



SOFTWARE TOOL ARTICLE

**REVISED** eXamine: Visualizing annotated networks in Cytoscape  
[version 2; referees: 4 approved]

Philipp Spohr<sup>1</sup>, Kasper Dinkla<sup>2</sup>, Gunnar W. Klau <sup>1</sup>, Mohammed El-Kebir<sup>3</sup>

<sup>1</sup>Algorithmic Bioinformatics, Heinrich Heine University, Düsseldorf, Germany

<sup>2</sup>Independent Researcher, Eindhoven, Netherlands

<sup>3</sup>Department of Computer Science, University of Illinois at Urbana-Champaign, Urbana, IL, USA

**v2** First published: 30 Apr 2018, 7:519 (doi: [10.12688/f1000research.14612.1](https://doi.org/10.12688/f1000research.14612.1))  
Latest published: 05 Jul 2018, 7:519 (doi: [10.12688/f1000research.14612.2](https://doi.org/10.12688/f1000research.14612.2))

**Abstract**

eXamine is a Cytoscape app that displays set membership as contours on top of a node-link layout of a small graph. In addition to facilitating interpretation of enriched gene sets of small biological networks, eXamine can be used in other domains such as the visualization of communities in small social networks. eXamine was made available on the Cytoscape App Store in March 2014, has since registered more than 7,700 downloads, and has been highly rated by more than 25 users. In this paper, we present eXamine's new automation features that enable researchers to compose reproducible analysis workflows to generate visualizations of small, set-annotated graphs.

**Keywords**

functional subnetwork modules, visualization, enrichment, graph drawing, network community, Euler diagram







This article is included in the [Cytoscape Apps gateway](#).

**Open Peer Review**

Referee Status:

	Invited Referees			
	1	2	3	4
<b>REVISED</b>				
<b>version 2</b>	report	report	report	report
published	↑	↑	↑	↑
05 Jul 2018				
<b>version 1</b>	?	?		?
published	report	report	report	report
30 Apr 2018				

- Barry Demchak** , University of California, San Diego, USA
- Richa Batra** , Helmholtz Center Munich, Germany  
Technical University of Munich, Germany
- Markus List** , Technical University Munich, Germany
- Alexander Pico** , Gladstone Institutes, USA

**Discuss this article**

Comments (0)

**Corresponding author:** Mohammed El-Kebir ([melkebir@illinois.edu](mailto:melkebir@illinois.edu))

**Author roles:** **Spohr P:** Data Curation, Methodology, Software, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing; **Dinkla K:** Conceptualization, Methodology, Software, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing; **Klau GW:** Conceptualization, Supervision, Validation, Writing – Original Draft Preparation, Writing – Review & Editing; **El-Kebir M:** Conceptualization, Data Curation, Supervision, Validation, Writing – Original Draft Preparation, Writing – Review & Editing

**Competing interests:** No competing interests were disclosed.

**Grant information:** The author(s) declared that no grants were involved in supporting this work.

**Copyright:** © 2018 Spohr P *et al.* This is an open access article distributed under the terms of the [Creative Commons Attribution Licence](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. Data associated with the article are available under the terms of the [Creative Commons Zero "No rights reserved" data waiver](#) (CC0 1.0 Public domain dedication).

**How to cite this article:** Spohr P, Dinkla K, Klau GW and El-Kebir M. **eXamine: Visualizing annotated networks in Cytoscape [version 2; referees: 4 approved]** *F1000Research* 2018, 7:519 (doi: [10.12688/f1000research.14612.2](https://doi.org/10.12688/f1000research.14612.2))

**First published:** 30 Apr 2018, 7:519 (doi: [10.12688/f1000research.14612.1](https://doi.org/10.12688/f1000research.14612.1))

**REVISED** Amendments from Version 1

This is a revised version addressing the various comments made by the referees. Previous Figures 1 and 2 are now merged into a single figure, and we included R code for the two cases.

**See referee reports**

## Introduction

The Cytoscape app eXamine visualizes a small graph and a collection of node sets. The main purpose of eXamine is to aid in the interpretation of a small subnetwork module extracted from a large biological network<sup>1</sup>. Cytoscape apps like jActiveModules<sup>2</sup> or external tools like Heinz<sup>3</sup> extract subnetwork modules from a protein-protein interaction network given gene expression data. To interpret the identified subnetwork module, a frequent follow-up analysis is to compute enrichment of the nodes of the identified module in terms of known annotations such as from the Gene Ontology (GO)<sup>4,5</sup> or from the Kyoto Encyclopedia of Genes and Genomes (KEGG)<sup>6</sup>. These annotations are a collection of node sets. eXamine provides a visual analysis approach that facilitates interpretation of the subnetwork module and the identified node sets by biologists. Given a small collection of node sets, eXamine generates a visualization of a small graph as a node-link layout together with contours for the selected sets (Figure 1). More specifically, the layout is computed using a variation of the algorithm described in 7. This algorithm preserves topological distances along inter-module links as much as possible, while making sure that none of the nodes overlap. In addition, spanning graphs are derived for those sets that have been selected by the user. These graphs are included in the computation of topological distances between nodes, pulling the nodes closer together. The spanning graphs are also used to draw the set contours, by adjusting the associated links to form the rounded shapes that visually encompass and connect nodes.

