

Artificial Intelligence with SQL WorkLoadExpert (WLX)

Roy Boxwell SEGUS INC PSP #11







Agenda



• What is Artificial Intelligence?

Db2 13 and Data Insights - Al for free?



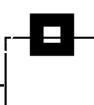
SEG and Al



Q&A



Agenda



• What is Artificial Intelligence?

Db2 13 and Data Insights - Al for free?



SEG and Al



Q&A



It is *the* buzzword of the moment!

Everyone has heard of, and probably already used, ChatGPT, MidJourney, Prezo, etc. etc.



According to industry gurus it will make a lot of people very rich and/or unemployed.



There are lots of web sites out there where you can play with it and see what it can do (or not do!)





Sadly, it is all bogus!

There is no such thing as AI... however, wherever you look, "There it is!"



Why is this?

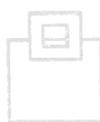
We, as programmers, are all acutely aware that it is a predictive program for the next word based on a massive input of 36 Billion data points. It looks and feels like an AI but all it is doing is <u>guessing</u> at what comes next based only on how it happened before.





With pictures it gets even better or worse depending on how you look at it!

I asked some of the leading graphic AIs to generate pictures for me based upon the keywords:



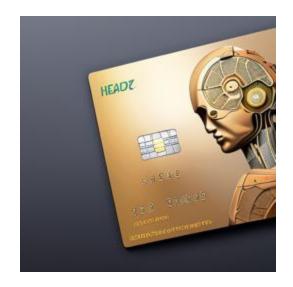
AI COBOL Credit card

What did I get?





This is not too bad:



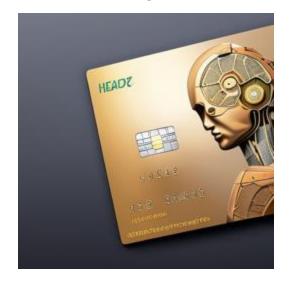








Now it gets a bit wobbly:

































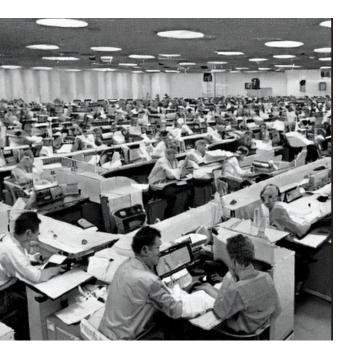


















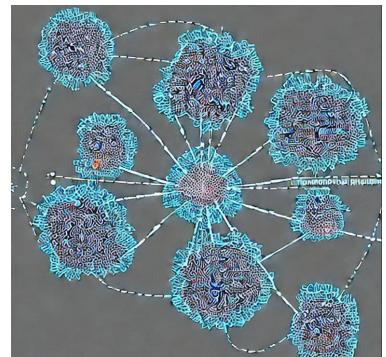
Then I tried "Cluster Detection":







Then I tried "Cluster Detection":

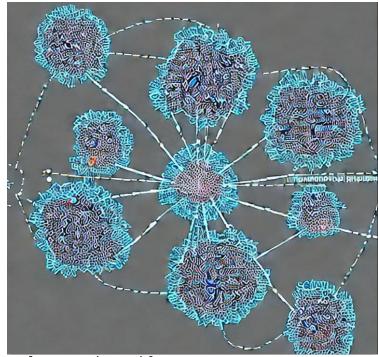








Then I tried "Cluster Detection":



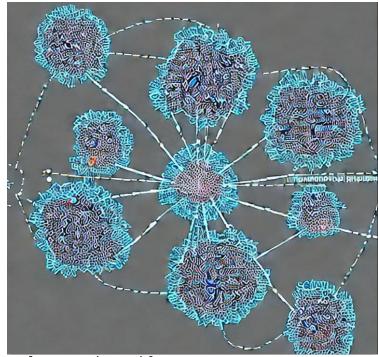
And Anomaly Detection:



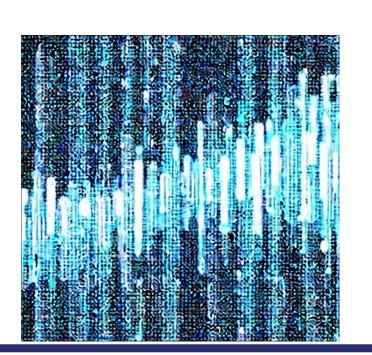




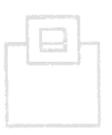
Then I tried "Cluster Detection":

