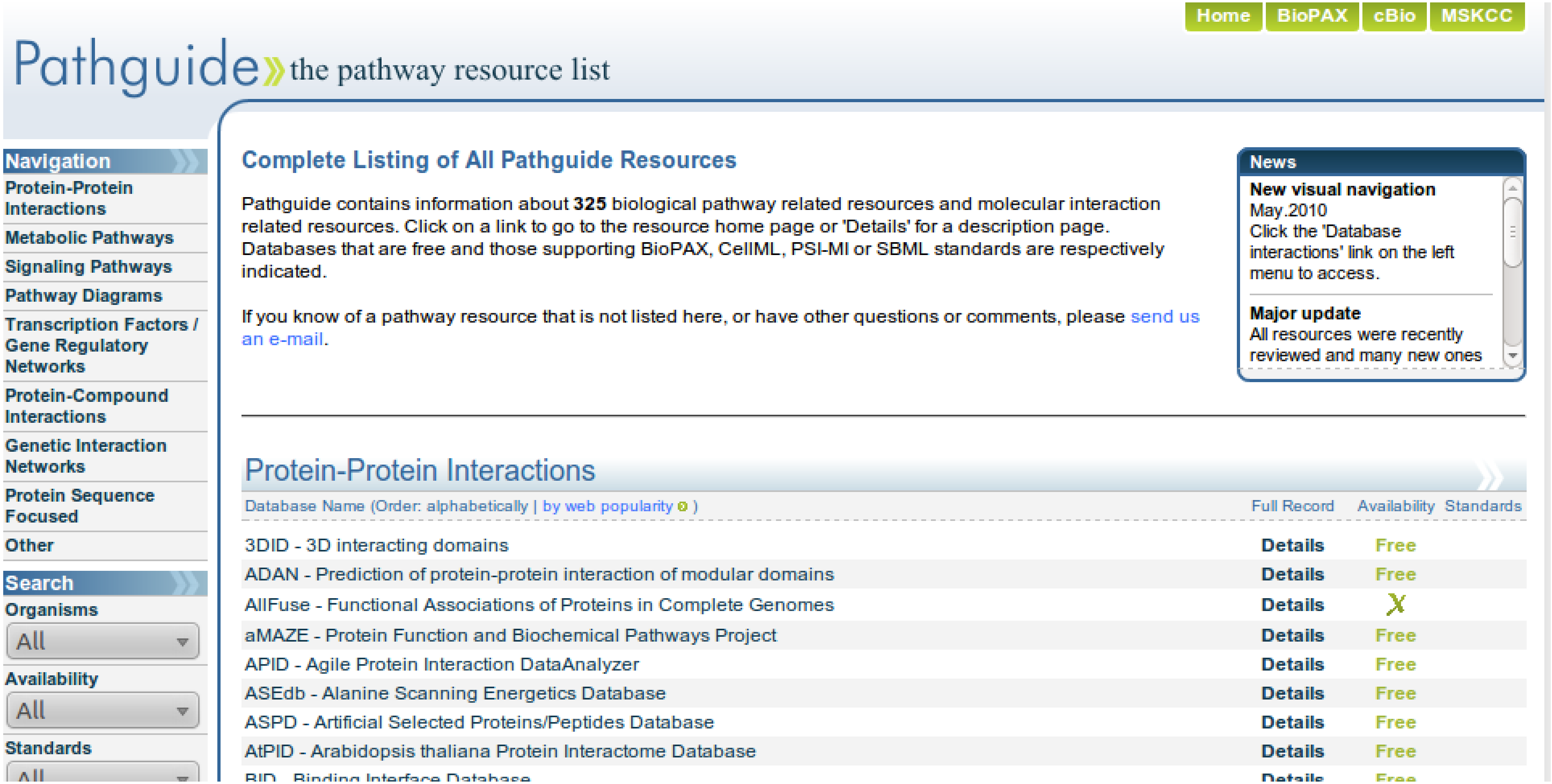
Who produces BioPAX? --cntd

BioGRID	Protein and genetic interaction, PSI-MI 2.5	192,000 interactions
SBCNY IMID	Signaling, BioPAX 2	1,700 interactions
EcoCyc	Metabolic and signaling, BioPAX 3	276 medium curated pathways
NetPath	Signaling, BioPAX 3	10 large curated pathways
BioModels	Metabolic and signaling, BioPAX 2	269 curated pathways
Panther	Signaling, BioPAX 3	165 curated medium pathways
Nature MoleculePages	Signaling, BioPAX 3	655 molecules with states
RegulonDB	Gene regulation, BioPAX 3	2644 regulatory interactions
MouseCyc	Metabolic, BioPAX 3	347 medium curated pathways
YeastCyc	Metabolic, BioPAX 3	200 medium curated pathways