As an alternative to eXamine, the existing group layout of Cytoscape can be used to show node partitions by visualizing disjoint sets in separate circles. The Venn and Euler diagram app<sup>8</sup> for Cytoscape visualizes overlapping sets. However, in both cases network and group analysis are visually separate. Finally, Cytoscape provides a group viewer<sup>1</sup> that aggregates groups into meta-nodes, without making group overlaps explicit. Here, we present a new version of eXamine that uses Cytoscape's recently introduced automation features. With these new features it becomes possible to create reproducible analysis workflows that generate appealing visualizations of small, set-annotated graphs. We demonstrate eXamine's new automation features using two use cases. The first use case replicates the case study provided in the original eXamine publication<sup>1</sup>. The second use case, the analysis of a social network, demonstrates that eXamine is applicable to other domains beyond computational biology.

## Methods

### Implementation

eXamine is implemented in Java and is available as an app for Cytoscape 3. We used WebCola algorithms<sup>11</sup> to simultaneously lay out nodes, links, and set contours. We refer to the original publication<sup>1</sup> for additional implementation details regarding the used visualization techniques. Since the typical analysis workflows of eXamine consist of relatively simple commands that do not require streaming of complex data, we implemented the automation features through the 'Command' interface rather than the 'Function' interface. As a result, eXamine's commands can be either directly from Cytoscape or through Cytoscape's REST (CyREST) interface from a Jupyter notebook or from the programming language R.

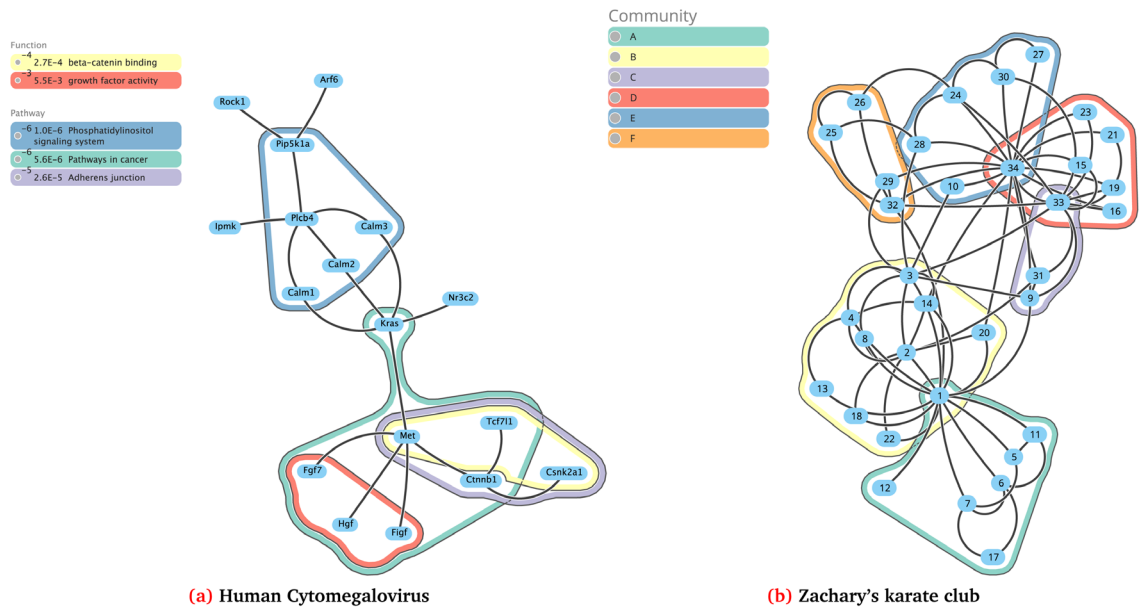
### Operation

As eXamine requires Cytoscape 3.6 to run, eXamine has the same system requirements as Cytoscape 3.6. eXamine can be operated through the Cytoscape graphical user interface (GUI) or through the new 'Command' interface of Cytoscape. We refer to the original publication<sup>1</sup> for GUI instructions. In the following, we will describe and use the new 'Command' interface of eXamine.

Table 1 provides a summary of the API of the eXamine commands and their parameters. To enable workflow authors to use eXamine's automation most effectively, we also generated Swagger-based documentation that describes the

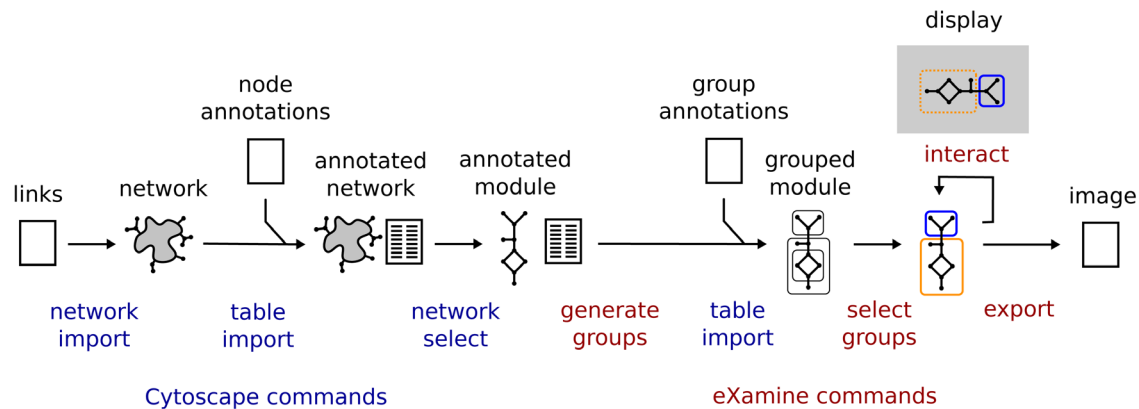
<sup>1</sup>[http://manual.cytoscape.org/en/stable/Creating\\_Networks.html#grouping-nodes](http://manual.cytoscape.org/en/stable/Creating_Networks.html#grouping-nodes)

