

# Semantic Web at Lilly

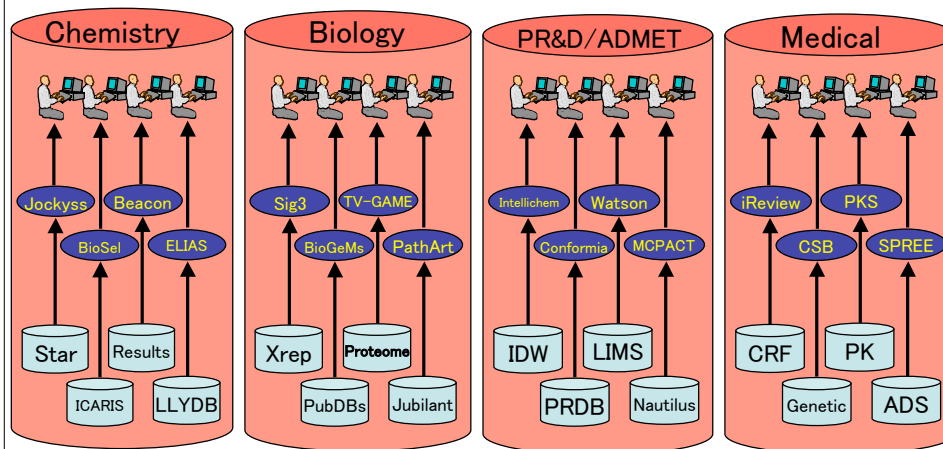
C-SHALS 2008

Ernst R. Dow, PhD  
dow@lilly.com  
Senior Information Consultant  
Discovery Informatics  
Eli Lilly and Company

## Challenges

- Existing drug discovery and development process cost too much to be sustainable in current climate
- Pharmaceutical research is a data intensive endeavor
- Data is being generated at an increasing rate
  - Difficult or impossible for any scientist to know all the sources – scientists asked to work more outside their own areas
  - Nucleic Acids Research, DB issue
    - 1078 databases, 110 more than last year
    - links to more than 80 databases have been updated
    - only 25 obsolete databases have been removed
  - Multiple ways of describing the same or similar data (same or similar depends on point of view)
    - MESH, PathArt disease, PharmaProjects indications, gene ontology, IDDB3 Pharmacology
    - How does one query across overlapping data?

## Where are Internal Data? Silos of Silos



- Tools, application, and data are standalone with limited interaction
- Scientists have great difficulty finding their data and associated tools
- Asking cross-domain questions ( e.g. Discovery + Medical ) very difficult
- Support becoming very impractical - estimated **400+** individual tools across silos
- Larger problem in older companies and regulated industries

## Data are generated faster than they can be understood

- Must find data that are relevant
  - tremendous duplication
  - what is the current answer?
  - wheat from chaff
- Find connections in data
  - visualization
  - words
  - statistics
- Difficulty measuring value of data, e.g. compare to compute speed
  - database quality
    - database 1 vs. database 2
    - agreement
    - quality measure of each element
- Data curation is expensive
- More than just having the data: ability to retrieve relevant decision-making information must be part of the value metric

# How do we address?

- Use Discovery Target Assessment Tool (DTAT)
  - *DTAT allows scientists to evaluate drug targets. DTAT allows scientists to select the scientific question of interest and returns data that is in the appropriate context.*
- Uses RDF to store information about targets, pharmacology, internal development, disease
- Built upon Life Science Grid (going open source ca. end of March)
- Plugins use “listeners” to respond to appropriate data type and serve information
- Question framework allows scientists to learn how each data source provides relevant data
  - Questions stay relatively constant, data and sources change.
  - If informatics is doing proper job, we are providing the best answers for the questions.

Show DTAT

The screenshot displays the Lilly Science Grid interface. At the top, there are menu options like 'File', 'Applications', 'Tools', and 'Help'. Below that, a search bar is visible with the text 'GLP1R' entered. The main area shows a complex network diagram with 'Diabetes, Type II' as a central node. Other nodes include 'Diabetes, general', 'Heart failure', 'Obesity', and 'Irritable bowel syndrome'. Various pharmaceutical companies are listed, such as 'Zelenski Pharmaceuticals', 'Novo Nordisk', and 'Eli Lilly'. A callout box labeled '1) enter target' points to the search bar. Another callout box labeled '2) all plugins run' points to the top navigation area. A third callout box labeled '3) select question' points to a list of questions at the bottom of the interface. The questions include: 'What other companies are working on this target?', 'What is the function of this target?', 'Is there a mouse model for this target?', 'Where are the Lilly SNPs for this target?', and 'What are the gene regulatory regions for this target?'. The interface also shows 'Data from PharmaProjects; visualization done by Lilly' and 'Star Trees created with InSight VizServer™'.