Who produces BioPAX? --cntd

BIND (86)	Molecular, genetic interaction, PSI-MI 2.5	200,000 interactions
DIP (87, 88)	Protein interactions, PSI-MI 2.5	71,000 interactions
PhosphoSite	Signaling, BioPAX 3	>2,000 kinase-substrate interactions
INOH (89)	Signaling, BioPAX 3	101 medium curated pathways
SPIKE	Signaling, BioPAX 3	26 medium curated pathways
PharmGKB	Signaling, BioPAX 3	72 medium curated pathways
~200 BioCycs (16, 90)	Metabolic and signaling, BioPAX 3	Mostly predicted
MIPS CYGD (91)	Protein interactions, PSI-MI 2.5	Statistics unavailable
I2D (92)	Protein interaction, PSI-MI 2.5	Predicted
WikiPathways	Diverse, BioPAX 3	1615 pathways

More at Pathguide.org



Complete Listing of All Pathguide Resources

Pathguide contains information about **325** biological pathway related resources and molecular interaction related resources. Click on a link to go to the resource home page or 'Details' for a description page. Databases that are free and those supporting BioPAX, CellML, PSI-MI or SBML standards are respectively indicated.

If you know of a pathway resource that is not listed here, or have other questions or comments, please send us an e-mail.

News

New visual navigation
May.2010
Click the 'Database interactions' link on the left menu to access.

Major update
All resources were recently reviewed and many new ones

Protein-Protein Interactions

Database Name (Order: alphabetically | by web popularity)	Full Record	Availability	Standards
3DID - 3D interacting domains	**Details**	Free	
ADAN - Prediction of protein-protein interaction of modular domains	**Details**	Free	
AllFuse - Functional Associations of Proteins in Complete Genomes	**Details**	X	
aMAZE - Protein Function and Biochemical Pathways Project	**Details**	Free	
APID - Agile Protein Interaction DataAnalyzer	**Details**	Free	
ASEdb - Alanine Scanning Energetics Database	**Details**	Free	
ASPD - Artificial Selected Proteins/Peptides Database	**Details**	Free	
AtPID - Arabidopsis thaliana Protein Interactome Database	**Details**	Free	
BID - Binding Interface Database	**Details**	Free	

What can I do with BioPAX?

- In which pathways does p53 play a role?
- How is the serine 17 phosphorylation of p53 controlled?
- Given a gene list, which pathways are overrepresented?
- Which pathways map to the Gene Ontology term ‘apoptosis’?
- How are B-Catenin and p53 is connected?
- Which network regions in yeast are transcriptionally active?
- How is a given set of somatically altered genes in colon cancer connected?

Advanced use-cases

- Find causal links between a given a set of genomic changes and matched proteomic profiles.
- Given metabolomic profiles detect differentially active pathways.
- Auto-generate orthologous pathways
- Rule-based modeling

Where do I start?

- OWL/Sparql – Good for querying, inference & data transformation.
- Paxtools Java Library – Full programmatic API – good for complex algorithms and embedding into applications. Also good for transforming Paxtools to simpler formats.

Choice is yours! We are happy to help, just ask!

Get Involved

- Visit biopax.org and join the discussion mailing list

  biopax-discuss@googlegroups.com
- Make pathway data available in BioPAX
- Build software that supports BioPAX
- Contribute BioPAX worked examples, documentation and specification reviews
- Spread the word about BioPAX
- Participate in annual Combine & Harmony meetings

Acknowledgements

- BioPAX community!
- Gary Bader
- Nadia Anwar
- Igor Rodchenkov
- Özgün Babur
- Chris Sander
- Harmony Organizers
- Mike Hucka
- Nicolas Le Novère

Funding: NIH/NIGMS