<sup>11</sup><http://ialab.it.monash.edu/webcola/>



(a) Human Cytomegalovirus

(b) Zachary's karate club



(c) Workflow for generating (a) and (b) using Cytoscape and eXamine automation commands

**Figure 1. eXamine displays set membership as contours on top of a node-link layout. (a–b)** The graphs generated in the two use cases using eXamine’s automation features. **(c)** The first step in the workflow consists of importing a network, followed by importing node annotations that associate each node with a set of groups. Next, we optionally select a smaller subnetwork to visualize. We generate an internal representation of the groups, and import additional group annotations. After selecting the groups to visualize, we export an image of the visualization. Alternatively, we can launch a window that allows the user to select different groups.

commands and arguments. This documentation can be accessed in the Cytoscape menu: Help → Automation → CyREST Command API. **Figure 1** shows a typical workflow of eXamine analysis, where the commands provided are in the ‘examine’ namespace (red) and the commands provided by Cytoscape are colored blue.

**Use cases**

To illustrate the new ‘Command’ interface of eXamine, we present two use cases. In the first use case, we describe a workflow to study a small subnetwork module extracted from the KEGG mouse network<sup>6</sup>. In the second use case, we describe a workflow to study Zachary’s karate club, a well-known social network. Both workflows are available as Jupyter notebooks and as R markdown documents. The workflows require Cytoscape (≥ v3.6) and a recent Python version (Python v2.7 or Python v3.6) or R (≥ v3.4). We note that the commands described in the two use cases can also be directly executed from Cytoscape via the ‘Commands dialog’ or a separate ‘Commands script’<sup>111</sup>.

<sup>111</sup>[http://manual.cytoscape.org/en/stable/Programmatic\\_Access\\_to\\_Cytoscape\\_Features\\_Scripting.html](http://manual.cytoscape.org/en/stable/Programmatic_Access_to_Cytoscape_Features_Scripting.html)

**Table 1. Application Programming Interface (API) of Cytoscape commands provided by eXamine.** Figure 1 shows an example workflow using the below commands.

Command	Argument (type)	Description
<b>examine generate groups</b>		Generates eXamine groups from a given set of columns of the node table
	<ul style="list-style-type: none"> <li>○ network (CyNetwork)</li> <li>○ selectedGroupColumns (List)</li> <li>○ useAllNodes (Boolean)</li> </ul>	<p>Network for which to generate the groups</p> <p>Columns of the node table from which the groups will be created (comma separated)</p> <p>Indicates whether to use all nodes for group generation, or only those nodes currently selected in Cytoscape</p>
<b>examine select groups</b>		Select groups to be visualized by eXamine
	<ul style="list-style-type: none"> <li>○ selectedGroup (List)</li> </ul>	Group identifiers to select (comma separated)
<b>examine remove groups</b>		Removes all groups generated by eXamine
	<ul style="list-style-type: none"> <li>○ network (CyNetwork)</li> </ul>	Network from which to remove the groups
<b>examine update settings</b>		Updates visualization settings
	<ul style="list-style-type: none"> <li>○ network (CyNetwork)</li> <li>○ labelColumn (String)</li> <li>○ urlColumn (String)</li> <li>○ scoreColumn (String)</li> <li>○ showScore (Boolean)</li> <li>○ selectedGroupColumns (List)</li> </ul>	<p>Network to visualize</p> <p>Column of the node table that contains the node labels</p> <p>Column of the node table that contains URLs with additional information about the nodes</p> <p>Column that contains annotation enrichment scores</p> <p>Indicates whether to show annotation enrichment scores</p> <p>Group columns of the node table to show in the visualization</p>
<b>examine interact</b>		Shows the selected network and groups in an interactive visualization window
<b>examine export</b>		Generates an SVG file with a visualization of the selected network and groups
	<ul style="list-style-type: none"> <li>○ path (String)</li> </ul>	Filename

### Use case 1: Dysregulated signaling in Human Cytomegalovirus

The *Human Cytomegalovirus (HCMV)* is a highly-contagious herpes virus. Previously<sup>1</sup>, we used eXamine to interpret a small subnetwork module (17 nodes and 18 edges) extracted from the KEGG mouse network using Heinz<sup>3</sup> given gene expression data of an HCMV-infected mouse cell line. Node sets of this subnetwork module were annotated using enriched pathways from KEGG and enriched terms from GO. Below, we provide Python code<sup>IV</sup> that uses Cytoscape's and eXamine's automation features to generate Figure 1. An R markdown document for this use case is available on the git repository<sup>V</sup>.

