

```

WikiLinkPath = "~/projects/complete/WikiLink/trunk/";
AppendTo[$Path, WikiLinkPath <> "mathematica/"];

<< WikiLink`

indexPath = "http://tqft.net/w/index.php";
(*"http://math.berkeley.edu/~scott/w/index.php"*)

CreateWikiConnection[indexPath, "Scott", "umozY1"]

WikiUserName[]
Scott

ClearWikiCache[] := Module[{},
  Clear[UserContributionsHTML];
  UserContributionsHTML[user_String] := UserContributionsHTML[user] = Import[indexPath <>
    "?title=Special%3AContributions&contribs=user&hideminor=0&namespace=&target=" <>
    user <> "&year=&month=-1&limit=1000", "Text"];
  FindAllPages[] := FindAllPages[] = Union@@(FindAllPages /@ {"0", "1", "2", "3"})
]
ClearWikiCache[]

```

AllUserContributions collects a few spurious pages, but mostly seems okay.

```

UserContributionsHTML["Scott"]

AllUserContributions[user_String] :=
  Union[StringCases[StringReplace[UserContributionsHTML[user],
    ___ ~~ "<ul>" ~~ p___ ~~ "<!-- end content -->" ~~ __ :> p],
    "title=\"\" ~~ ShortestMatch[title__] ~~ "\">" :> title]]

AllUserContributions["Scott"]