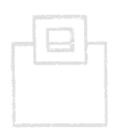


And Anomaly Detection:









Then I tried a different generator with "COBOL AI DB2":

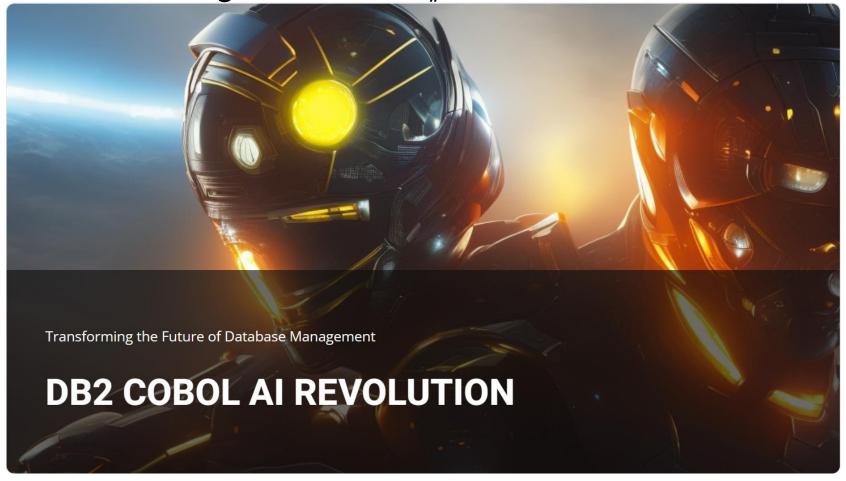








Then I tried a different generator with "COBOL AI DB2":

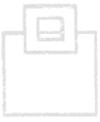








WOW! But it went off on one as well:







WOW! But it went off on one as well:









WHAT IS DB2 COBOL?

DB2 COBOL is a powerful, high-performance database management system that integrates with the COBOL programming language, providing a robust and reliable platform for data storage and retrieval.

DB2 COBOL combines a powerful database with the COBOL language for a solid data management solution.



So you can see, it might write you a love letter, an interpretation of "Der Gärtner" from Joseph Freiherr von Eichendorff and translate code snippets Between Python and C++, but it is *not* Intelligent!

The second secon

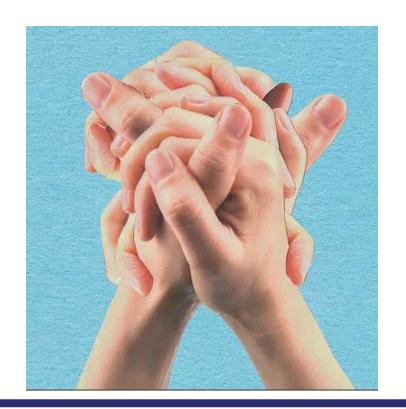
Fingers anyone?





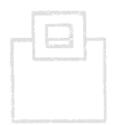
So you can see, it might write you a love letter, an interpretation of "Der Gärtner" from Joseph Freiherr von Eichendorff and translate code snippets Between Python and C++, but it is *not* Intelligent!

Fingers anyone?









The last page of the generated presentation looked like:









The last page of the generated presentation looked like:

EMBRACE THE DB2 COBOL AI REVOLUTION

By adopting the powerful combination of DB2 COBOL and artificial intelligence, businesses can harness the potential of this cutting-edge technology to revolutionize their operations and achieve greater success in the digital era.













Xing article from 2023-04-18 by Thomas Knüwer:

Künstliche Intelligenz – mal nüchtern betrachtet

Wie intelligent sind die aktuellen AI-Programme?

Das ist eine Frage der Definition von Intelligenz. Weshalb ich diese Vokabel im Zusammenhang mit Software auch für schwierig halte. Eigentlich müsste man sagen:

Die aktuellen KI-Projekte haben eine Intelligenz von null.

Denn: Sie verstehen ja nicht, was sie da gerade tun. Sie arrangieren Worte und Bilder um eine vorgegebene Aufgabe möglichst gut zu erfüllen.

Doch wie gesagt: Ich halte Intelligenz für die falsche Maßgröße im Rahmen der Debatte. Sinnvoller ist aus meiner Sicht die Frage:

Wie leistungsfähig sind die aktuellen AI-Programme?