1. To begin, we define a helper function that creates HTTP POST requests.

```
import requests
PORT_NUMBER = 1234
BASE = "http://localhost:" + str(PORT_NUMBER) + "/v1/"
BASE_URL = "https://raw.githubusercontent.com/lscwi/"
          + "eXamine/master/doc/tutorial/"
```

<sup>IV</sup><https://github.com/lscwi/eXamine/blob/master/doc/tutorial/eXamineNotebook/eXamineTutorial.ipynb>

<sup>V</sup>[https://github.com/lscwi/eXamine/blob/master/doc/tutorial/eXamineNotebook/eXamineTutorial\\_R.Rmd](https://github.com/lscwi/eXamine/blob/master/doc/tutorial/eXamineNotebook/eXamineTutorial_R.Rmd)

```
def executeRestCommand(namespace="", command="", args={}):
    postString = BASE + "commands/" + namespace + "/" + command
    firstarg = True
    for arg in args:
        postString += ("?" if firstarg else "&") + arg + "=" + args[arg]
        if (firstarg): firstarg = False
    res = requests.post(postString, json=args)
    return res
```

2. We import the full KEGG mouse network using the 'network import url' command provided by Cytoscape. This a protein-protein interaction network that is typically used for the analysis of high-throughput biological data in the context of a biological network. The Cytoscape app KEGGscope provides functionality for importing pathways from KEGG<sup>9</sup>.

```
executeRestCommand("network", "import url",
                  {"indexColumnSourceInteraction" : "2",
                   "indexColumnTargetInteraction" : "2",
                   "url" : BASE_URL + "edges.txt"})
```

3. We import node annotation, containing set membership information, using the 'table import url' command provided by Cytoscape. We note that the entries of the columns of type list (indicated by 'sl' in 'dataTypeList') are separated by a pipe character.

```
executeRestCommand("table", "import url",
                  {"firstRowAsColumnNames" : "true",
                   "keyColumnIndex" : "1",
                   "startLoadRow" : "1",
                   "dataTypeList" : "s,s,f,f,f,s,s,s,sl,sl,sl,sl",
                   "url" : BASE_URL + "nodes_induced.txt"})
```

4. Using the 'network select' command provided by Cytoscape, we select the nodes that comprise the subnetwork module identified by Heinz<sup>3</sup>.

```
executeRestCommand("network" , "select" , {"nodeList" : "Module:small"})
```

5. Using the 'examine generate groups' command, we generate group nodes for each member of the specified sets that occur in the module (the nodes we previously selected).

```
executeRestCommand("examine", "generate groups",
                  {"selectedGroupColumns" : "Function,Pathway"})
```

6. We import additional annotations that describe the generated groups using the 'import table url' command of Cytoscape.

```
executeRestCommand("table", "import url",
                  {"firstRowAsColumnNames" : "true",
                   "keyColumnIndex" : "1", "startLoadRow" : "1",
                   "url" : BASE_URL + "sets_induced.txt"})
```

7. As described in 1, we consider groups from the sets *Pathway* and *Function* using the 'examine update settings' command. In addition, we specify the columns that contain node labels and URLs as well as group scores and annotations.

```
executeRestCommand("examine", "update settings",
                  {"labelColumn" : "Symbol", "urlColumn" : "URL",
                   "scoreColumn" : "Score", "showScore" : "true",
                   "selectedGroupColumns" : "Component, Function"})
```

8. Using the 'examine select groups' command, we select the same groups as described in 1.

```
executeRestCommand("examine", "select groups", {"selectedGroups" :
        "GO:0008013,GO:0008083,mmu04070,mmu05200,mmu04520"})
```

9. Finally, we export the visualization to a scalable vector graphics (SVG) file using the ‘examine export’ command. [Figure 1a](#) shows the resulting graph.

```
executeRestCommand("examine", "export", {"path": "fig1a.svg"})
```

Alternatively, using the ‘examine interact’ command, we can launch an interactive visualization window that allows us to select different groups.

```
executeRestCommand("examine", "interact", {})
```

### Use case 2: Zachary’s karate club

We consider the graph ‘Zachary’s karate club’, which is an undirected social network of friendships between 34 members of a karate club at a US university in the 1970s<sup>10</sup>. Here, we use eXamine to visualize the six overlapping communities of this network identified in [11](#). We provide Python code below<sup>vi</sup>. The corresponding R markdown document is available on the git repository<sup>vii</sup>.

1. We use the same helper function as defined in the previous use case.
2. We import the network using the ‘network import url’ command provided by Cytoscape.

```
executeRestCommand("network", "import url",
  {"indexColumnSourceInteraction" : "2",
   "indexColumnTargetInteraction" : "2",
   "url" : BASE_URL + "edges_karate.gml"})
```

3. We import node annotation, containing set membership information, using the ‘table import url’ command provided by Cytoscape. Entries of the column of type list (indicated by ‘sl’ in ‘dataTypeList’) are separated by a pipe character.

```
executeRestCommand("table", "import url",
  {"firstRowAsColumnNames" : "true",
   "keyColumnIndex" : "1", "startLoadRow" : "1",
   "dataTypeList" : "s,sl",
   "url" : BASE_URL + "nodes_karate_induced.txt"})
```

4. Using the ‘network select’ command provided by Cytoscape, we select all nodes of the network.

```
executeRestCommand("network" , "select", {"nodeList":"all"})
```

