



PathText



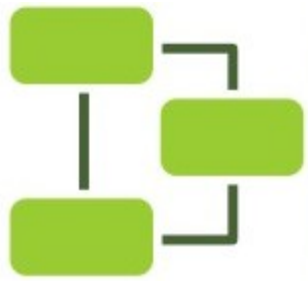
Tsujii Laboratory



NaCTeM
The National Centre for Text Mining



PAYAO
Okinawa Institute of Science and Technology
Community Tagging System to SBML models
Kitano, Matsuoka, Kikuchi, and
other Payao team members.

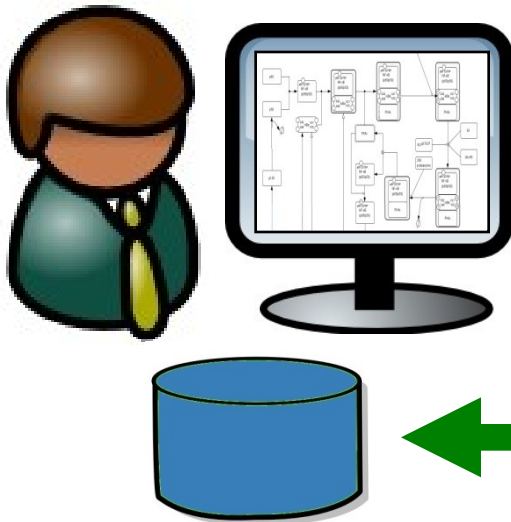


PathText

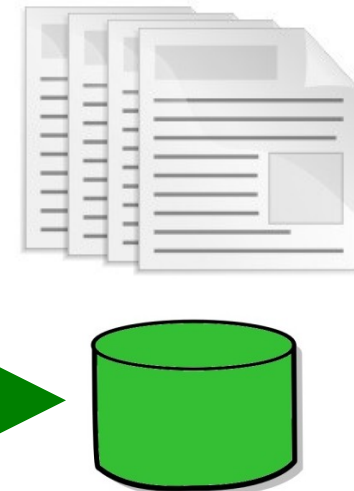
The Goal

To discover publications relevant to entities in a pathway model using a pathway model visualization for a user interface and text mining tools for knowledge discovery

Pathway Visualization



Publications





Publication Counts Demonstration

PathText: **Publication Counts** **View Publications** MaxNumber: Color:

Tag Comments Only **Search** **Advanced Search** Login User: brian **Login Info** **Export** **ChangePassword** **Back To Home** **Logout**

Tag Info **Model Info**

brian's Tag Sets **add tagset**

Other Tag Sets

Disabled tags

TagTitle	TagSet	PubMedIds

NFkB_for_PathText_project

The diagram illustrates the NF-κB signaling pathway. It starts with the activation of IκBα, which is phosphorylated at S32, S36, K21, and K22. This phosphorylated IκBα then binds to p65 S278P NF-κB (p65/p50). The complex is then ubiquitinated and degraded by the 26S proteasome. The released p65 S278P NF-κB (p65/p50) then translocates to the nucleus, where it is phosphorylated at S32 and S36, and acetylated by PKAc. The acetylated p65 S278P NF-κB (p65/p50) then binds to SCFβ-TrCp, leading to the ubiquitination and degradation of the complex. The diagram also shows the phosphorylation of p65 S278P NF-κB (p65/p50) at S32 and S36, and the ubiquitination of the complex by the 26S proteasome.

First, the user logs in to Payao (or another compatible pathway visualization) and selects a “PathText enabled” model.

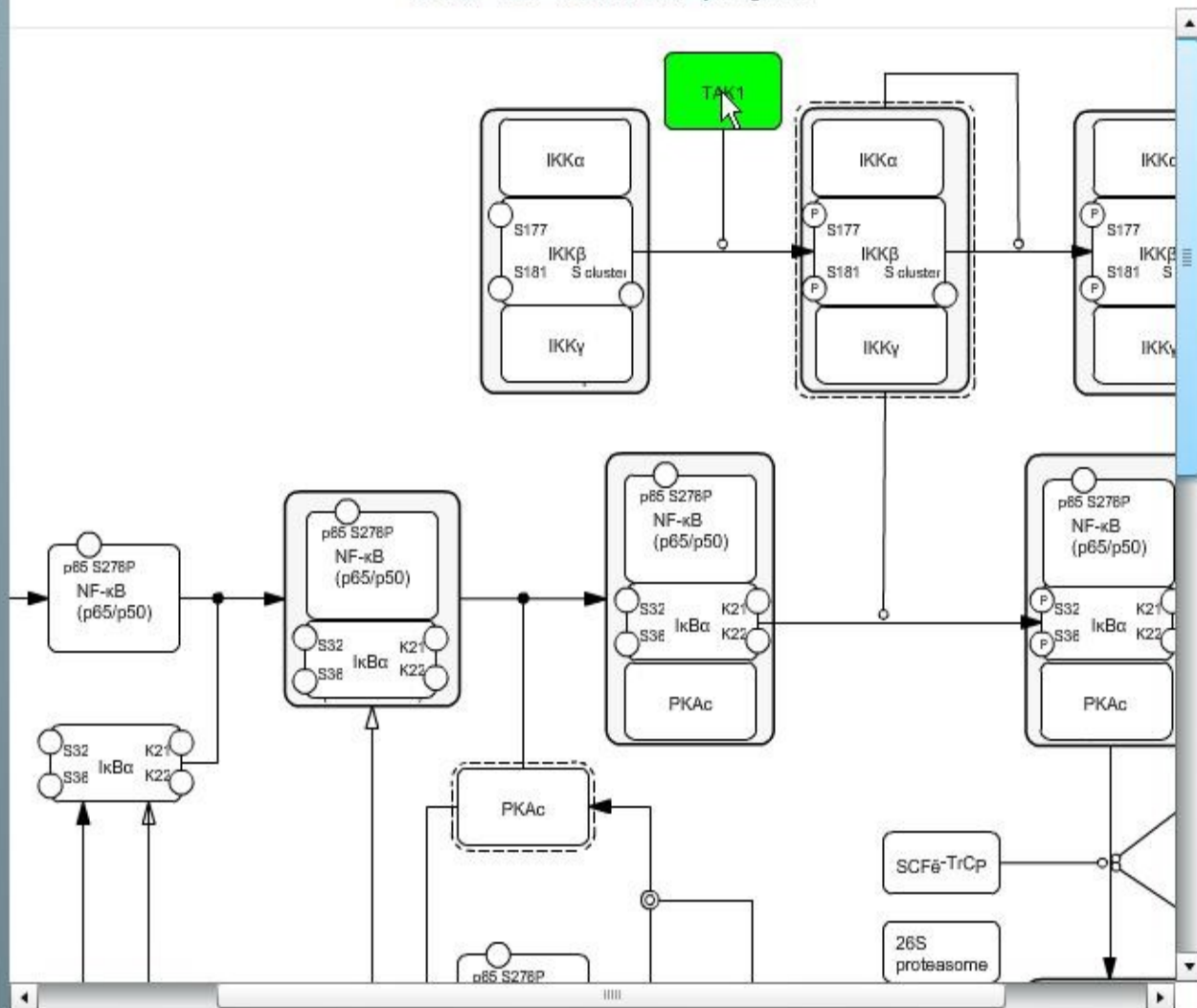
brian's Tag Sets add tagset

Other Tag Sets

Disabled tags

TagTitle	TagSet	PubMedIds
----------	--------	-----------

NFkB_for_PathText_project



The user then selects one or multiple nodes from the model, in this case “TAK1”