```

{2000 Maniacs: copy, 2000 Maniacs: Letter to RA's, 3rd Evans Stair Running Competition, 4th Evans Stair Running Competition, 5th Evans Stair Running Competition, Analysing the MathLink protocol, Appendix 0: Student Committee Roster and Contact Info, APPENDIX 1: CHALK ART COMPETITION FLIER, Appendix 3c: CAC Judge Thank-You, APPENDIX 4: CHALK ART COMPETITION VOTING SLIPS, APPENDIX 5: CONGRATULATORY LETTERS, Arxiv-oai2rdf, Arxivwiki.org issues, ASCII map, Atlas of subfactors, Baby, Baby wishlist, Balkans, Birthdays, Blob complex, Bookmarklets, Books to give away, Buildering, Building an arxivwiki server, Career Day, Career Day 2005, Cell Phone, CHALK ART COMPETITION, Christmas list, Cologne visit, Conferences, Conjectures, Copy of Yossi's planning page for the Nose, Craigs List Ad Draft, Doi hashes, DOI hashes, Electronic money, Entering a (graded bipartite) graph, FOAF, Frankland Range, Future trips, Generators and Relations Talk 2, Give away, Google Print, Google Print competitors, Google Print development history, Google Print extensions, Google Print extensions announcement, Google Print hacking, Google Print Recorder, Graded graphs, Graph planar agebra, Graph planar algebra, Hacking the parser, Half Dome with Vlad, Help:Contents, Help:Editing, Homework 2, House Meeting Oct 2005, Image:3rd Evans Stair Race Results.pdf, Image:Career day04.gif, Image:Math121a Dec7 hw.pdf, Image:Math121a Review.pdf, Image:PCMI thumbnail.png, Inaugural Evans Stair Running Competition, INTRODUCTION, Java Security API, Jobs, Juggling for Geeks (abstract), Jython subclasses can't call Java super constructors, Khovanov Homology in Austin, Khovanov Workshop, Leaning Tower, Leaning Tower Trip Report, Letter of invitation to groups, Links, List of countries, List of things we need, Local Khovanov homology Talk 1, Lost luggage, Main Page, Math, Math121a, Math 54 Fall 2005 Homework Solutions, Math 54 Fall 2005 Quizzes, Mathematics, Maths, Matt's bed, Maui 2011, MediaWiki:Common.js, Meta:Help:Special characters, Metaphor/scratch/Example, M:Help:Table, Mills Real Analysis, Moving, Music list, Neighbourhood fruit, Nina's broken computer, Nina's Journal, NJToNJ, Notes for my talk, Old homework assignments, Old housing leads, Oldmathpapers.org issues tracker, Open Problems in Fusion Categories, Orange-Lemon Cake, Performers, Phone numbers, Phone Numbers, Photos, Pivotal Categories, Planar Algebras and Subfactors 2, Pivotal Categories, Planar Algebras and Subfactors 3, Precipitous Bluff, Previous teaching, Quantum Topology conference, Quantum Topology in Auckland, Quantum Topology in Wellington, Ranen's webpage, Renaming Tungsten, Salathe planning, Sandbox, Santa Barbara House Hunting, ScottWiki:All system messages, Scraping the arXiv, Setting up a Derby server, Spam morning of, October 30, Special:Contributions, Special:Contributions/CimVaf, Special:Contributions/Demgeorge5, Special:Contributions/FnnCsh, Special:Contributions/Ii7U8o, Special:Contributions/Jheena789, Special:Contributions/Lo0C0a, Special:Contributions/MarilenaGallo, Special:Contributions" rel="next, Student spam morning of, October 30, Su 3-referee, Subfactors in Maui, Taipa, Talk:Berkeley restaurant guide, Talk:Notes for my talk, Talk:Previous teaching, Talk:Vegan chocolate pecan brownies, Template:All messages, Template:Arxiv-link, Template:Current-address, Template:Current-office, Template:Lowercase, Template:Message box, The Nose, Todo, TQFT mini-course, Transient, Transient/Summer-meeting, Trips list, Tungsten installation instructions, Tungsten Todo List, UC Berkeley Buildering, USA, User:213.73.238.178 (not yet written), User:Scott, User talk:CimVaf (not yet written), User talk:Demgeorge5 (not yet written), User talk:FnnCsh (not yet written), User talk:Ii7U8o (not yet written), User talk:Jheena789 (not yet written), User talk:Lo0C0a (not yet written), User talk:MarilenaGallo (not yet written), User talk:Scott, Who Am I, Wikifying the Knot Atlas, Wish list, Work in progress, Zodiac}

```

hostName = "tqft.net";

nextPageURL [text_] := Module[{candidates}, candidates = StringCases [text, "<a href=\"" ~
  nextURL : ("/w/index.php?title=Special:Allpages&from=" ~ ShortestMatch [__]) ~
  "\" title=\"Special:Allpages\">Next page" :>
  "http://" <> hostName <> StringReplace [nextURL, "&" -> "&"], 1];
If [Length [candidates] == 0, $Failed, candidates [1]]
]

Clear [FindAllPages]
FindAllPages [] := FindAllPages [] = Union @@ (FindAllPages /@ {"0", "1", "2", "3"})

FindAllPages [nameSpaceNumber_] :=
Module[{targets = {}, allPagesURL =
  "http://" <> hostName <> "/w/index.php?title=Special%3AAllpages&from=&namespace=" <>
  nameSpaceNumber, allPagesText, newTargets},
While [allPagesURL != $Failed,
Print ["Looking at: " <> allPagesURL];
allPagesText = Import [allPagesURL, "Text"];
allPagesURL = nextPageURL [allPagesText];
newTargets = StringCases [allPagesText,
  "title=\"" ~ pagename : ShortestMatch [__] ~ "\" :> pagename];
Print ["Found " <> ToString [Length [newTargets]] <> " new targets."];
targets = targets ~ Join ~ newTargets;
];
Complement [targets, {"Special:Allpages"}]
]

FindAllPages [] // TableForm