5. Using the ‘examine generate groups’ command, we generate group nodes for each member of the specified sets that annotate the nodes of the network.

```
executeRestCommand("examine", "generate groups",
  {"selectedGroupColumns" : "Community"})
```

6. We consider groups from the set *Community* using the ‘examine update settings’ command.

```
executeRestCommand("examine", "update settings",
  {"labelColumn" : "label", "urlColumn" : "label",
   "showScore" : "false",
   "selectedGroupColumns" : "Community"})
```

7. We select all six groups using the ‘examine select groups’ command.

```
executeRestCommand("examine", "select groups",
  {"selectedGroups" : "A,B,C,D,E,F"})
```

---

<sup>vi</sup><https://github.com/lscwi/eXamine/blob/master/doc/tutorial/eXamineNotebook/eXamineTutorial2.ipynb>

<sup>vii</sup>[https://github.com/lscwi/eXamine/blob/master/doc/tutorial/eXamineNotebook/eXamineTutorial2\\_R.Rmd](https://github.com/lscwi/eXamine/blob/master/doc/tutorial/eXamineNotebook/eXamineTutorial2_R.Rmd)

8. Finally, we export the visualization to an SVG file using the 'examine export' command. [Figure 1b](#) shows the resulting graph.

```
executeRestCommand("examine", "export", {"path": "fig1b.svg"})
```

### Discussion and conclusions

eXamine is limited to visualizing small, relatively sparse networks. It is not possible to use eXamine to construct a comprehensive layout if the network consists of hundreds of nodes or if there are dozens of annotation sets to visualize at the same time. This is a natural limitation of any visualization approach based on node-link diagrams and set contours.

eXamine currently uses Cytoscape's CyREST API to import networks and their annotations. If apps that provide gene enrichment analysis functionality, such as BiNGO<sup>12</sup>, would expose this functionality through the CyREST API, we envision updating the workflow to include this type of upstream analysis. Finally, it would be good to enhance the API to return richer R and Python data types rather than a flag indicating whether the command succeeded. For instance, upon a 'generate groups' it would be good to actually return a dataframe containing the groups. This, however, will require switching from the 'Command' interface to the 'Functions' interface.

### Summary

eXamine is a Cytoscape app for a set-oriented visual analysis approach for small annotated graphs that displays set membership as contours on top of a node-link layout ([Figure 1](#)). In this paper, we presented new automation features for eXamine that are accessible through Cytoscape's REST API ([Table 1](#)). As such, researchers can embed eXamine in reproducible and well-documented workflows that generate appealing visualizations of small, set-annotated graphs ([Figure 1](#)). We demonstrated two such workflows in the context of computational network biology and social network analysis.

### Data and software availability

1. Cytoscape app store: <http://apps.cytoscape.org/apps/examine>
2. Source code: <https://github.com/lscwi/eXamine>
3. Link to latest commit as of release: <https://github.com/lscwi/eXamine/commit/c1c3e643250c668c77a40cb982f9d122993e22a5>
4. Link to archived source code as at time of publication: <https://doi.org/10.5281/zenodo.1302788><sup>13</sup>
5. License: GPL v2.0
6. Link to use case data: <https://github.com/lscwi/eXamine/blob/master/doc/tutorial>.
7. Link to Jupyter notebooks: <https://github.com/lscwi/eXamine/blob/master/doc/tutorial/eXamineNotebook/eXamineTutorial.ipynb> and <https://github.com/lscwi/eXamine/blob/master/doc/tutorial/eXamineNotebook/eXamineTutorial2.ipynb>.
8. Link to R markdown documents: [https://github.com/lscwi/eXamine/blob/master/doc/tutorial/eXamineNotebook/eXamineTutorial\\_R.Rmd](https://github.com/lscwi/eXamine/blob/master/doc/tutorial/eXamineNotebook/eXamineTutorial_R.Rmd) and [https://github.com/lscwi/eXamine/blob/master/doc/tutorial/eXamineNotebook/eXamineTutorial2\\_R.Rmd](https://github.com/lscwi/eXamine/blob/master/doc/tutorial/eXamineNotebook/eXamineTutorial2_R.Rmd).

---

### Author contributions

KD and PS implemented the automation features of eXamine. MEK and GWK supervised the project. GWK, KD, MEK and PS wrote the paper. All authors read and approved the manuscript.

### Competing interests

No competing interests were disclosed.

### Grant information

The author(s) declared that no grants were involved in supporting this work.