PathText: Publication Counts View Publications MaxNumber: 200 Color: 200

Tag Comments Only Search Advanced Search Login User: brian Login Info Export ChangePassword Back To Home Logout

Tag Info Model Info

brian's Tag Sets add tagset

Other Tag Sets

Disabled tags

TagTitle	TagSet	PubMedIds
----------	--------	-----------

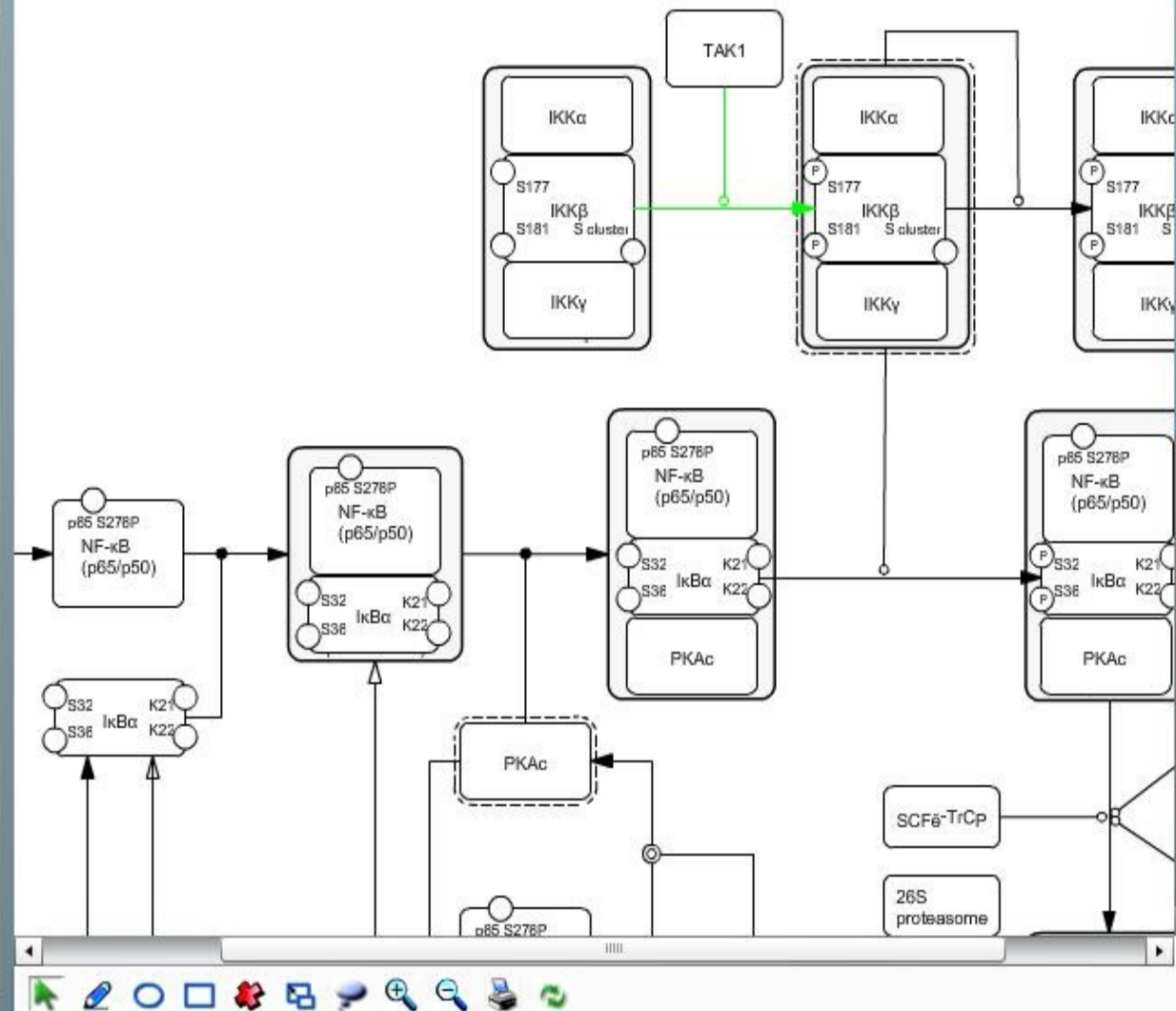
NFkB_for_PathText_project

The user then presses the “Publication Counts” button and a request is sent to PathText. A list of Nodes with publications co-occurring with the user’s selection is returned. These results are displayed as document icons color coded according to the number of publications discovered.



View Publications Demonstration

NFkB_for_PathText_project



First, a user selects one or multiple nodes in the pathway, in this case a reaction describing the activation/phosphorylation of a complex with TAK1 as a modifier. The user then presses the “View Publications” button

Positive and negative regulation of IkappaB kinase activity through IKKbeta subunit phosphorylation.

In mammalian cells, phosphorylation of two sites at the activation loop of IKKb was essential for activation of IKK by tumor necrosis factor and interleukin-1. [View Annotation](#)

Whereas the Ser177-> Ala177 (S177A) mutation in IKKb slightly decreased cytokine responsiveness, the S181A mutation had a more severe effect, and the replacement of both sites abolished IKK activation altogether. [View Annotation](#)

Science, 1999 Apr 9, 284(5412):309-13, PMID: 10195894

Delhase M, Hayakawa M, Chen Y, Karin M.

Laboratory of Gene Regulation and Signal Transduction, Department of Pharmacology, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0636, USA.

Annotation by K. Oda

TAK1 is a ubiquitin-dependent kinase of MKK and IKK

Significantly, the TAK1 kinase once activated by Ubc13-Uev1A-mediated ubiquitination was able to phosphorylate IKKb specifically at S177 and S181. Thus, TAK1 is a ubiquitin-dependent kinase of IKKb. [View Annotation](#)

Nature, 2001 Jul 19, 412(6844):346-51, PMID: 11460167

Wang C, Deng L, Hong M, Akkaraju GR, Inoue J, Chen ZJ

Department of Molecular Biology, University of Texas Southwestern Medical Center, Dallas, Texas 75390-9148, USA

Annotation by K. Oda

Mammalian TAK1 activates Snf1 protein kinase in yeast and phosphorylates AMP-activated protein kinase in vitro.

Mammalian TAK1 activates Snf1 protein kinase in yeast and phosphorylates AMP-activated protein kinase in vitro

The Journal of biological chemistry, 2006 Sep 1, 281(35):25336-43. Epub 2006 Jul 11, PMID: 16835226

Momcilovic M, Hong SP, Carlson M.

Search Term(s): s:TAK1 v:Phosphorylate

[View Publication](#) | [Create Annotation](#) | [Mark "Not Relevant"](#)

Constitutive activation of TAK1 by HTLV-1 tax-dependent overexpression of TAB2 induces activation of JNK-ATF2 but not IKK-NF-kappaB.

... Transforming growth factor-beta-activated kinase 1 (TAK1) has been shown to play a critical role in these transcription factors. Here, we found that TAK1 was constitutively activated in Tax-positive HTLV-1-transformed T cells. Tax induced persistent overexpression of TAK1-binding protein 2 (TAB2), but not TAB3, which is essential for TAK1 activation. Surprisingly, TAK1 was not involved in the activation of NF-kappaB. On the other hand, JNK and p38 mitogen-activated protein kinases were activated by TAK1. In addition, ATF2, but not CREB, was a target for the TAK1-JNK pathway, and p38 negatively regulated TAK1 activity through TAB1 phosphorylation. These results indicate that Tax-mediated TAK1 activation is important for the activation of ATF2 rather than NF-kappaB.