**Lilly Science Grid**

Search Type:  Search Text:  Search Options

Found 1: GLPIR

Load Ontology Tree for GLPIR

Public Information | Biological Assets | Lilly Assessments | DTAT Help

Clustered Results

- ▶ safety GLPIR OR "glucagon-like peptide 1 receptor" (264)
  - ▶ Obesity, Antagonists (39)
  - ▶ Modulators (43)
  - ▶ Activity (31)
  - ▶ GLP (28)
  - ▶ Related disorders (23)
  - ▶ Cells (19)
  - ▶ Diabetes (21)
  - ▶ Agonists (25)
  - ▶ Receptor ligands (12)
  - ▶ Inhibitors (20)
  - ▶ Encoding (13)
  - ▶ Pharmaceutical (12)
  - ▶ Gene Expression (7)
  - ▶ Binding (6)
  - ▶ G Protein Coupled Receptor (7)
  - ▶ GLP1R (6)
  - ▶ Ointoline (6)
  - ▶ Amino Acids (4)
  - ▶ Analogs Thereof (3)
  - ▶ Cancer, Breast (4)
  - ▶ More

Web News

Top 264 results retrieved for the query **safety GLPIR OR "glucagon-like peptide 1 receptor"** (Details)

Select/deselect all on this page Selected results: 0

- [Advisory Committee Briefing Document](#) [new window] [frame] [preview] [clustered]
 

Page 1. Advisory Committee Briefing Document Cardiovascular Safety of Rosiglitazone ... Drug Safety and Risk Management Advisory Committee Meeting on July 30, 2007 ...  
 Date: 2007-07-26  
[www.fda.gov/.../2007-4308b1-01-sponsor-background.pdf](http://www.fda.gov/.../2007-4308b1-01-sponsor-background.pdf) - FDA CDER 1
- [Exendin\(9-39\)Amide as a Glucagon-Like Peptide-1 \(GLP-1\) Receptor Antagonist in Humans](#) [new window] [frame] [preview] [clustered]
 

Changed: July 24, 2007  
 Condition: Hyperglycemia  
 Status: Completed  
[clinicaltrials.gov/show/NCT00393445](http://clinicaltrials.gov/show/NCT00393445) - ClinicalTrials.gov 1
- [Proteins and nucleic acids encoding same](#) [new window] [frame] [preview] [clustered]
 

Applicant: 002 CuraGen Corporation  
 Category: C07H  
 Dates: -23-104  
 Patent: US727693B2  
[www.micropat.com/...highlight=&forward\\_url=&patnum=US727693B2&-MicroPatent%201,%20MicroPatent%205](http://www.micropat.com/...highlight=&forward_url=&patnum=US727693B2&-MicroPatent%201,%20MicroPatent%205)
- [Clinical Assessment of GSK716155 for Type 2 Diabetes Mellitus](#) [new window] [frame] [preview] [clustered]
 

Changed: September 14, 2007  
 Condition: Type 2 Diabetes Mellitus  
 Status: Recruiting  
[clinicaltrials.gov/show/NCT00530309](http://clinicaltrials.gov/show/NCT00530309) - ClinicalTrials.gov 2
- [Exendin improves  \$\beta\$ -cell response in subjects with impaired glucose tolerance](#) [new window] [frame] [preview] [clustered]
 

Applicant: S04 AMYLIN PHARMACEUTICALS INC  
 Category: A61K  
 Dates: -23-104

Get It Quick

- What are text mining results for this target?
- What are the changes in the past month?
- What are the toxicology results for this target?
- What are the safety results for this target?**
- What are the biomarker results for this target?

Browser Search My Ontology Get It Quick

**Lilly Science Grid**

Search Type:  Search Text:  Search Options

Found 4:
 

- glucagon-like peptide 1 receptor (GLPIR)
- glucagon-like peptide 2 receptor (GLP2R)
- glucagon (GCS)
- glucagon receptor (GCCR)

Load Ontology Tree for glucagon-like peptide 1 receptor (GLPIR)

Public Information | **Biology** | Biological Assets | Lilly Assessments | Wikis | DTAT Help

Array Based Expression | Literature Expression | Ensembl Viewer | XRef | Gene Correlation | Biomarkers | Isoform Viewer | General

Back Forward Home Current Locale: Corporate

**PGLEnsembl@Lilly Human Contigview** e.g. AL133772.15.1.44776.AL355240.17.1.11242

PGLEnsembl@Lilly release 48 - Dec 2007 HOME - BLAST - SITEMAP - HELP

Your Ensembl

- Login or Register
- About User Accounts

Chromosome 6  
39,123,595 - 39,164,497

- View of Chromosome 6
- Graphical overview
- Resequencing alignment
- View alignment with ...
  - 7 *suspeian mammals*
  - Pecan*
  - 10 *amniota vertebrates*
  - Pecan*
  - Rattus norvegicus*
  - Musca domestica*
  - Loxodonta africana*
  - Echinops telfairi*
  - Oryctolagus cuniculus*
  - Dasyatis novemcinctus*
  - Canis familiaris*
  - Pan troglodytes*
  - Gallus gallus*
  - Ornithorhynchus anatinus*
  - Bos taurus*
  - Felis catus*
  - Tupaia belangeri*
  - Eriacus europaeus*
  - Cavia porcellus*