## References

---

1. Dinkla K, El-Kebir M, Bucur CI, *et al.*: **eXamine: exploring annotated modules in networks**. *BMC Bioinformatics*. 2014; **15**: 201.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
2. Ideker T, Ozier O, Schwikowski B, *et al.*: **Discovering regulatory and signalling circuits in molecular interaction networks**. *Bioinformatics*. 2002; **18**(Suppl 1): S233–S240.  
[PubMed Abstract](#) | [Publisher Full Text](#)
3. Dittrich MT, Klau GW, Rosenwald A, *et al.*: **Identifying functional modules in protein-protein interaction networks: an integrated exact approach**. *Bioinformatics*. 2008; **24**(13): i223–31.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
4. Ashburner M, Ball CA, Blake JA, *et al.*: **Gene ontology: tool for the unification of biology**. The Gene Ontology Consortium. *Nat Genet*. 2000; **25**(1): 25–29.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
5. The Gene Ontology Consortium: **Expansion of the Gene Ontology knowledgebase and resources**. *Nucleic Acids Res*. 2017; **45**(D1): D331–D338.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
6. Kanehisa M, Furumichi M, Tanabe M, *et al.*: **KEGG: new perspectives on genomes, pathways, diseases and drugs**. *Nucleic Acids Res*. 2017; **45**(D1): D353–D36.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
7. Dwyer T, Koren Y, Marriott K: **IPSEP-COLA: an incremental procedure for separation constraint layout of graphs**. *IEEE Trans Vis Comput Graph*. 2006; **12**(5): 821–828.  
[PubMed Abstract](#) | [Publisher Full Text](#)
8. Venn and Euler Diagrams: **Venn and Euler Diagrams**.  
[Reference Source](#)
9. Nishida K, Ono K, Kanaya S, *et al.*: **KEGGscape: a Cytoscape app for pathway data integration [version 1; referees: 1 approved, 2 approved with reservations]**. *F1000Res*. 2014; **3**: 144.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
10. Zachary WW: **An information flow model for conflict and fission in small groups**. *J Anthropol Res*. 1977; **33**(4): 452–473.  
[Publisher Full Text](#)
11. Chen W, Liu Z, Sun X, *et al.*: **A game-theoretic framework to identify overlapping communities in social networks**. *Data Min Knowl Discov*. 2010; **21**(2): 224–240.  
[Publisher Full Text](#)
12. Maere S, Heymans K, Kuiper M: **BINGO: a Cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks**. *Bioinformatics*. 2005; **21**(16): 3448–3449.  
[PubMed Abstract](#) | [Publisher Full Text](#)
13. El-Kebir M, Dinkla K, Klau GW, *et al.*: **Is-cw/eXamine: Updated documentation (Version v2.1.2)**. *Zenodo*. 2018.  
[Data Source](#)

# Open Peer Review

Current Referee Status:



## Version 2

Referee Report 19 July 2018

doi:10.5256/f1000research.16899.r35797



**Richa Batra**  1,2

<sup>1</sup> Institute of Computational Biology, Helmholtz Center Munich, Neuherberg, Germany

<sup>2</sup> Technical University of Munich, Munich, Germany

Thank you for updating the manuscript and adding the R code.

**Competing Interests:** No competing interests were disclosed.

**Referee Expertise:** patient stratification, disease endotyping, clustering, network analysis, pathway enrichment, classification

**I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.**

Referee Report 13 July 2018

doi:10.5256/f1000research.16899.r35798



**Alexander Pico** 

Gladstone Institutes, San Francisco, CA, USA

Thanks for addressing the points of my review.

**Competing Interests:** No competing interests were disclosed.

**I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.**

Referee Report 10 July 2018

doi:10.5256/f1000research.16899.r35796



**Markus List** 

Chair of Experimental Bioinformatics, TUM School of Life Sciences, Technical University Munich, Munich, Germany

I have no additional comments to make.

**Competing Interests:** No competing interests were disclosed.

**I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.**

Referee Report 06 July 2018

doi:10.5256/f1000research.16899.r35799



**Barry Demchak** 

Department of Medicine, University of California, San Diego, La Jolla, CA, USA

Thank you. These changes satisfy my concerns. Thank you for taking such care with this material.

**Competing Interests:** No competing interests were disclosed.

**I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.**

---

## Version 1

Referee Report 25 June 2018

doi:10.5256/f1000research.15901.r33592



**Alexander Pico** 

Gladstone Institutes, San Francisco, CA, USA

Many researchers arrive at point in enrichment analysis where they have tables of enriched ontology terms and pathways, but don't know how to go beyond presenting those data as tables or perhaps bar graphs. The use of networks is an intuitive and powerful visualization option. The eXamine app for Cytoscape provides a nice approach for displaying a simplified version of enrichment results. A couple steps are missing, however, and a few minor edits are suggested.

Missing steps:

1. All sources of enrichment that I'm aware of provide rows of terms and a list of genes in a single field per row. It would be nice if you provided a bit of Python to transform that common input data type to that required by eXamine.
2. I was wondering about genes that belong to more than one term (often the case). I looks like your tool handles these via a piped list. Is that right? Please describe the correct syntax in the text so readers don't have to guess.

Minor suggestions:

1. Groups in Cytoscape are no longer provided as a separate plugin, but rather are now integrated into the core of Cytoscape. Your reference to "RBVI Cytoscape plugins" should be updated to simply mention the Cytoscape manual and point to the section on groups:  
[http://manual.cytoscape.org/en/stable/Creating\\_Networks.html#grouping-nodes](http://manual.cytoscape.org/en/stable/Creating_Networks.html#grouping-nodes)
2. The LaTeX formatting for your Python code is unfortunately rendering single quotes in a way that throws `SyntaxError` in Python. Can you update the rendering so one can simply copy/paste snippets?
3. In the next version of this paper, I'd highly recommend using `py2cytoscape` (<https://github.com/cytoscape/py2cytoscape>). This will make the code you have write much more concise and easier to maintain or adapt. Likewise, I'd really like to see R examples and you can leverage `RCy3` to make that code easy to write as well (<http://bioconductor.org/packages/release/bioc/html/RCy3.html>).