The Journal of Biological Chemistry, 2007 Aug 31, 282(35):25177-81. Epub 2007 Jul 11, PMID: 17626013

Suzuki S, Singhirunusorn P, Mori A, Yamaoka S, Kitajima I, Saiki I, Sakurai H.

Search Term(s): TAK1

[View Publication](#) | [Create Annotation](#) | [Mark "Not Relevant"](#)

A new window (or tab) is then opened in the user's browser and the PathText results are displayed.

Positive and negative regulation of IkappaB kinase activity through IKKbeta subunit phosphorylation.

In mammalian cells, phosphorylation of two sites at the activation loop of IKKb was essential for activation of IKK by tumor necrosis factor and interleukin-1. [View Annotation](#)

Whereas the Ser177-> Ala177 (S177A) mutation in IKKb slightly decreased cytokine responsiveness, the S181A mutation had a more severe effect, and the replacement of both sites abolished IKK activation altogether. [View Annotation](#)

Science, 1999 Apr 9, 284(5412):309-13, PMID: 10195894

Delhase M, Hayakawa M, Chen Y, Karin M.

Laboratory of Gene Regulation and Signal Transduction, Department of Pharmacology, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0636, USA.

Annotation by K. Oda

TAK1 is a ubiquitin-dependent kinase of MKK and IKK

Significantly, the TAK1 kinase once activated by Ubc13-Uev1A mediated ubiquitination was able to phosphorylate IKKb specifically at S177 and S181. Thus, TAK1 is a ubiquitin-dependent kinase of IKKb. [View Annotation](#)

Nature, 2001 Jul 19, 412(6844):346-51, PMID: 11460167

Wang C, Deng L, Hong M, Akkaraju GR, Inoue J, Chen ZJ

Department of Molecular Biology, University of Texas Southwestern Medical Center, Dallas, Texas 75390-9148, USA

Annotation by K. Oda

Mammalian TAK1 activates Snf1 protein kinase in yeast and phosphorylates AMP-activated protein kinase in vitro.

Mammalian **TAK1** activates Snf1 protein kinase in yeast and phosphorylates AMP-activated protein kinase in vitro

The Jour

Momcilov

Search T

Constitu

IKK-NF-

... Transf

that TAK1

2 (TAB2)

hand, JNK

pathway,

important

for the activation

of ATF2 rather

than NF-kappaB.

Search Term(s): TAK1

The first two results are publications that were annotated manually by a biologist using the PathText annotation tool. The tool assists the user in linking a specific node in a pathway model with annotations in a publication. The information is then stored in the PathText annotation repository. This repository is then queried by the PathText service and these results are generated.

The Journal of Biological Chemistry, 2007 Aug 31, 282(35):25177-81. Epub 2007 Jul 11, PMID: 17626013

Suzuki S, Singhirunnosorn P, Mori A, Yamaoka S, Kitajima I, Saiki I, Sakurai H.

Search Term(s): TAK1

The next result comes from a “subject verb object” query generated by PathText and processed by a web service, in this case Medie.

The “View Publication” link will display the original document, such as the PDF file on the PubMed website.

The “Create Annotation” link will take the user to the PathText annotation system which facilitates creating a link between the publication and a node or nodes in a pathway model.

The “Mark Not Relevant” link allows the user to mark the search result as “Not Relevant”. This data will be stored by the PathText system and used to generate more accurate search results for future users.

Wang C, Deng L, Hong M, Akkaraju GK, Inoue J, Chen ZJ
Department of Molecular Biology, University of Texas Southwestern Medical Center, Dallas, Texas 75390-9148, USA
Annotation by K. Oda

Mammalian TAK1 activates Snf1 protein kinase in yeast and phosphorylates AMP-activated protein kinase in vitro.

Mammalian **TAK1** activates Snf1 protein kinase in yeast and phosphorylates AMP-activated protein kinase in vitro

The Journal of biological chemistry, 2006 Sep 1, 281(35):25336-43. Epub 2006 Jul 11, PMID: 16835226

Momcilovic M, Hong SP, Carlson M.

Search Term(s): s:TAK1 v:Phosphorylate

[View Publication](#) | [Create Annotation](#) | [Mark "Not Relevant"](#)

Constitutive activation of TAK1 by HTLV-1 tax-dependent overexpression of TAB2 induces activation of JNK-ATF2 but not IKK-NF-kappaB.

... Transforming growth factor-beta-activated kinase 1 (TAK1) has been shown to play a critical role in these transcription factors. Here, we found that TAK1 was constitutively activated in Tax-positive HTLV-1-transformed T cells. Tax induced persistent overexpression of TAK1-binding protein 2 (TAB2), but not TAB3, which is essential for TAK1 activation. Surprisingly, TAK1 was not involved in the activation of NF-kappaB. On the other hand, JNK and p38 mitogen-activated protein kinases were activated by TAK1. In addition, ATF2, but not CREB, was a target for the TAK1-JNK pathway, and p38 negatively regulated TAK1 activity through TAB1 phosphorylation. These results indicate that Tax-mediated TAK1 activation is important for the activation of ATF2 rather than NF-kappaB.

The Journal of Biological Chemistry, 2007 Aug 31, 282(35):25177-81. Epub 2007 Jul 11, PMID: 17626013

Suzuki S, Singhirunusorn P, Mori A, Yamaoka S, Kitajima I, Saiki I, Sakurai H.

Search Term(s): TAK1

[View Publication](#) | [Create Annotation](#) | [Mark "Not Relevant"](#)

Positive and negative regulation of IkappaB kinase activity through IKKbeta subunit phosphorylation.

In mammalian cells, phosphorylation of two sites at the activation loop of IKKb was essential for activation of IKK by tumor necrosis factor and interleukin-1. [View Annotation](#)

Whereas the Ser177-> Ala177 (S177A) mutation in IKKb slightly decreased cytokine responsiveness, the S181A mutation had a more severe effect, and the replacement of both sites abolished IKK activation altogether. [View Annotation](#)

Science, 1999 Apr 9, 284(5412):309-13, PMID: 10195894

Delhase M, Hayakawa M, Chen Y, Karin M.

Laboratory of Gene Regulation and Signal Transduction, Department of Pharmacology, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0636, USA.

Annotation by K. Oda

TAK1 is a ubiquitin-dependent kinase of MKK and IKK

Significantly, the TAK1 kinase once activated by Ubc13-Uev1A-mediated ubiquitination was able to phosphorylate IKKb specifically at S177 and

The final result comes from a web service that processes space separated query terms, such as Facta and Kleio.

As with the previous result, the user can view the publication, generate annotations or mark the result as "Not Relevant"

The Journal of biological chemistry, 2006 Sep 1, 281(35):25336-43. Epub 2006 Jul 11, PMID: 16835226

Momcilovic M, Hong SP, Carlson M.

Search Term(s): s:TAK1 v:Phosphorylate

[View Publication](#) | [Create Annotation](#) | [Mark "Not Relevant"](#)

Constitutive activation of TAK1 by HTLV-1 tax-dependent overexpression of TAB2 induces activation of JNK-ATF2 but not IKK-NF-kappaB.

... Transforming growth factor-beta-activated kinase 1 (TAK1) has been shown to play a critical role in these transcription factors. Here, we found that TAK1 was constitutively activated in Tax-positive HTLV-1-transformed T cells. Tax induced persistent overexpression of TAK1-binding protein 2 (TAB2), but not TAB3, which is essential for TAK1 activation. Surprisingly, TAK1 was not involved in the activation of NF-kappaB. On the other hand, JNK and p38 mitogen-activated protein kinases were activated by TAK1. In addition, ATF2, but not CREB, was a target for the TAK1-JNK pathway, and p38 negatively regulated TAK1 activity through TAB1 phosphorylation. These results indicate that Tax-mediated TAK1 activation is important for the activation of ATF2 rather than NF-kappaB.