Overview

Chr. 6 band

58,77 Mb 59,79 Mb 59,79 Mb 59,79 Mb 59,79 Mb

Markers

Ensembl Genes: LPT09, LGL01, LGL02, LGL03, LGL04, LGL05, LGL06, LGL07, LGL08, LGL09, LGL10, LGL11, LGL12, LGL13, LGL14, LGL15, LGL16, LGL17, LGL18, LGL19, LGL20, LGL21, LGL22, LGL23, LGL24, LGL25, LGL26, LGL27, LGL28, LGL29, LGL30, LGL31, LGL32, LGL33, LGL34, LGL35, LGL36, LGL37, LGL38, LGL39, LGL40, LGL41, LGL42, LGL43, LGL44, LGL45, LGL46, LGL47, LGL48, LGL49, LGL50, LGL51, LGL52, LGL53, LGL54, LGL55, LGL56, LGL57, LGL58, LGL59, LGL60, LGL61, LGL62, LGL63, LGL64, LGL65, LGL66, LGL67, LGL68, LGL69, LGL70, LGL71, LGL72, LGL73, LGL74, LGL75, LGL76, LGL77, LGL78, LGL79, LGL80, LGL81, LGL82, LGL83, LGL84, LGL85, LGL86, LGL87, LGL88, LGL89, LGL90, LGL91, LGL92, LGL93, LGL94, LGL95, LGL96, LGL97, LGL98, LGL99, LGL100, LGL101, LGL102, LGL103, LGL104, LGL105, LGL106, LGL107, LGL108, LGL109, LGL110, LGL111, LGL112, LGL113, LGL114, LGL115, LGL116, LGL117, LGL118, LGL119, LGL120, LGL121, LGL122, LGL123, LGL124, LGL125, LGL126, LGL127, LGL128, LGL129, LGL130, LGL131, LGL132, LGL133, LGL134, LGL135, LGL136, LGL137, LGL138, LGL139, LGL140, LGL141, LGL142, LGL143, LGL144, LGL145, LGL146, LGL147, LGL148, LGL149, LGL150, LGL151, LGL152, LGL153, LGL154, LGL155, LGL156, LGL157, LGL158, LGL159, LGL160, LGL161, LGL162, LGL163, LGL164, LGL165, LGL166, LGL167, LGL168, LGL169, LGL170, LGL171, LGL172, LGL173, LGL174, LGL175, LGL176, LGL177, LGL178, LGL179, LGL180, LGL181, LGL182, LGL183, LGL184, LGL185, LGL186, LGL187, LGL188, LGL189, LGL190, LGL191, LGL192, LGL193, LGL194, LGL195, LGL196, LGL197, LGL198, LGL199, LGL200, LGL201, LGL202, LGL203, LGL204, LGL205, LGL206, LGL207, LGL208, LGL209, LGL210, LGL211, LGL212, LGL213, LGL214, LGL215, LGL216, LGL217, LGL218, LGL219, LGL220, LGL221, LGL222, LGL223, LGL224, LGL225, LGL226, LGL227, LGL228, LGL229, LGL230, LGL231, LGL232, LGL233, LGL234, LGL235, LGL236, LGL237, LGL238, LGL239, LGL240, LGL241, LGL242, LGL243, LGL244, LGL245, LGL246, LGL247, LGL248, LGL249, LGL250, LGL251, LGL252, LGL253, LGL254, LGL255, LGL256, LGL257, LGL258, LGL259, LGL260, LGL261, LGL262, LGL263, LGL264, LGL265, LGL266, LGL267, LGL268, LGL269, LGL270, LGL271, LGL272, LGL273, LGL274, LGL275, LGL276, LGL277, LGL278, LGL279, LGL280, LGL281, LGL282, LGL283, LGL284, LGL285, LGL286, LGL287, LGL288, LGL289, LGL290, LGL291, LGL292, LGL293, LGL294, LGL295, LGL296, LGL297, LGL298, LGL299, LGL300, LGL301, LGL302, LGL303, LGL304, LGL305, LGL306, LGL307, LGL308, LGL309, LGL310, LGL311, LGL312, LGL313, LGL314, LGL315, LGL316, LGL317, LGL318, LGL319, LGL320, LGL321, LGL322, LGL323, LGL324, LGL325, LGL326, LGL327, LGL328, LGL329, LGL330, LGL331, LGL332, LGL333, LGL334, LGL335, LGL336, LGL337, LGL338, LGL339, LGL340, LGL341, LGL342, LGL343, LGL344, LGL345, LGL346, LGL347, LGL348, LGL349, LGL350, LGL351, LGL352, LGL353, LGL354, LGL355, LGL356, LGL357, LGL358, LGL359, LGL360, LGL361, LGL362, LGL363, LGL364, LGL365, LGL366, LGL367, LGL368, LGL369, LGL370, LGL371, LGL372, LGL373, LGL374, LGL375, LGL376, LGL377, LGL378, LGL379, LGL380, LGL381, LGL382, LGL383, LGL384, LGL385, LGL386, LGL387, LGL388, LGL389, LGL390, LGL391, LGL392, LGL393, LGL394, LGL395, LGL396, LGL397, LGL398, LGL399, LGL400, LGL401, LGL402, LGL403, LGL404, LGL405, LGL406, LGL407, LGL408, LGL409, LGL410, LGL411, LGL412, LGL413, LGL414, LGL415, LGL416, LGL417, LGL418, LGL419, LGL420, LGL421