```

2000 Maniacs: copy  
3rd Evans Stair Running Competition  
4th Evans Stair Running Competition  
5th Evans Stair Running Competition  
960 61st St  
About the project, what you can do, where to find things  
Advice for the Planner  
Analysing the MathLink protocol  
Appendix 1: Living Institutions  
Appendix 1: Living Institutions  
Appendix 3a: FAF Trustee invitation  
APPENDIX 3b: CAC JUDGES INVITATION.  
ART WORKSHOPS  
Arxiv-oai2rdf  
Arxivwiki.org issues  
ASCII  
ASCII art  
ASCII map  
Atlas of subfactors  
Baby  
Baby wishlist  
Balkans  
BAM/PFA Student Committee Manual  
Banana Bread  
Berkeley restaurant guide  
Birthdays  
Blob complex

Bookmarklets  
Books by Jones  
Books to give away  
BUDGET  
Bugaboos finances  
Buildering  
Building an arxivwiki server  
But medical records are kept at the doctors office not her house 4  
CAC Rules and Regulations  
Career Day  
Carrot Soup  
Cell Phone  
Cheese Fondue  
Cheese/Garlic Biscuits  
Chem  
Chocolate Guinness Cake  
Christmas list  
Classifying Spinor Structures (abstract)  
Climbing the Prow  
Code snippets  
Cologne visit  
Computing properties of graphs  
CONCLUSION  
Conferences  
Conjectures  
Conversations with Dror  
Cookbook  
Copy of Yossi's planning page for the Nose  
Corn Bread  
Craigs List Ad Draft  
Cranberry Bread  
Current events  
CURRENT PROJECTS  
Dahi Batata Puri  
Doi hashes  
DOI hashes  
Eggplant Dengaku  
Electronic money  
Entering a (graded bipartite) graph  
Eu tbm lembrei dele quando escrevi KKKKKK nós ama lentes online iNqual uu MentiraEuAmoMais  
Find background information on current events  
FOAF  
Foo  
Frankland Range  
Future trips  
Gear list  
Generators and Relations Talk 1  
Generators and Relations Talk 2  
Getting started hacking Tungsten  
Give away  
Gnocchi  
Google Print  
Google Print competitors  
Google Print copyright issues  
Google Print development history  
Google Print extensions  
Google Print extensions announcement  
Google Print extenstions announcement  
Google Print hacking  
Google Print hacking history

Google Print Library Project  
Google Print Recorder  
Google Print right click enabling  
Go to a page with this exact name if exists  
Graded graphs  
Graph planar algebra  
Graph planar algebra  
Hacking the parser  
Half Dome with Vlad  
Hectare  
Hexadecimal  
Homework  
Homework 1  
Homework 2  
Homework 3  
House  
House Meeting Oct 2005  
Housing leads  
How to find the Dipsea Trail  
Important Dates  
Inaugural Evans Stair Running Competition  
INTRODUCTION  
Java Security API  
Jobs  
Jones-Wenzl projections in the Temperley-Lieb algebra  
Juggling for Geeks (abstract)  
Jython subclasses can't call Java super constructors  
Kate and Simon's trip list  
Kearsarge Pass to Whitney Portal  
Khovanov Homology in Austin  
Khovanov Workshop  
Kingdom (biology)  
Laksa  
Laksa paste  
Leaning Tower  
Leaning Tower Trip Report  
Letter of invitation to groups  
Links  
List of all special pages [q]  
List of countries  
List of things we need  
Load a random page [x]  
Local Khovanov homology Talk 1  
Local Khovanov homology Talk 2  
Local spam blacklist  
Lost luggage  
Love when talks about Mayweather Pacman 18  
Main Page  
Maneki Neko  
Math  
Math121a  
Math 53 Fall 2005 Quizzes  
Math 54 Fall 2005  
Math 54 Fall 2005 Homework Solutions  
Math 54 Fall 2005 Midterm reviews  
Math 54 Fall 2005 Office Hours  
Math 54 Fall 2005 Quizzes  
Math dictionary  
Mathematics  
Maths