The Journal of Biological Chemistry, 2007 Aug 31, 282(35):25177-81. Epub 2007 Jul 11, PMID: 17626013

Suzuki S, Singhirunnosorn P, Mori A, Yamaoka S, Kitajima I, Saiki I, Sakurai H.

Search Term(s): TAK1

[View Publication](#) | [Create Annotation](#) | [Mark "Not Relevant"](#)

The logo consists of three green squares arranged in a triangle, connected by thin black lines to form a square-like shape.

PathText

So how does it work?

Probably out of time by now,

So please ask me later:

satre@idi.ntnu.no

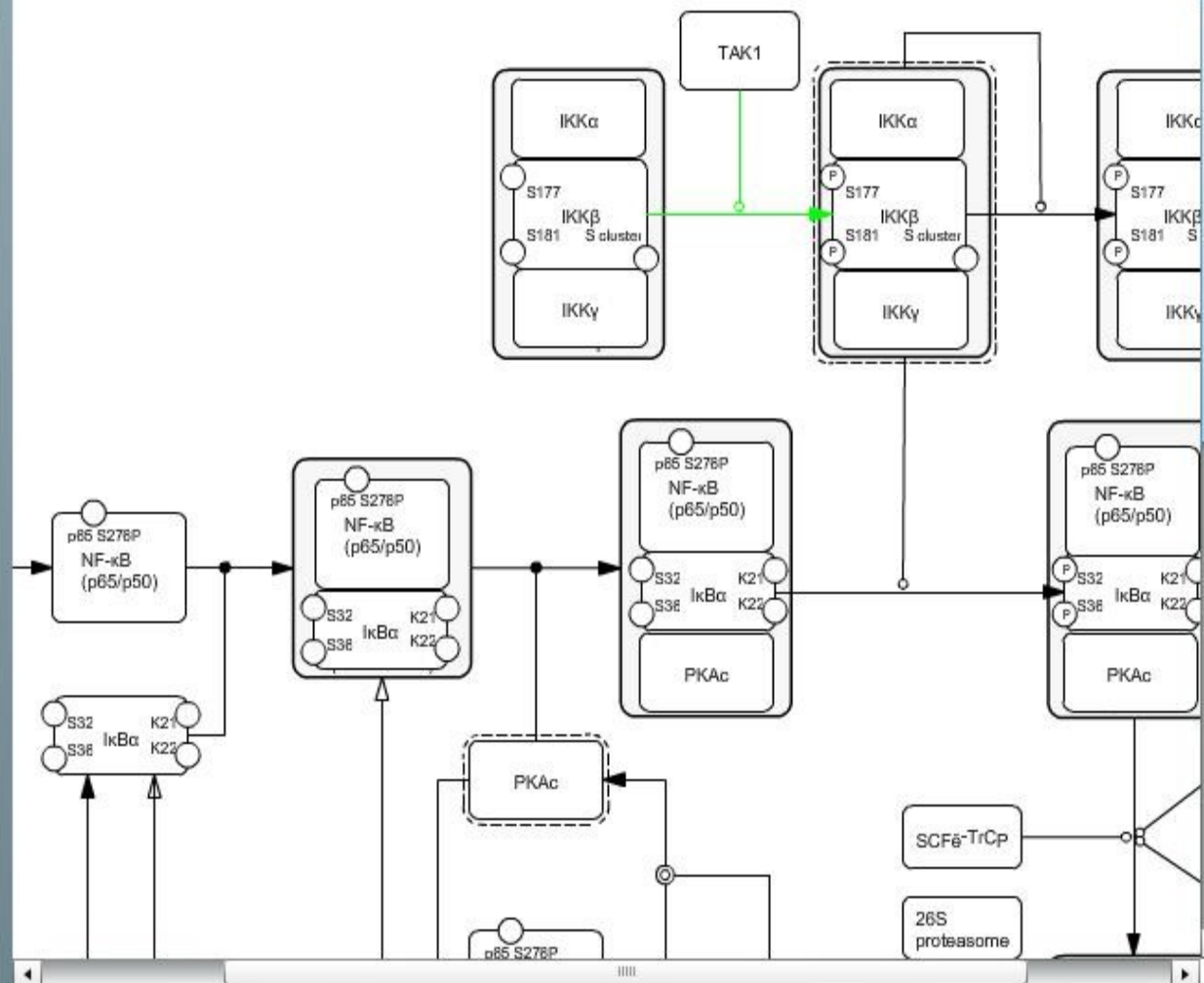
**A request is sent by Payao to the PathText web service.
This request contains the selected model and a unique identifier for each selected node.**

brian's Tag Sets **add tagset**

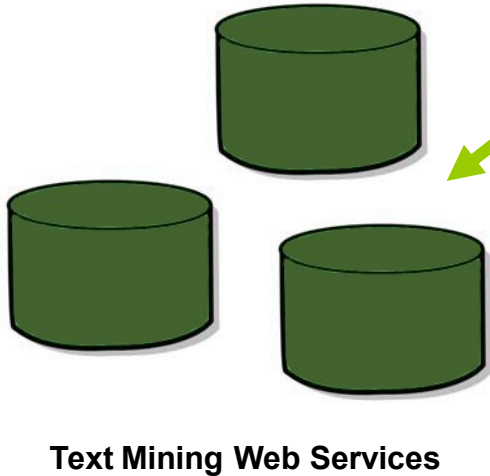
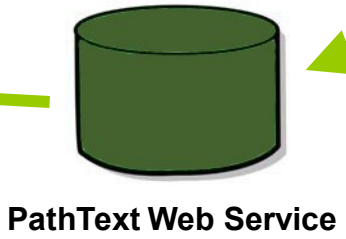
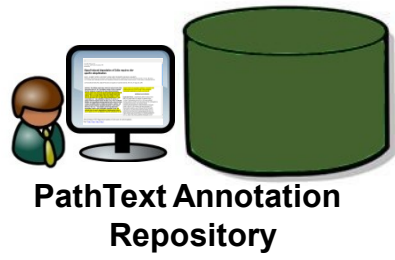
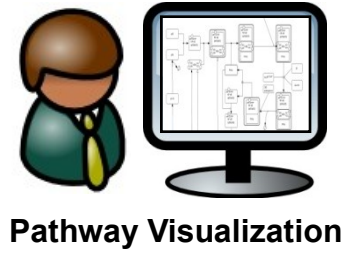
Other Tag Sets