, LGL422, LGL423, LGL424, LGL425, LGL426, LGL427, LGL428, LGL429, LGL430, LGL431, LGL432, LGL433, LGL434, LGL435, LGL436, LGL437, LGL438, LGL439, LGL440, LGL441, LGL442, LGL443, LGL444, LGL445, LGL446, LGL447, LGL448, LGL449, LGL450, LGL451, LGL452, LGL453, LGL454, LGL455, LGL456, LGL457, LGL458, LGL459, LGL460, LGL461, LGL462, LGL463, LGL464, LGL465, LGL466, LGL467, LGL468, LGL469, LGL470, LGL471, LGL472, LGL473, LGL474, LGL475, LGL476, LGL477, LGL478, LGL479, LGL480, LGL481, LGL482, LGL483, LGL484, LGL485, LGL486, LGL487, LGL488, LGL489, LGL490, LGL491, LGL492, LGL493, LGL494, LGL495, LGL496, LGL497, LGL498, LGL499, LGL500, LGL501, LGL502, LGL503, LGL504, LGL505, LGL506, LGL507, LGL508, LGL509, LGL510, LGL511, LGL512, LGL513, LGL514, LGL515, LGL516, LGL517, LGL518, LGL519, LGL520, LGL521, LGL522, LGL523, LGL524, LGL525, LGL526, LGL527, LGL528, LGL529, LGL530, LGL531, LGL532, LGL533, LGL534, LGL535, LGL536, LGL537, LGL538, LGL539, LGL540, LGL541, LGL542, LGL543, LGL544, LGL545, LGL546, LGL547, LGL548, LGL549, LGL550, LGL551, LGL552, LGL553, LGL554, LGL555, LGL556, LGL557, LGL558, LGL559, LGL560, LGL561, LGL562, LGL563, LGL564, LGL565, LGL566, LGL567, LGL568, LGL569, LGL570, LGL571, LGL572, LGL573, LGL574, LGL575, LGL576, LGL577, LGL578, LGL579, LGL580, LGL581, LGL582, LGL583, LGL584, LGL585, LGL586, LGL587, LGL588, LGL589, LGL590, LGL591, LGL592, LGL593, LGL594, LGL595, LGL596, LGL597, LGL598, LGL599, LGL600, LGL601, LGL602, LGL603, LGL604, LGL605, LGL606, LGL607, LGL608, LGL609, LGL610, LGL611, LGL612, LGL613, LGL614, LGL615, LGL616, LGL617, LGL618, LGL619, LGL620, LGL621, LGL622, LGL623, LGL624, LGL625, LGL626, LGL627, LGL628, LGL629, LGL630, LGL631, LGL632, LGL633, LGL634, LGL635, LGL636, LGL637, LGL638, LGL639, LGL640, LGL641, LGL642, LGL643, LGL644, LGL645, LGL646, LGL647, LGL648, LGL649, LGL650, LGL651, LGL652, LGL653, LGL654, LGL655, LGL656, LGL657, LGL658, LGL659, LGL660, LGL661, LGL662, LGL663, LGL664, LGL665, LGL666, LGL667, LGL668, LGL669, LGL670, LGL671, LGL672, LGL673, LGL674, LGL675, LGL676, LGL677, LGL678, LGL679, LGL680, LGL681, LGL682, LGL683, LGL684, LGL685, LGL686, LGL687, LGL688, LGL689, LGL690, LGL691, LGL692, LGL693, LGL694, LGL695, LGL696, LGL697, LGL698, LGL699, LGL700, LGL701, LGL702, LGL703, LGL704, LGL705, LGL706, LGL707, LGL708, LGL709, LGL710, LGL711, LGL712, LGL713, LGL714, LGL715, LGL716, LGL717, LGL718, LGL719, LGL720, LGL721, LGL722, LGL723, LGL724, LGL725, LGL726, LGL727, LGL728, LGL729, LGL730, LGL731, LGL732, LGL733, LGL734, LGL735, LGL736, LGL737, LGL738, LGL739, LGL740, LGL741, LGL742, LGL743, LGL744, LGL745, LGL746, LGL747, LGL748, LGL749, LGL750, LGL751, LGL752, LGL753, LGL754, LGL755, LGL756, LGL757, LGL758, LGL759, LGL760, LGL761, LGL762, LGL763, LGL764, LGL765, LGL766, LGL767, LGL768, LGL769, LGL770, LGL771, LGL772, LGL773, LGL774, LGL775, LGL776, LGL777, LGL778, LGL779, LGL780, LGL781, LGL782, LGL783, LGL784, LGL785, LGL786, LGL787, LGL788, LGL789, LGL790, LGL791, LGL792, LGL793, LGL794, LGL795, LGL796, LGL797, LGL798, LGL799, LGL800, LGL801, LGL802, LGL803, LGL804, LGL805, LGL806, LGL807, LGL808, LGL809, LGL810, LGL811, LGL812, LGL813, LGL814, LGL815, LGL816, LGL817, LGL818, LGL819, LGL820, LGL821, LGL822, LGL823, LGL824, LGL825, LGL826, LGL827, LGL828, LGL829, LGL830, LGL831, LGL832, LGL833, LGL834, LGL835, LGL836, LGL837, LGL838, LGL839, LGL840, LGL841, LGL842, LGL843, LGL844, LGL845, LGL846, LGL847, LGL848, LGL849, LGL850, LGL851, LGL852, LGL853, LGL854, LGL855, LGL856, LGL857, LGL858, LGL859, LGL860, LGL861, LGL862, LGL863, LGL864, LGL865, LGL866, LGL867, LGL868, LGL869, LGL870, LGL871, LGL872, LGL873, LGL874, LGL875, LGL876, LGL877, LGL878, LGL879, LGL880, LGL881, LGL882, LGL883, LGL884, LGL885, LGL886, LGL887, LGL888, LGL889, LGL890, LGL891, LGL892, LGL893, LGL894, LGL895, LGL896, LGL897, LGL898, LGL899, LGL900, LGL901, LGL902, LGL903, LGL904, LGL905, LGL906, LGL907, LGL908, LGL909, LGL910, LGL911, LGL912, LGL913, LGL914, LGL915, LGL916, LGL917, LGL918, LGL919, LGL920, LGL921, LGL922, LGL923, LGL924, LGL925, LGL926, LGL927, LGL928, LGL929, LGL930, LGL931, LGL932, LGL933, LGL934, LGL935, LGL936, LGL937, LGL938, LGL939, LGL940, LGL941, LGL942, LGL943, LGL944, LGL945, LGL946, LGL947, LGL948, LGL949, LGL950, LGL951, LGL952, LGL953, LGL954, LGL955, LGL956, LGL957, LGL958, LGL959, LGL960, LGL961, LGL962, LGL963, LGL964, LGL965, LGL966, LGL967, LGL968, LGL969, LGL970, LGL971, LGL972, LGL973, LGL974, LGL975, LGL976, LGL977, LGL978, LGL979, LGL980, LGL981, LGL982, LGL983, LGL984, LGL985, LGL986, LGL987, LGL988, LGL989, LGL990, LGL991, LGL992, LGL993, LGL994, LGL995, LGL996, LGL997, LGL998, LGL999, LGL1000, LGL1001, LGL1002, LGL1003, LGL1004, LGL1005, LGL1006, LGL1007, LGL1008, LGL1009, LGL1010, LGL1011, LGL1012, LGL1013, LGL1014, LGL1015, LGL1016, LGL1017, LGL1018, LGL1019, LGL1020, LGL1021, LGL1022, LGL1023, LGL1024, LGL1025, LGL1026, LGL1027, LGL1028, LGL1029, LGL1030, LGL1031, LGL1032, LGL1033, LGL1034, LGL1035, LGL1036, LGL1037, LGL1038, LGL1039, LGL1040, LGL1041, LGL1042, LGL1043, LGL1044, LGL1045, LGL1046, LGL1047, LGL1048, LGL1049, LGL1050, LGL1051, LGL1052, LGL1053, LGL1054, LGL1055, LGL1056, LGL1057, LGL1058, LGL1059, LGL1060, LGL1061, LGL1062, LGL1063, LGL1064, LGL1065, LGL1066, LGL1067, LGL1068, LGL1069, LGL1070, LGL1071, LGL1072, LGL1073, LGL1074, LGL1075, LGL1076, LGL1077, LGL1078, LGL1079, LGL1080, LGL1081, LGL1082, LGL1083, LGL1084, LGL1085, LGL1086, LGL1087, LGL1088, LGL1089, LGL1090, LGL1091, LGL1092, LGL1093, LGL1094, LGL1095, LGL1096, LGL1097, LGL1098, LGL1099, LGL1100, LGL1101, LGL1102, LGL1103, LGL1104, LGL1105, LGL1106, LGL1107, LGL1108, LGL1109, LGL1110, LGL1111, LGL1112, LGL1113, LGL1114, LGL1115, LGL1116, LGL1117, LGL1118, LGL1119, LGL1120, LGL1121, LGL1122, LGL1123, LGL1124, LGL1125, LGL1126, LGL1127, LGL1128, LGL1129, LGL1130, LGL1131, LGL1132, LGL1133, LGL1134, LGL1135, LGL1136, LGL1137, LGL1138, LGL1139, LGL1140, LGL1141, LGL1142, LGL1143, LGL1144, LGL1145, LGL1146, LGL1147, LGL1148, LGL1149, LGL1150, LGL1151, LGL1152, LGL1153, LGL1154, LGL1155, LGL1156, LGL1157, LGL1158, LGL1159, LGL1160, LGL1161, LGL1162, LGL1163, LGL1164, LGL1165, LGL1166, LGL1167, LGL1168, LGL1169, LGL1170, LGL1171, LGL1172, LGL1173, LGL1174, LGL1175, LGL1176, LGL1177, LGL1178, LGL1179, LGL1180, LGL1181, LGL1182, LGL1183, LGL1184, LGL1185, LGL1186, LGL1187, LGL1188, LGL1189, LGL1190, LGL1191, LGL1192, LGL1193, LGL1194, LGL1195, LGL1196, LGL1197, LGL1198, LGL1199, LGL1200, LGL1201, LGL1202, LGL1203, LGL1204, LGL1205, LGL1206, LGL1207, LGL1208, LGL1209, LGL1210, LGL1211, LGL1212, LGL1213, LGL1214, LGL1215, LGL1216, LGL1217, LGL1218, LGL1219, LGL1220, LGL1221, LGL1222, LGL1223, LGL1224, LGL1225, LGL1226, LGL1227, LGL1228, LGL1229, LGL1230, LGL1231, LGL1232, LGL1