**Is the rationale for developing the new software tool clearly explained?**

Yes

**Is the description of the software tool technically sound?**

Yes

**Are sufficient details of the code, methods and analysis (if applicable) provided to allow replication of the software development and its use by others?**

Yes

**Is sufficient information provided to allow interpretation of the expected output datasets and any results generated using the tool?**

Yes

**Are the conclusions about the tool and its performance adequately supported by the findings presented in the article?**

Yes

**Competing Interests:** No competing interests were disclosed.

**I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.**

Author Response 01 Jul 2018

**Mohammed El-Kebir**, University of Illinois at Urbana-Champaign, USA

We thank the referee for the positive feedback. Below our point-by-point response to the referee's comments.

- Regarding the missing details on generating the enrichment input, the DAVID webtool (<https://david.ncifcrf.gov/>) yields a file where the rows are genes containing a list of terms.

Implementing a generic Python function that transforms enrichment output that comes in various to our required format is a challenging task. Ideally, the enrichment step should be part of the workflow, as described in Discussion.

- We handle genes with multiple terms using lists, and the default separator used by Cytoscape's import table functionality is indeed a pipe character. We updated the tutorials and text to describe this.
- We updated the reference to "RBVI Cytoscape plugins".
- We replaced all single quotes by double quotes, which are retained by the used LaTeX package for displaying the code (package is 'listing'). Simple copy-paste of the commands into the Python interpreter should work now.
- We thank the referee for the suggestions regarding R and Python. We now use the RCy3 library for the two use cases in R. As for py2cytoscape, we found that the functionality that we need is not supported directly and requires separate script files for each command. We found that using the REST interface directly was easier.

**Competing Interests:** No competing interests.

Referee Report 21 June 2018

doi:[10.5256/f1000research.15901.r34808](https://doi.org/10.5256/f1000research.15901.r34808)



**Markus List** 

Chair of Experimental Bioinformatics, TUM School of Life Sciences, Technical University Munich, Munich, Germany

In their manuscript "eXamine: Visualizing Annotated Networks in Cytoscape", Spohr et al. describe a new version of their Cytoscape tool eXamine. eXamine allows user to enrich the visualization of small networks via colored contours which represent specific annotations of subnetworks, e.g. functional annotation. eXamine is very useful because it not only highlights group properties in a visually appealing way but also allows the representation to be changed interactively. The software always attempts to achieve an optimal layout in which overlaps are kept at a minimum. Here, the authors have adapted eXamine to make use of the new Cytoscape automation features which allow for programmatic control of the plotting process, thus allows for generating reproducible results that can further be embedded in scripts.

The manuscript is well written and describes the new features and their usage clearly.

The code of eXamine is hosted on github and contains the two use cases from the manuscript in the form of popular Jupyter notebooks. Each step of the examples is well described and illustrates to the user how eXamine can be used in a scripting environment such as Python. In particular less obvious steps like extracting network ids from a JSON result help new users to achieve results fairly quickly.

I have only minor comments:

- Unfortunately the second use case (Jupyter notebook) in the git repository did not work due to the use of a local file path instead of a public URL.

- Add the end of the first paragraph the manuscript says “by dilating and eroding the associated links”. I’m not sure what is meant here with erosion.
- Use case 1, first paragraph, there is a typo: feautres -> features
- The github readme should briefly describe the jupyter notebooks and link to them.

**Is the rationale for developing the new software tool clearly explained?**

Yes

**Is the description of the software tool technically sound?**

Yes

**Are sufficient details of the code, methods and analysis (if applicable) provided to allow replication of the software development and its use by others?**

Yes

**Is sufficient information provided to allow interpretation of the expected output datasets and any results generated using the tool?**

Yes

**Are the conclusions about the tool and its performance adequately supported by the findings presented in the article?**

Yes

**Competing Interests:** No competing interests were disclosed.

**I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.**

Author Response 01 Jul 2018

**Mohammed El-Kebir**, University of Illinois at Urbana-Champaign, USA

We thank the referee for the positive feedback. Below our point-by-point response to the referee's comments.

- We fixed the link to the git repository for the second use case.
- We changed the wording of the sentence where we previously used 'erosion'.
- We fixed the typo. Thank you for pointing this out.
- We updated the Github README.md file.

**Competing Interests:** No competing interests.

Referee Report 20 June 2018

doi:10.5256/f1000research.15901.r34810



 **Richa Batra**  1,2

<sup>1</sup> Institute of Computational Biology, Helmholtz Center Munich, Neuherberg, Germany

<sup>2</sup> Technical University of Munich, Munich, Germany

This paper presents new features to an existing cytoscape app namely eXamine. eXamine was initially developed to visualize small graphs and node sets. This new version of the app has automation features for reproducible research. Given the importance of reproducible research, this new version of the app is highly relevant.

Figure 2 should really be Figure 1. For someone new to eXamine, it aptly summarizes the app. Also the summary at the end is very clear.

It would have been great to have a small dataset and step by step guide through the cytoscape app. Then small R script that I can run to generate the same results. Could you extend it for R users? Do you plan it in future work?

If the paper is about automation of eXamine via python, then it should reflect in title. The current name of the paper is misleading.