Disabled tags

TagTitle	TagSet	PubMedIds
----------	--------	-----------



PathText

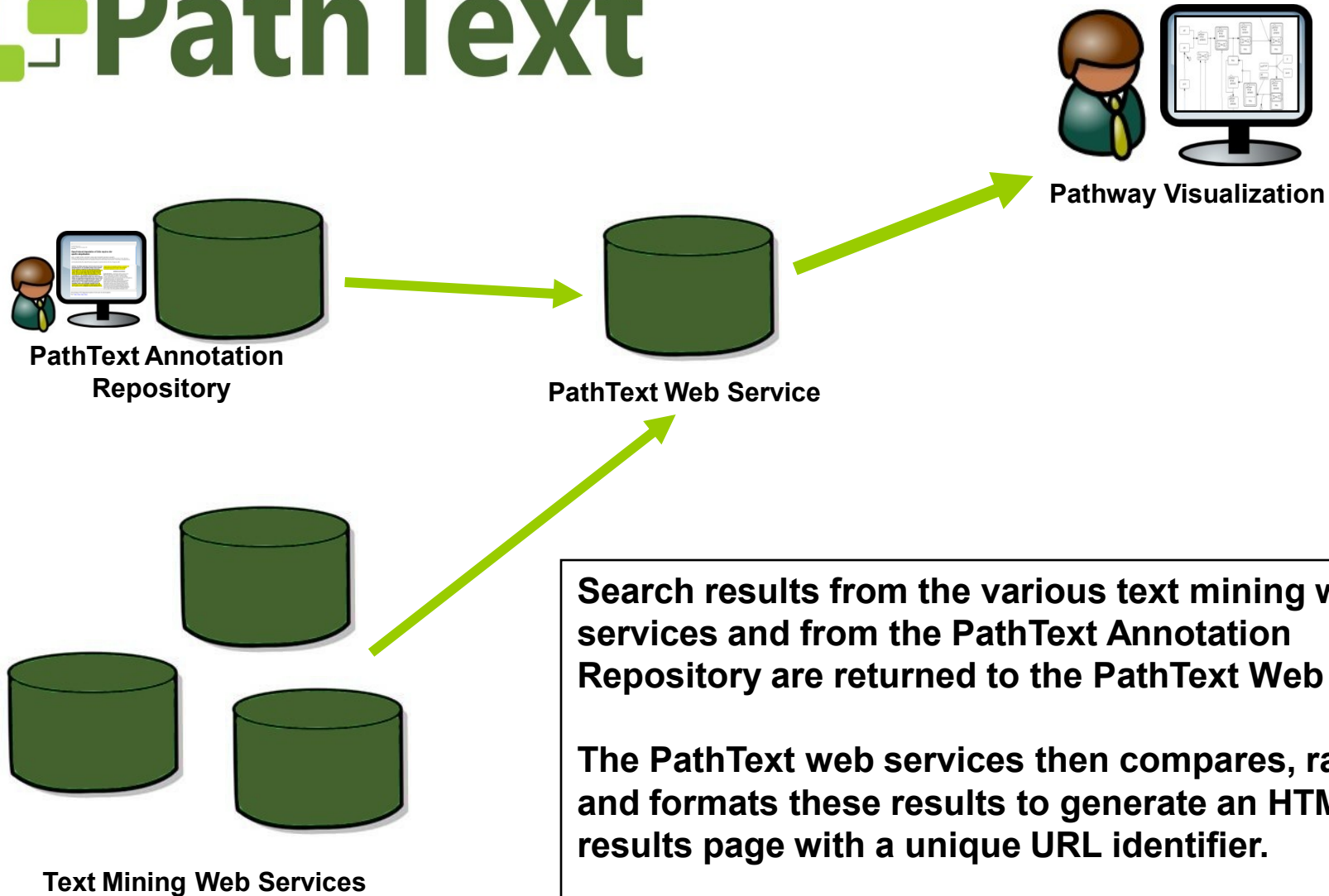


The PathText web service receives the request containing the selected nodes.

This data is then processed using the PathText Model Repository to convert the node ID(s) into queries compatible with various text mining web services.

PathText sends a request for manually generated annotations to the “PathText Annotation Repository”

PathText



Search results from the various text mining web services and from the PathText Annotation Repository are returned to the PathText Web Service.

The PathText web services then compares, ranks and formats these results to generate an HTML results page with a unique URL identifier.

This unique URL is then sent to the Pathway Visualization (Payao) and the result page is displayed in a new window on the user's browser.



PathText API

PathText



Request from Pathway Visualization to PathText web service

URI Request

server_id=(string)
request_id=(string)
request_type=(publications, count)
user_id=(string)
model_id=(string)
start_date=(mmddyyyy)
end_date=(mmddyyyy)
node[]=(string array)
author=(string)
publication=(string)
max_results=(integer)

XML Response (for publications)

```
<pathtext_result>  
<request_id>(string)</request_id>  
<result_url>(url string)</result_url>  
</pathtext_result>
```

XML Response (for count)

```
<pathtext_result>  
<request_id>(string)</request_id>  
<node_count>  
    <node_id>(string)</node_id>  
    <pub_count>(integer)</pub_count>  
</node_count>  
<node_count>  
    <node_id>(string)</node_id>  
    <pub_count>(integer)</pub_count>  
</node_count>  
</pathtext_result>
```

PathText



Text Mining Web Services

List of Associated Proteins/Publication Counts from Uniprot ID(s) or Name String(s)

List of Publications from Uniprot ID(s) or Name String(s)

List of Publications from Subject,Verb,Object Query

(Future Work)

Protein Name to Uniprot ID / Uniprot ID to Protein Name

List of Protein Interactions from Uniprot ID

Protein/Gene Name Normalizer

Protein/Gene Name Synonym Service

PathText



Text Mining Web Services

List of Associated Proteins/Publication Counts from Uniprot ID(s) or Name String(s)

URI Request

query=(string)
cat=(string)

query contains a list of uniprot ids or protein labels separated by a space character (%20) and grouped with parenthesis

cat contains the species to be searched, e.g. human

Here is an example:

?query=(uniprot_p19838%20uniprot_q04206)%20or%20p65&cat=human

XML Response

```
<publication_count>  
<node_count>  
    <type>(string)</type>  
    <id>(string)</id>  
    <label>(string)</label>  
</node_count>  
<node_count>  
    <type>(string)</type>  
    <id>(string)</id>  
    <label>(string)</label>  
</node_count>  
</publication_count>
```

PathText



Text Mining Web Services List of Publications from Uniprot ID(s) or Name String(s)

URI Request

query=(string)

query contains a list of uniprot ids or protein labels separated by a space character (%20) and grouped with parenthesis

Here is an example:

?query=(uniprot_p19838%20uniprot_q04206)%20or%20p65

XML Response

```
<publications>
<publication>
  <pmid>(string)</pmid>
  <pub_url>(string)</pub_url>
  <pub_title>(string)</pub_title>
  <journal>(string)</journal>
  <authors>
    <author>(string)</author>
    <author>(string)</author>
  </authors>
  <excerpt>(string)</excerpt>
</publication>
<publication> ...</publication>
</publications>
```

PathText



Text Mining Web Services List of Publications from Subject, Verb, Object Query

URI Request

subject=(string)
verb=(string)
object=(string)

subject contains a list of uniprot ids or protein labels separated by a space character (%20) and grouped with parenthesis

Here is an example:

? subject=p65%20p50&object=NF-kappaB&verb=bind

XML Response

```
<publications>  
<publication>  
  <pmid>(string)</pmid>  
  <pub_url>(string)</pub_url>  
  <pub_title>(string)</pub_title>  
  <journal>(string)</journal>  
  <authors>  
    <author>(string)</author>  
    <author>(string)</author>  
  </authors>  
  <excerpt>(string)</excerpt>  
</publication>  
<publication> ...</publication>  
</publications>
```


PathText

The Pathway Text Mining Bridge

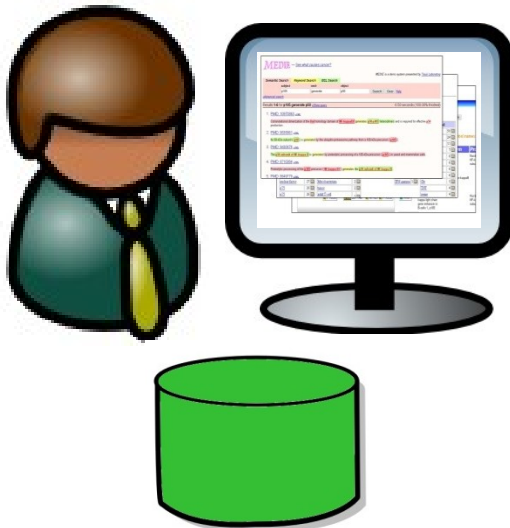
Text Mining Tools

Here I would like to insert a description of the benefits of text mining tools. Something similar to these excerpts from Sophia's "Towards Text Mining Terabytes of Text Documents" presentation:

Text mining is about knowledge discovery from large collections of unstructured text.