233, LGL1234, LGL1235, LGL1236, LGL1237, LGL1238, LGL1239, LGL1240, LGL1241, LGL1242, LGL1243, LGL1244, LGL1245, LGL1246, LGL1247, LGL1248, LGL1249, LGL1250, LGL1251, LGL1252, LGL1253, LGL1254, LGL1255, LGL1256, LGL1257, LGL1258, LGL1259, LGL1260, LGL1261, LGL1262, LGL1263, LGL1264, LGL1265, LGL1266, LGL1267, LGL1268, LGL1269, LGL1270, LGL1271, LGL1272, LGL1273, LGL1274, LGL1275, LGL1276, LGL1277, LGL1278, LGL1279, LGL1280, LGL1281, LGL1282, LGL1283, LGL1284, LGL1285, LGL1286, LGL1287, LGL1288, LGL1289, LGL1290, LGL1291, LGL1292, LGL1293, LGL1294, LGL1295, LGL1296, LGL1297, LGL1298, LGL1299, LGL1300, LGL1301, LGL1302, LGL1303, LGL1304, LGL1305, LGL1306, LGL1307, LGL1308, LGL1309, LGL1310, LGL1311, LGL1312, LGL1313, LGL1314, LGL1315, LGL1316, LGL1317, LGL1318, LGL1319, LGL1320, LGL1321, LGL1322, LGL1323, LGL1324, LGL1325, LGL1326, LGL1327, LGL1328, LGL1329, LGL1330, LGL1331, LGL1332, LGL1333, LGL1334, LGL1335, LGL1336, LGL1337, LGL1338, LGL1339, LGL1340, LGL1341, LGL1342, LGL1343, LGL1344, LGL1345, LGL1346, LGL1347, LGL1348, LGL1349, LGL1350, LGL1351, LGL1352, LGL1353, LGL1354, LGL1355, LGL1356, LGL1357, LGL1358, LGL1359, LGL1360, LGL1361, LGL1362, LGL1363, LGL1364, LGL1365, LGL1366, LGL1367, LGL1368, LGL1369, LGL1370, LGL1371, LGL1372, LGL1373, LGL1374, LGL1375, LGL1376, LGL1377, LGL1378, LGL1379, LGL1380, LGL1381, LGL1382, LGL1383, LGL1384, LGL1385, LGL1386, LGL1387, LGL1388, LGL1389, LGL1390, LGL1391, LGL1392, LGL1393, LGL1394, LGL1395, LGL1396, LGL1397, LGL1398, LGL1399, LGL1400, LGL1401, LGL1402, LGL1403, LGL1404, LGL1405, LGL1406, LGL1407, LGL1408, LGL1409, LGL1410, LGL1411, LGL1412, LGL1413, LGL1414, LGL1415, LGL1416, LGL1417, LGL1418, LGL1419, LGL1420, LGL1421, LGL1422, LGL1423, LGL1424, LGL1425, LGL1426, LGL1427, LGL1428, LGL1429, LGL1430, LGL1431, LGL1432, LGL1433, LGL1434, LGL1435, LGL1436, LGL1437, LGL1438, LGL1439, LGL1440, LGL1441, LGL1442, LGL1443, LGL1444, LGL1445, LGL1446, LGL1447, LGL1448, LGL1449, LGL1450, LGL1451, LGL1452, LGL1453, LGL1454, LGL1455, LGL1456, LGL1457, LGL1458, LGL1459, LGL1460, LGL1461, LGL1462, LGL1463, LGL1464, LGL1465, LGL1466, LGL1467, LGL1468, LGL1469, LGL1470, LGL1471, LGL1472, LGL1473, LGL1474, LGL1475, LGL1476, LGL1477, LGL1478, LGL1479, LGL1480, LGL1481, LGL1482, LGL1483, LGL1484, LGL1485, LGL1486, LGL1487, LGL1488, LGL1489, LGL1490, LGL1491, LGL1492, LGL1493, LGL1494, LGL1495, LGL1496, LGL1497, LGL1498, LGL1499, LGL1500, LGL1501, LGL1502, LGL1503, LGL1504, LGL1505, LGL1506, LGL1507, LGL1508, LGL1509, LGL1510, LGL1511, LGL1512, LGL1513, LGL1514, LGL1515, LGL1516, LGL1517, LGL1518, LGL1519, LGL1520, LGL1521, LGL1522, LGL1523, LGL1524, LGL1525, LGL1526, LGL1527, LGL1528, LGL1529, LGL1530, LGL1531, LGL1532, LGL1533, LGL1534, LGL1535, LGL1536, LGL1537, LGL1538, LGL1539, LGL1540, LGL1541, LGL1542, LGL1543, LGL1544, LGL1545, LGL1546, LGL1547, LGL1548, LGL1549, LGL1550, LGL1551, LGL1552, LGL1553, LGL1554, LGL1555, LGL1556, LGL1557, LGL1558, LGL1559, LGL1560, LGL1561, LGL1562, LGL1563, LGL1564, LGL1565, LGL1566, LGL1567, LGL1568, LGL1569, LGL1570, LGL1571, LGL1572, LGL1573, LGL1574, LGL1575, LGL1576, LGL1577, LGL1578, LGL1579, LGL1580, LGL1581, LGL1582, LGL1583, LGL1584, LGL1585, LGL1586, LGL1587, LGL1588, LGL1589, LGL1590, LGL1591, LGL1592, LGL1593, LGL1594, LGL1595, LGL1596, LGL1597, LGL1598, LGL1599, LGL1600, LGL1601, LGL1602, LGL1603, LGL1604, LGL1605, LGL1606, L