Matt's bed  
Maui 2011  
Meta:Help:Special characters  
Metaphor/scratch/Example  
M:Help:Formula  
M:Help:Table  
M:Help:Variable  
Midterm 2 Review Problems  
Mills Real Analysis  
Move the student manual  
Moving  
MP s broadband campaign roll out boost The Westmorland Gazette 82  
Multiplying matrices properly  
Music list  
Neighbourhood fruit  
Nem sabia Furnished Apartments que vc era timãã mas tamo JUNTO PORRA  
Nina's broken computer  
Nina's grad school stuff  
Nina's Journal  
NinaToNJ  
NJToNJ  
Nose  
Notes culled from the SuperTopo forum  
Notes for my talk  
Old homework assignments  
Old housing leads  
Oldmathpapers.org issues tracker  
Open Problems in Fusion Categories  
Orange-Lemon Cake  
Otra semana business speakers que empieza y sigue injustamente detrás de las rejas Libertad  
Performers  
PGP public key  
Phone numbers  
Phone Numbers  
Photos  
PLANNING OVERVIEW  
Precipitous Bluff  
Previous teaching  
PUBLICITY  
Public transport  
Qual questions  
Quantum Topology conference  
Quantum Topology in Auckland  
Quantum Topology in Wellington  
Ranen's webpage  
Recipes  
Reembedding  
Renaming Tungsten  
Restricted pages example  
Road Trip Food Guide  
Road Trip Food Guide/Berkeley  
Road Trip Food Guide/Central Valley  
Road Trip Food Guide/CVS  
Road Trip Food Guide/Highway 120  
Road Trip Food Guide/Highway 4  
Road Trip Food Guide/Highway 50  
Road Trip Food Guide/Highway 80  
Road Trip Food Guide/Highway 88  
Road Trip Food Guide/North Lake Tahoe  
Road Trip Food Guide/North Shore

Road Trip Food Guide/Reno  
Road Trip Food Guide/South Lake Tahoe  
Road Trip Food Guide/Way North  
RT Talk is cheap 6  
Russian Tea Cakes  
Salathe planning  
Sandbox  
Sandbox/wikilink1  
Sandbox/wikilink2  
Sandbox/wikilink3  
Sandbox/wikilink4  
Santa Barbara House Hunting  
Scones  
ScottWiki:About  
ScottWiki Atom Feed  
ScottWiki (en)  
ScottWiki:General disclaimer  
ScottWiki:Privacy policy  
ScottWiki RSS Feed  
Scraping the arXiv  
Search Inside The Book hacking  
Search ScottWiki [f]  
Search the pages for this text  
Setting up a Derby server  
Setting up katlas  
SITB page limit  
Social calendar  
Spam morning of, October 30  
Special:AllPages  
Student spam morning of, October 30  
Student spam, October 28  
Su 3-referee  
Subfactors in Maui  
Subfactor Ski Retreat  
Summer 2005 trips proposals  
Sydney Christmas 2005  
Taipa  
Taisho Chic  
Talk:960 61st St  
Talk:About  
Talk:ASCII map  
Talk:BAM/PFA Student Committee Manual  
Talk:Berkeley restaurant guide  
Talk:Birthdays  
Talk:Community Portal  
Talk:Electronic money  
Talk:General disclaimer  
Talk:Google Print  
Talk:Google Print Recorder  
Talk:Hacked by FatalDevil  
Talk:Java Security API  
Talk:List of things we need  
Talk:Main Page  
Talk:MaL  
Talk:Math 54 Fall 2005  
Talk:Mathematics  
Talk:Mills Real Analysis  
Talk:Move the student manual  
Talk:Nina's grad school stuff  
Talk:Notes for my talk