**Is the rationale for developing the new software tool clearly explained?**

Partly

**Is the description of the software tool technically sound?**

Yes

**Are sufficient details of the code, methods and analysis (if applicable) provided to allow replication of the software development and its use by others?**

Yes

**Is sufficient information provided to allow interpretation of the expected output datasets and any results generated using the tool?**

Yes

**Are the conclusions about the tool and its performance adequately supported by the findings presented in the article?**

Yes

**Competing Interests:** No competing interests were disclosed.

**Referee Expertise:** patient stratification, disease endotyping, clustering, network analysis, pathway enrichment, classification

**I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.**

Author Response 01 Jul 2018

**Mohammed El-Kebir**, University of Illinois at Urbana-Champaign, USA

We thank the referee for the positive feedback. Below our point-by-point response to the referee's comments.

- We have merged Figures 1 and 2 in the revised manuscript.
- We included R example code in the revised manuscript.

**Competing Interests:** No competing interests.

Referee Report 18 May 2018

doi:10.5256/f1000research.15901.r33593



**Barry Demchak** 

Department of Medicine, University of California, San Diego, La Jolla, CA, USA

This article describes the use of new Commands exposed by the eXamine app.

The article is well written and adequately describes the Commands. My comments are advisory, and I hope they can improve the paper.

The procedure for using Commands should be well known to a proficient Cytoscape user. However, there's an important class of user that wants to use Automation but isn't a Cytoscape expert. To help him/her along good reference Cytoscape manual (Commands Tool), and good to point out that the commands can be used from the Commands dialog, a separate Commands script (via the Cytoscape command line or the Tools | Execute Command File menu item), or via Python. (See last paragraph of Methods section.) Good, too, to remind the user that eXamine commands are in the examine namespace, and that different namespaces resolve to Cytoscape and other apps.

This paper is about Python usage. Good to remind user that similar calls can be made from R.

In the Introduction, it would be helpful to justify eXamine automation by giving an example of workflows that become possible. This motivates the user. This could be as simple as summarizing Use cases.

For Discussion, it may be worth speculating on the value of providing Python and R libraries that act as cover functions for the eXamine REST calls. Python and R programmers really want to think in terms of Python and R, not REST.

The color coding in Figure 2 is very effective. Can you also color code the corresponding REST calls in the Use Case section? This may require changing the color scheme so the examples can be easily read. (Maybe Red/Green or Red/Blue??)

In Use Cases, good to say where the Jupyter notebooks are ... even if you identify them at the end of the paper.

As a side note, I notice that all eXamine endpoints are GET. The Commands convention we're using now is POST, with GET being deprecated. In this case, it doesn't matter much, as the parameter list isn't long



and there isn't any return result that would benefit from JSON encoding. For a future release, good to consider POST versions, too. That way, if an error occurs, you'll be able to return it in a CIResponse structure.

For the BASE\_URL, the datasets are necessary to run the examples. Can you list them at the end of the paper? Of course, they must be persistently available for the life of the paper, correct??

In Use Case 1, step 2, it would be good to give a sentence or two explaining how/why a user would have created this dataset.

That's it ... nice job!

**Is the rationale for developing the new software tool clearly explained?**

Partly

**Is the description of the software tool technically sound?**

Yes

**Are sufficient details of the code, methods and analysis (if applicable) provided to allow replication of the software development and its use by others?**

Yes

**Is sufficient information provided to allow interpretation of the expected output datasets and any results generated using the tool?**

Yes

**Are the conclusions about the tool and its performance adequately supported by the findings presented in the article?**

Yes

**Competing Interests:** No competing interests were disclosed.

**I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.**

Author Response 01 Jul 2018

**Mohammed El-Kebir**, University of Illinois at Urbana-Champaign, USA

We thank the referee for the positive feedback. Below our point-by-point response to the referee's comments.

- We included the following sentences along with a reference to the Cytoscape documentation:

*'Figure 1 shows a typical workflow of eXamine analysis, where the commands provided are in the `examine` namespace (red) and the commands provided by Cytoscape are colored blue.'*

*'We note that the commands described in the two use cases can also be directly executed from Cytoscape via the 'Commands dialog' or a separate 'Commands script'.<sup>3</sup>'*

- We now provide R code for the two use cases.
- We reworded the last paragraph of the introduction to more clearly describe the benefits of eXamine's new automation features.
- We wrote the following sentences in Discussion:

*'Finally, it would be good to enhance the API to return richer R and Python data types rather than a flag indicating whether the command succeeded. For instance, upon a 'table import' it would be good to actually return a dataframe. This, however, will require switching from the 'Command' interface to the 'Function' interface.'*

- Thank you for the excellent suggestion about the coloring. We colored eXamine commands red and Cytoscape commands blue.
- We included URLs to the Jupyter and R markdown documents.
- We now use POST commands.
- We now include links to the datasets. They are indeed persistently archived using zenodo.
- We included the following sentences in step 2 of use case 1:

*'This a protein-protein interaction network that is typically used for the analysis of high-throughput biological data in the context of a biological network. The Cytoscape app KEGGscape provides functionality for importing pathways from KEGG [9].'*

**Competing Interests:** No competing interests.

The benefits of publishing with F1000Research:

- Your article is published within days, with no editorial bias
- You can publish traditional articles, null/negative results, case reports, data notes and more
- The peer review process is transparent and collaborative
- Your article is indexed in PubMed after passing peer review
- Dedicated customer support at every stage

For pre-submission enquiries, contact [research@f1000.com](mailto:research@f1000.com)

**F1000Research**