**Lilly Science Grid**

File Applications Tools Help

Browse

**Public Information** | Biology | Biological Assets | Lilly Assessments | GVKBMPugin01

IDDB3 Viewer | Compare Targets | Vivisimo Viewer | Isoform Viewer | PharmaProjects Tree | PharmaProjects Viewer

| Drug Name   | Originator  | Edit Date  | World Status            | Therapy           |
|---|---|------------|-------------------------|-------------------|
| BAL-AR-18   | Laboratoires Baldacci Sp...   | 10/12/2001 | No Development Reported | Diabetes mellitus |
| insulin (controlled release, Biosphere), Bioglan                              | Bioglan Pharmaceuticals CoBradley Pharmaceuticals IncOriginator developing and marketing own product    | 10/14/2004 | Discontinued            | Diabetes mellitus |
| oral insulin, Biota   | Biota Holdings LtdBiota H...  | 1/27/2003  | No Development Reported | Diabetes mellitus |
| insulin lispro  | Eli Lilly & CoEli Lilly & CoOriginator developing and marketing own product                             | 11/22/2006 | Launched                | Diabetes mellitus |
| TP-107  | TheraPei Pharmaceutical...  | 11/17/2006 | Discovery               | Diabetes mellitus |
| sipoglitazar  | Takeda Pharmaceutical Co LtdTakeda Pharmaceutical Co LtdOriginator developing and marketing own product | 9/19/2006  | Discontinued            | Diabetes mellitus |
| emertamine  | Takeda Pharmaceutical...  | 10/5/2004  | Discontinued            | Diabetes mellitus |
| glycogen synthase kinase-3 (GSK3) inhibitors (diabetes), Boehringer Ingelheim | Boehringer Ingelheim CorpBoehringer Ingelheim CorpOriginator developing and marketing own product       | 6/23/2006  | Discovery               | Diabetes mellitus |
| BM-13677  | Boehringer Mannheim G...  | 4/4/2002   | Discontinued            | Diabetes mellitus |
| SGLT-2 inhibitor (diabetes), CoBristol-Myers Squibb                           | Bristol-Myers Squibb CoBristol-Myers Squibb   | 1/10/2006  | Phase 1 Clinical        | Diabetes mellitus |

**Hypoglycemic agent** - Showing 392 records.

Get It Quick

What drugs are being studied for this pharmacology?  
What are text mining results for this pharmacology?

My Ontology | Get It Quick

**Lilly Science Grid**

File Applications Tools Help

Browse

DHTs | IDDB3\_Pharmacology | **MESH**

Biological Sciences | Chemicals and Drugs | Diseases

Animal Diseases | Bacterial Infections and Mycoses | Cardiovascular Diseases | Congenital, Hereditary, and Neonatal Diseases and Disorders of Environmental Origin | Endocrine System Diseases

ACP1 | Adrenal Gland Diseases | Diabetes Mellitus | Dwarfism | Endocrine Gland Neoplasms | Gonadal Disorders | Parathyroid Diseases | Pituitary Diseases | Polyendocrinopathies, Autoimmune | Thyroid Diseases

Pathways

Acquired Immuno Deficiency Syndrome | Alzheimers | Arrhythmia | Asthma | Atherosclerosis | Bipolar Disorder | Breast Cancer | Chronic Myeloid Leukemia | Chronic Obstructive Pulmonary Disease | Chronic Obstructive Pulmonary Disease (Chronic Bronchitis) | Chronic Obstructive Pulmonary Disease (Emphysema) | Colon Cancer | Crohns Disease | Depression | Diabetes Type II

ACAC Mediated Pathway | Acetylcholine Signaling Pathway | Adiponectin Signaling Pathway | Angiotensin Signaling Pathway | Beta3-Adrenergic Receptor Mediated Pathway | Caveolin1 Mediated Pathway

Links | Ontology | Pathways | Ref Seq | Gene Rif

| Function  | Description  | Reference                  | Evidence                            | Citation                |
|-----------|--|----------------------------|-------------------------------------|-------------------------|
| Function  | G-protein coupled receptor activity                  | <a href="#">GO:0004930</a> | inferred from electronic annotation |                         |
| Function  | glucagon receptor activity                           | <a href="#">GO:0004937</a> | inferred from electronic annotation |                         |
| Function  | peptide receptor activity, G-protein coupled         | <a href="#">GO:0008528</a> | inferred from electronic annotation |                         |
| Process   | G-protein coupled receptor protein signaling pathway | <a href="#">GO:0007186</a> | inferred from electronic annotation |                         |
| Process   | adenylate cyclase activation                         | <a href="#">GO:0007190</a> | traceable author statement          | <a href="#">8405712</a> |
| Process   | cAMP-mediated signaling                              | <a href="#">GO:0018933</a> | inferred from direct assay          | <a href="#">7589461</a> |
| Process   | elevation of cytosolic calcium ion concentration     | <a href="#">GO:0007204</a> | inferred from direct assay          | <a href="#">7589461</a> |
| Process   | learning and/or memory                               | <a href="#">GO:0007611</a> | inferred from electronic annotation |                         |
| Process   | regulation of heart contraction                      | <a href="#">GO:0008016</a> | inferred from electronic annotation |                         |
| Process   | response to stress                                   | <a href="#">GO:0006950</a> | inferred from electronic annotation |                         |
| Process   | signal transduction                                  | <a href="#">GO:0007185</a> | inferred from electronic annotation |                         |
| Component | integral to membrane                                 | <a href="#">GO:0016021</a> | traceable author statement          | <a href="#">8405712</a> |