Talk:PGP public key  
Talk:Photos  
Talk:Previous teaching  
Talk:Recipes  
Talk:Road Trip Food Guide/South Lake Tahoe  
Talk:TARGET EVENTS  
Talk:Tibet trip planning  
Talk:Todo  
Talk:Transient  
Talk:Tungsten  
Talk:Vegan chocolate pecan brownies  
Temperley-Lieb algebras and representation theory (abstract)  
The list of recent changes in the wiki. [r]  
The Nose  
The place to find out.  
This is a special page, you cannot edit the page itself  
Thursdays at MUSE  
Tibet trip planning  
To ask the landlord  
Todo  
Tofu Nuggets  
TQFT mini-course  
Transient  
Transient/Summer-meeting  
Trips  
Trips list  
Trips List  
Tungsten  
Tungsten Dreams And Hopes  
Tungsten installation instructions  
Tungsten session: Map and Range  
Tungsten session: Part and Plus  
Tungsten session: Set and Pattern  
Tungsten tags  
Tungsten Todo List  
Tungsten using Eclipse  
UC Berkeley Buildering  
Upload files [u]  
User:Anping  
User:Benw  
User:Brockert  
User:Drorbn  
User:Nina  
User:ReputationManagement  
User:Scott  
User talk:128.32.173.239  
User talk:131.215.241.26  
User talk:169.229.113.100  
User talk:169.229.140.27  
User talk:69.110.11.59  
User talk:Anping  
User talk:Anping0  
User talk:Anping00  
User talk:Benw  
User talk:CBV  
User talk:CX  
User talk:HFGJ  
User talk:HHGJGH  
User talk:Kaixinde  
User talk:Kaixindea



User talk:Kaixindeab  
 User talk:Kaixindeabc  
 User talk:Kaixindeabcd  
 User talk:Lplp  
 User talk:Scott  
 User talk:Wowo123  
 User talk:Yvonne  
 Vegan chocolate pecan brownies  
 Vfr  
 Virginia Sugar Cookies  
 Visit the Main Page [z]  
 What's Your Theme Song -  
 Who Am I  
 Wikifying the Knot Atlas  
 Wish list  
 Work in progress  
 You are encouraged to log in, it is not mandatory however. [o]  
 Zodiac

**Length[FindAllPages[]]**

571

```
(spam =
  Cases[FindAllPages[], s_ /; StringMatchQ[s, p_ /; StringLength[p] > 40 & StringFreeQ[
    ToLowerCase[p], "about the" | "yossi" | "jones" | "jython" | "temperley" | "evans" |
    "wikilink" | "sydney" | "homework" | "maui" | "math" | "generators" |
    "house" | "local" | "road" | "talk" | "spam"]]]) // TableForm
```

Angry Birds Game - How to Target Villains in Angry Birds  
 Best Annuity Rates UK 2012, Pension Annuities Rates - what is annuity protection  
 Eu tbm lembrei dele quando escrevi KKKKKK nós ama lences online iNqual uu MentiraEuAmoMais  
 Find background information on current events  
 Go to a page with this exact name if exists  
 Hoje a Assembleia Lesgislativa do Estaddo lançou Edital de Convocação p interessados em con  
 Nem sabia Furnished Apartments que vc era timãão mas tamo JUNTO PORRA  
 Otra semana business speakers que empieza y sigue injustamente detrás de las rejas Libertad  
 The list of recent changes in the wiki. [r]  
 This is a special page, you cannot edit the page itself  
 You are encouraged to log in, it is not mandatory however. [o]

**SetDirectory["~/projects/complete/WikiLink/trunk"]**

/Users/scott/projects/complete/WikiLink/trunk

```
Run /@ (("./wikidelete.sh " <> StringReplace[#, " " -> "_"]) & /@ spam)
```

```
{0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 512}
```

```
(spam = DeleteCases[Cases[FindAllPages[], s_ /; ! StringFreeQ[s, "Personal"]], "Jobs"]) //
  TableForm
```

```
{}
```

```
Cases [Take [FindAllPages [], All],
s_ /; ! StringFreeQ [Import ["http://tqft.net/wiki/" <> StringReplace[s, " " -> "_"], "Text"],
"Has Been Moved"]]
```

FetchURL::conopen : The connection to URL http://tqft.net/wiki/Angry\_Birds\_Game\_-\_How\_to\_Target\_Villains\_in\_Angry\_Birds cannot be opened. If the URL is correct, you might need to configure your firewall program, or you might need to set a proxy in the Internet connectivity tab of the Preferences dialog (or by calling SetInternetProxy). For HTTPS connections, you might need to inspect the authenticity of the server's SSL certificate and choose to accept it.