The primary goal of text mining is to
–extract knowledge that is hidden in text
–present the distilled knowledge to users in a concise form.

The benefits of text mining are it enables users/scientists to
–collect, maintain, interpret, curate and discover
knowledge needed for research or education efficiently and systematically.



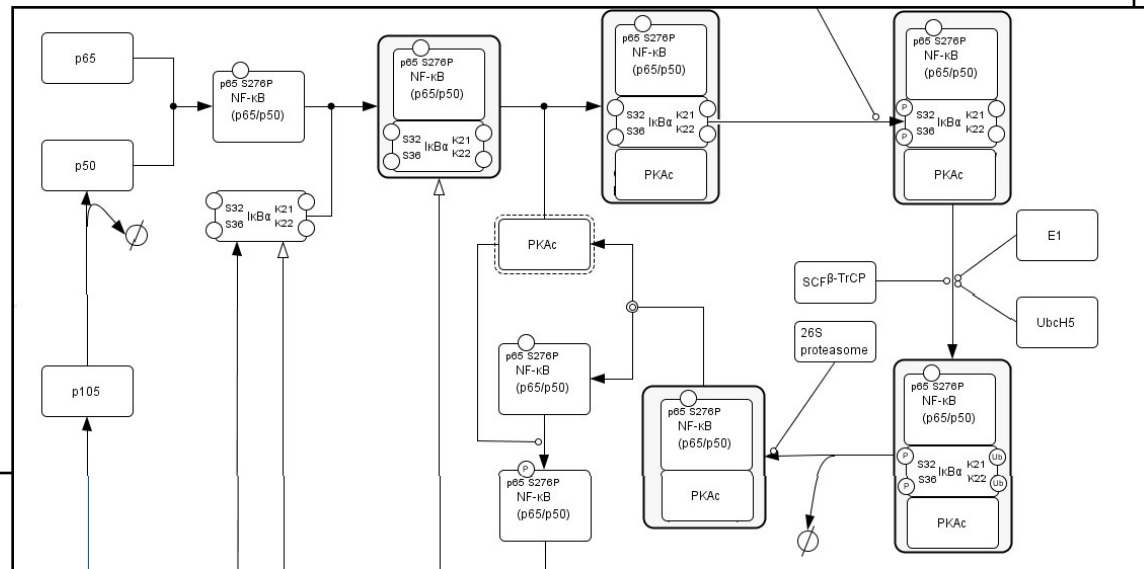
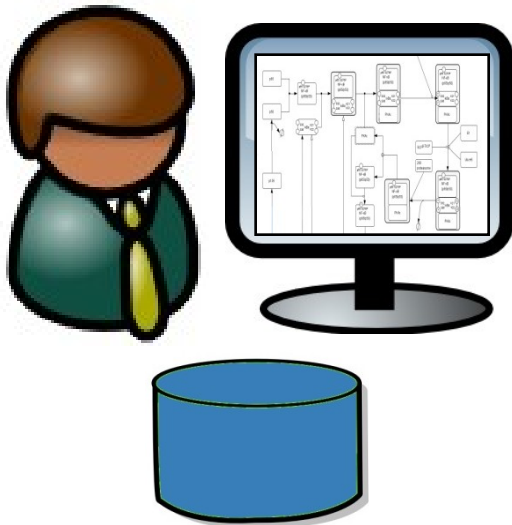
A screenshot of the MEDIE text mining tool interface. The title is "MEDIE - See what causes cancer?". Below the title, there are search options: "Semantic Search", "Keyword Search", and "GCL Search". The "Keyword Search" section has a search bar with "p105 generate p50" entered. Below the search bar, there are search results for "p105 generate p50". The results are numbered 1 to 5, each with a PMID and a brief description of the text snippet. The first result is "PMID: 10970863" and the snippet is "Cotranslational demethylation of the Bcl homology domain of NF-kappa-B1 generates p50 p105 heterodimers and is required for effective p50 production." The second result is "PMID: 9535861" and the snippet is "The 50 kDa subunit (p50) is generated by the ubiquitin-proteasome pathway from a 105-kDa precursor (p105)." The third result is "PMID: 9430676" and the snippet is "The p50 subunit of NF-kappa-B is generated by proteolytic processing of a 105-kDa precursor (p105) in yeast and mammalian cells." The fourth result is "PMID: 8710364" and the snippet is "Proteolytic processing of the p105 precursor (NF-kappa-B1) generates the p50 subunit of NF-kappa-B." The fifth result is "PMID: 9649770" and the snippet is "p105 generate p50". Below the search results, there is a table with columns for "Molecular factor", "Molecular factor", "Molecular factor", "Molecular factor", and "Molecular factor". The table contains several rows of data, including "p50", "p105", and "p50".

Pathway Visualization Tools

Here I would like to insert a description of the usefulness of Visualization tools (SBGN) primarily as a user input method.

Something similar to the goal statement for SBGN:

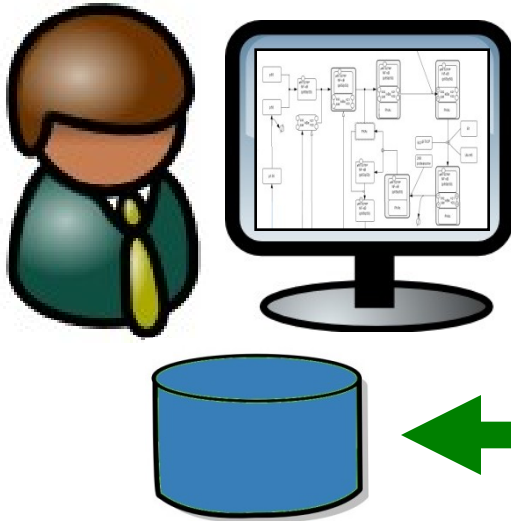
The goal of the SBGN effort is to help standardize a graphical notation for computational models in systems biology. Such a standard notation will have broad impact. For example, it will add rigor and consistency to the usually ad hoc diagrams that often accompany research articles in publications. It will also help bring consistency to the user interfaces of different software tools and databases. SBGN is a natural complement to SBML.