## The Theory

|                             |   |                        |
|-----------------------------|---|------------------------|
| IDDB3                       |   | PharmaProjects         |
| • drug name                 | → | • drug name            |
| • synonym(s)                |   | • synonym(s)           |
| • originator(s)             | → | • originator(s)        |
| • world status(s)           |   | • world status(s)      |
| • therapy(s)                | → | • therapy(s)           |
| • pharmacology(s) structure | → | • pharmacology(s) flat |
|                             |   | • target(s)            |

Make a database and match on these fields and we are done!

### Reality

- Now we can select a drug name and both databases will respond
  - but drug names are often worded differently or disagree
  - same issues with originator, world status, therapy, pharmacology
- DTAT is target focused, but IDDB3 doesn't have target!
  - We can select a target in PharmaProjects, use pharmacology to find the matching records in IDDB3
  - Works about 30% of the time. All of these fields are text fields and can be different
  - currently use Lilly curated target mapping to pharmacology
- We can spend the human resources and do the mapping, put in the database and we are done!
- But these databases are updated so our database is out of date
- Updates often conflict with what has been established in the DB. Which is correct?

## Why not use a relational DB?

- RDF allows data to be stored and queried without first requiring a schema
  - In a relational model, the structure has to be agreed upon first before any data can ever be stored. With RDF, the triple construct is simple and allows you to store all data immediately.
- RDF is more agile and allows for incremental changes
  - Additional data and/or attributes can be added quickly with minimal impact. Adding new relationships is nothing more than adding a new triple. None of the existing structure or data is changed. In relational, adding new relationships requires that the structure/schema of the database be changed.
- The RDF triple is standard across all implementations
  - Relational database representations are typically unique to the designer (i.e. each company's EMPLOYEE domain will have it's own representation of tables, columns, and relationships). This requires custom coding and renders interoperability virtually impossible. However, a triple is the simplest of structure (i.e. subject, predicate, object) and all RDF uses it.
- Metadata is stored in the RDF
  - The attributes of a given piece of information is contained in the RDF itself. Therefore, the ability to identify these attributes and explore the data is much simpler than having to understand the relational organization of the data
- RDF is intended to be more expressive than relational
- A triple is designed to be "sentence-like" and relates data in a straight-forward way
- Compound example, LSG can listen to new type

Thanks to Phil Brooks


### Question: What is the penalty for conflicting data?

- What if 2 or more databases disagree? Do we need to fix them? Can we live with the ambiguity?
- Can we simply show the scientist the similar data and let the scientist make the call that is appropriate for that time?
- Would the scientist be motivated to fix the answer? Issues around how scientists could change the RDF store? Need to make it very easy and have an immediate benefit.
- Estimate that over 90% of the data will not be used in the database, so we shouldn't spend the money / time to resolve conflicts.

### “Ideal”

- Ability to purchase high quality triples
- Ability to “update” triples
- Ability to show conflicting data in an intuitive manner
- Realization that not all conflicting data is worth resolving – agree to disagree
- Understanding of how recent data fits into existing knowledge – most beneficial to experts
- Scientists directly contribute to the RDF

# Life Science Grid

- LSG – an asynchronous web services (message oriented) “smart” client-side application deployed using Microsoft ClickOnce deployment strategy.
  - Software Development: Microsoft Visual Studio 2005
  - Client: Windows XP SP2, .NET Framework 2.0, WSE 3.0
  - Server: Windows 2003 Enterprise Edition, SP1, .NET Framework 2.0 and IIS 6
  - Databases Supported:
    - MySQL 5.0
    - Microsoft SQL Server 2005 Express Edition
    - Oracle Database 10g Express Edition
- 
- Framework to be released under the BSD license
  - Framework will include sample public domain plugins
  - Documentation “how to” for software developers

# Conclusions

- Lilly is using RDF as part of our scientific solutions
- Life Science Grid is being open sourced
- Lilly is interested in collaborations in this area

# Abstract

- Pharmaceutical researchers are facing tremendous challenges in being able to answer interesting scientific questions in drug discovery. This is because the knowledge that they seek is typically held across many disparate data stores that are managed by a range of different laboratories and organizations. Integrating the data stores is difficult because they use different file formats, identifiers, terminologies, and data types. They also contain data of differing quality based on the study design and the experimental methodology used to generate the data. These challenges are compounded by the fast rate of growth of data due to automation, new databases due to novel techniques, duplication of data, and overlapping data and concepts. It is imperative that scientists can easily identify data sources of interest, pull them together in interesting ways, and interact with the results intuitively. In order to be meaningful for a domain expert it is necessary to have capabilities that would allow a scientist to view new data or data that has been changed. This presentation will describe how Semantic Web technologies are playing a role in addressing some of these scientific requirements at Lilly.