StringFreeQ::strse : String or list of strings expected at position 1 in StringFreeQ[\$Failed, Has Been Moved]. >>

FetchURL::conopen : The connection to URL http://tqft.net/wiki/Best\_Annuity\_Rates\_UK\_2012,\_Pension\_Annuities\_Rates\_-\_what\_is\_annuity\_protection cannot be opened. If the URL is correct, you might need to configure your firewall program, or you might need to set a proxy in the Internet connectivity tab of the Preferences dialog (or by calling SetInternetProxy). For HTTPS connections, you might need to inspect the authenticity of the server's SSL certificate and choose to accept it.

StringFreeQ::strse : String or list of strings expected at position 1 in StringFreeQ[\$Failed, Has Been Moved]. >>

FetchURL::conopen : The connection to URL http://tqft.net/wiki/Eu\_tbm\_lembrei\_dele\_quando\_escrevi\_KKKKKK\_nos\_ama\_lenses\_online\_iNqual\_uu\_MentiraEuAmoMais cannot be opened. If the URL is correct, you might need to configure your firewall program, or you might need to set a proxy in the Internet connectivity tab of the Preferences dialog (or by calling SetInternetProxy). For HTTPS connections, you might need to inspect the authenticity of the server's SSL certificate and choose to accept it.

General::stop : Further output of FetchURL::conopen will be suppressed during this calculation. >>

StringFreeQ::strse : String or list of strings expected at position 1 in StringFreeQ[\$Failed, Has Been Moved]. >>

General::stop : Further output of StringFreeQ::strse will be suppressed during this calculation. >>

```
{Act 1 Insurance, Adult Dating Seattle -, American Flag Free Clipart - talheav,
Anti-intellectualism And School -, Art Career Nursing - www job bank,
Atkins Diet Format -, Black Christian Single Dating, Cannon Beach Castles,
Career Path Seminar - principals jobs, Chili Palmer S Free Story Annex -,
Country Rose Liquor Store -, Dating Pagan -, Dating Rules Teenager Us -,
Discreet Dating Online, Drinking Safe Water Hose For Rv -,
Efiling Income Singapore Tax -, Emergency Chart -, Female Tauren Art -,
Fleet Feet Bend Oregon -, How To Make Benches From Headboards, Icp Chat Room,
Investment Property Usa -, Irving Kushner -, Learn To Use A Computer Keyboard -,
Luminescent Bay Puerto Rice Pictures, Mexican Art Wedding Favors -, Mimosa Torrent -,
Norwood Yard Ohio -, Online Chat For Dating -, Savior Easter Music -,
South Beach Diet Restaurant -, The Names Of King Arhtur -, Thomas Scalea -,
Top Philippine Chat Sites, Vince Martz, What's Your Theme Song -}
```

```
spam = %;
```

```
Run [%[[1]]]
```

```
0
```

```
Run [
  "open http://tqft.net/wiki/You_ll_figure_it_out_You_hve_a_law_degree_You_can_do_it_51"]
0

LocatePagesByContent[content_] :=
  Cases[WikiStringCases[FindAllPages[], content], {title_String, {__}} => title]

LocatePagesByContent["Omath"]
{}

LocatePagesByContent["<div" ~~ ShortestMatch[___] ~~ "/div>"]
LocatePagesByContent[<div ~~ ShortestMatch[___] ~~ /div>]

LocatePagesByUserAndContent[user_String, spamPattern_] :=
  Union[Cases[WikiStringCases[AllUserContributions[user], spamPattern],
    {title_String, {__}} => title]]

LocatePagesByUserAndContent["Kuza22", "wikitiki"]
{}
```