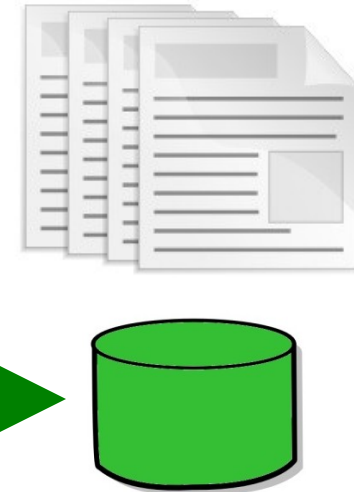
The Goal

To discover publications relevant to entities in a pathway model using a pathway model visualization for a user interface and text mining tools for knowledge discovery

Pathway Visualization



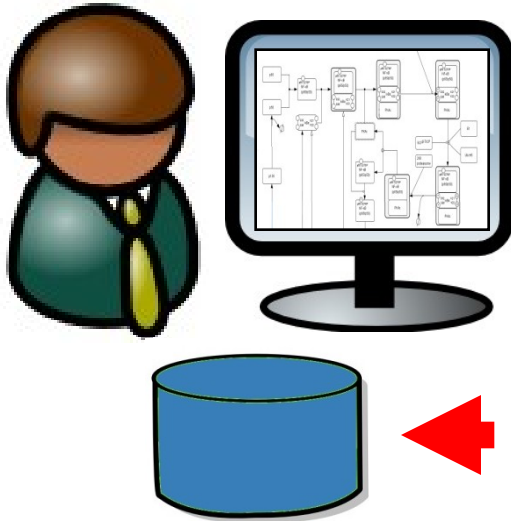
Text Mining Tools



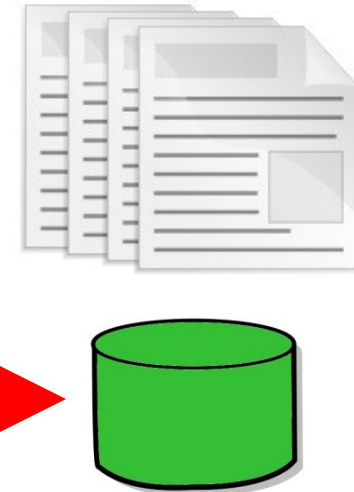
The Problem

Most pathway visualization tools do not contain the information required to communicate with text mining tools directly

Pathway Visualization



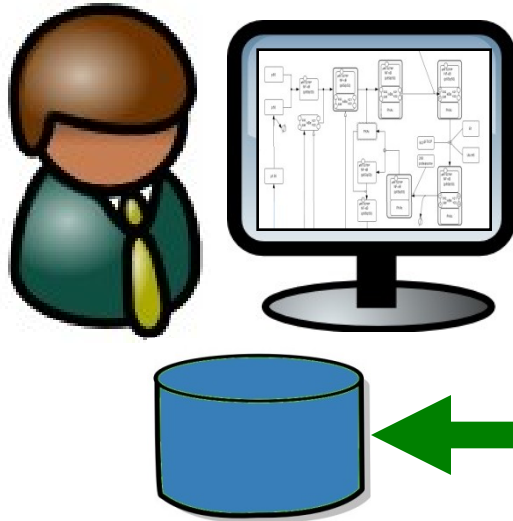
Text Mining Tools



The PathText Solution

The PathText system provides an “association table” that can bridge visualization interfaces and text mining tools for the discovery of relevant publications

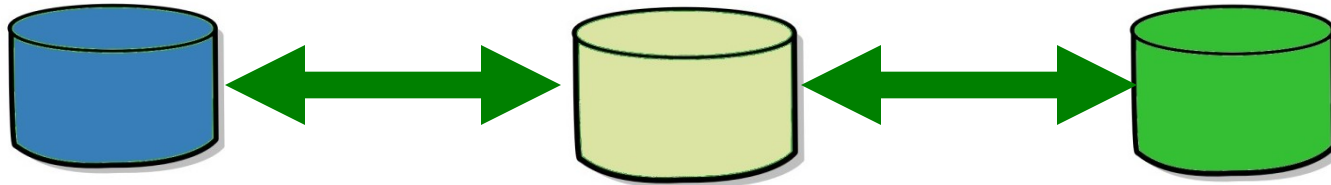
Pathway Visualization



Text Mining Tools



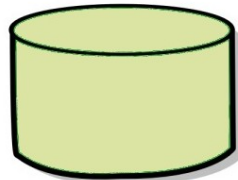
PathText



Step 1: Association Table Generation

Using data contained in the SBML/CellDesigner model description and descriptions generated by the model designer, PathText builds a semantically rich description of all nodes and reactions in the model.

PathText



PathText Association Table



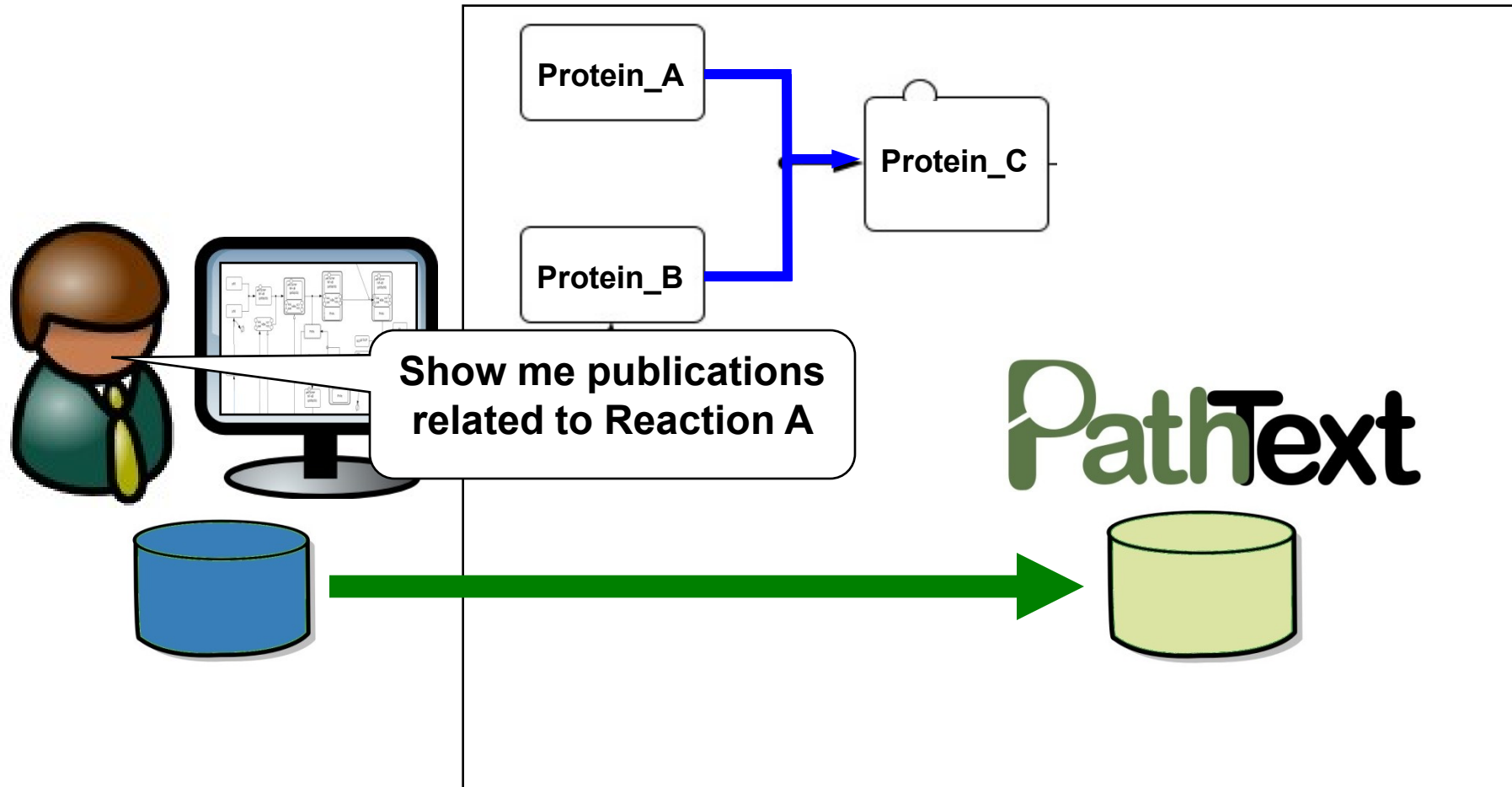
**Computer Assisted
Description Generation**

SBML/CellDesigner Model Data

```
Reaction_A
Reaction Type: Heterodimer Association
Reactants: Protein_A, Protein_B
Products: Protein_C ...
```

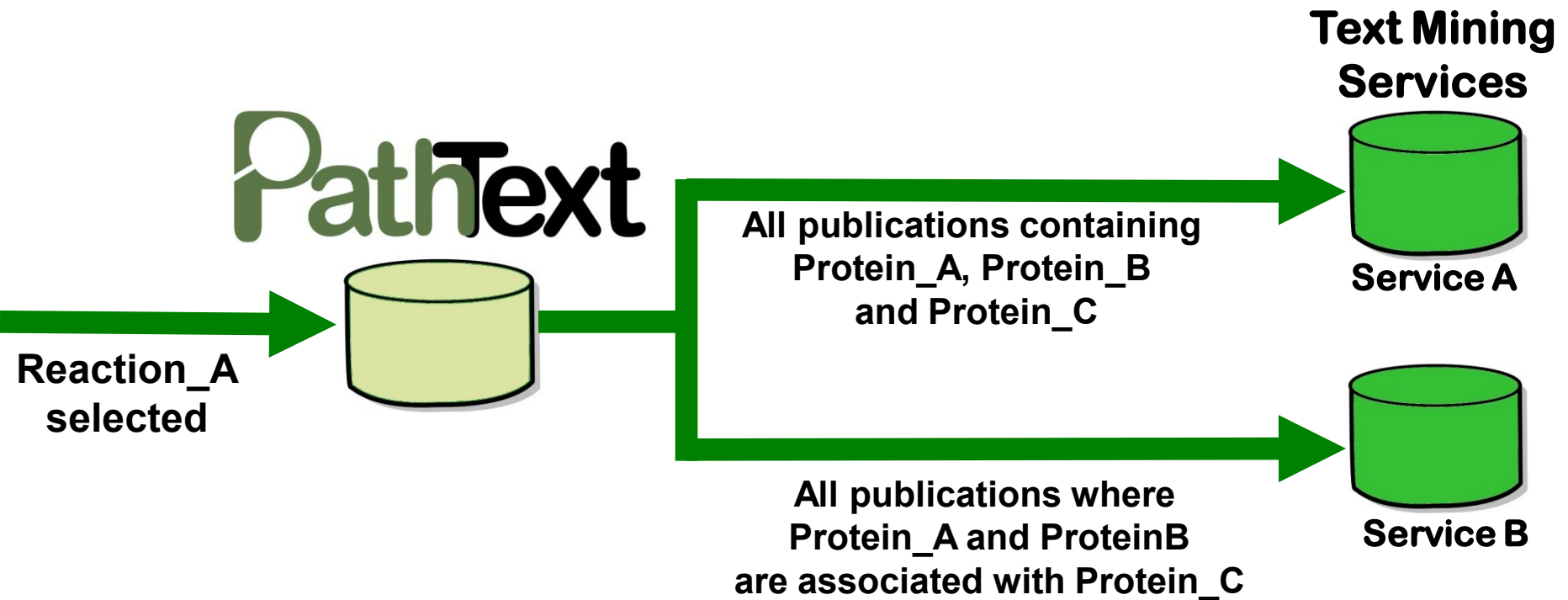
Step 2: User Selection

A user selects one or more nodes in the pathway visualization. The selection is then sent via web service request to the PathText system in the form of unique ids for each selected entity



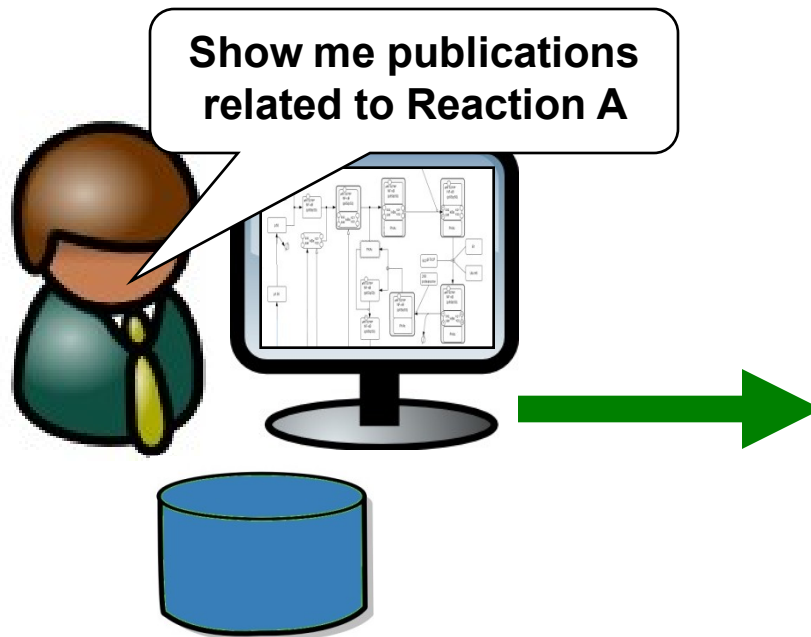
Step 3: Querying Text Mining Services

PathText then uses the association table to convert these ids into queries tailored to the requirements of each Text Mining Service



Step 4: User query matched with Publications

Finally, the results are then returned to the user and can either be displayed in the form of abstracts and links to full text, or displayed directly in the pathway visualization



Inhibition of the canonical IKK/NF kappa B pathway sensitizes human cancer cells to doxorubicin.

The NF kappa B family is composed by five subunits (p65/RelA, c-Rel, RelB, p105-p50/NF kappa B(1), p100-p52/NF kappa B(2)) and controls the expression of many genes that participate in cell cycle, apoptosis, and other key cellular processes. ... Transient down-regulation of members of the canonical pathway (p65, p52, c-Rel and IKKgamma/NEMO) by siRNA in HeLa cells increased doxorubicin cytotoxicity. ... To conclude, NF kappa B inhibition sensitized cells to doxorubicin, implying directly p65, p52, c-Rel and IKKgamma/NEMO subunits in chemoresistance, but not RelB. ...

PMID:17890907 Cell Cycle 2007 Jul

Transcriptional profiling of the nuclear factor-kappaB pathway identifies a subgroup of primary lymphoma of the central nervous system with low BCL10 expression.

... Compared with nonmalignant germinal center centroblasts, expression of BCL10, REL, IAP1, and TRAF1 was significantly lower in PCNSLs, whereas that of BAX, BCLXL, BCL2, MALT1, CARD9, CARD10, CARD11, CARD14, CCND2, cFLIP, RELA, RELB, NFKB1, NFKB2, and IRF4 was higher. ... Thus, these quantitative RT-PCR data with expression of genes of the NF-kappaB family as well as NF-kappaB-regulated genes together with immunohistochemical detection of nuclear RELA and REL indicate activation of the NF-kappaB pathway in PCNSLs, which may contribute to their high proliferative activity and the low level of apoptosis.

PMID:17356384 J. Neuropathol. Exp. Neurol. 2007 Mar

Essential roles of c-Rel in TLR-induced IL-23 p19 gene expression in dendritic cells.

... Unexpectedly, mutation of either of these two c-Rel binding sites completely abolished the p19 promoter activity induced by five TLRs (2, 3, 4, 6, and 9) and four members of the NF-kappaB family (c-Rel, p65, p100, and p105). ...

PMID:17182554 J. Immunol. 2007 Jan 1

NFKB1 is a direct target of the TAL1 oncoprotein in human T leukemia cells.

We recently showed that a subset of human T acute lymphoblastic leukemia (T-ALL) cell lines