Bei den meisten derzeit öffentlich zugänglichen Programmen kommt erst Begeisterung auf und dann recht schnell Ernüchterung. Meine jährliche Trendprognose habe ich zum Beispiel diesmal mit KI-Bildern illustriert und die Ergebnisse fand ich eher wenig euphorisierend.

https://www.xing.com/news/article/kuenstliche-intelligenz-mal-nuechte-1







0

Xing article from 2023-04-18 by Thomas Knüwer:

Artificial intelligence – a sober look

How intelligent are current AI programs?

This is a question of the definition of intelligence. Which is why I also find this vocabulary difficult in the context of software. Actually, one would have to say:

Current AI projects have an intelligence of zero.

After all, they don't understand what they are doing. They are arranging words and images to fulfill a given task as well as possible.

But as I said, I think intelligence is the wrong measure for this debate. More meaningful, in my view, is the question:

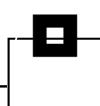
How powerful are the current AI programs?

With most of the programs currently available to the public, enthusiasm arises at first and then disillusionment quite quickly. This time, for example, I illustrated my annual trend forecast with AI images and found the results rather less euphoric.

https://www.xing.com/news/article/kuenstliche-intelligenz-mal-nuechte-1



Agenda

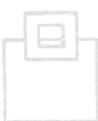


What is Artificial Intelligence?

Db2 13 and Data Insights - Al for free?



SEG and Al



Q&A



Db2 13 FL500 brings three new Scalar Built-in Functions (BiFs). These are the new SQL Data Insights (DI)

First, you need to make sure you have all the prereqs in place. These are basically a couple of APARs for the IBM Z Deep Neural Network Library (zDNN), the z/OS Supervisor, IBM OpenBLAS, z/OS OpenSSH and IBM 64-bit SDK for z/OS Java.



zDNN and OpenBLAS come with z/OS 2.4/2.5, but without the required APARs, the libraries may be empty.

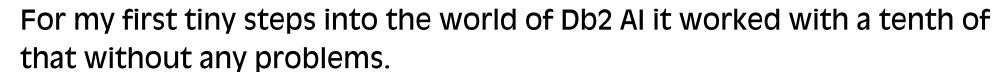


SQL Data Insights (SQL DI) is a kind of no-charge add-on to Db2 13, so you need to order and install it separately (FMID HDBDD18).



Then you need to install and customize SQL DI, starting with the definition of a (technical) user along with its appropriate authorization.

They ask for 100 GB of storage for the zFS home directory, but I think you'll only need that when you start to run AI model training on vast amounts of data.



The requirements listed for CPU and system memory aren't much smaller, and I'm experiencing a very measurable CPU consumption whenever the Model training on an object starts.







Next up is a bunch of RACF defs and the first occasion where the dreaded "certificates" raises its ugly head!

While waiting for the RACDCERT to be sorted out you can now execute the DDL in the DSNTIJAI member in the SDSNSAMP lib. This creates the pseudocatalog tables and other artifacts that SQL DI requires.



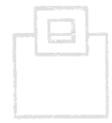
You also have to make sure that WLM has the correct Utility DD Cards and the NUMTCB=1 override for the LOAD. It is *very* annoying to wait one hour and then have the LOAD fail due to a DD Card being missing!





Now the real install process begins with the "standard" SMP/E install of SQL DI and the execution of the Install Script in the USS environment.

Be very careful with the values you use for all of these and stick with them to the bitter end!



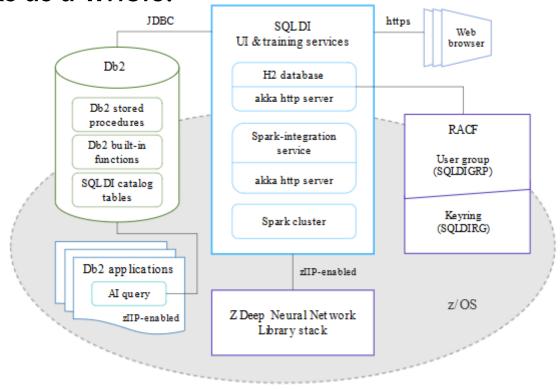
Finally, you can create an STC for SQL DI which is highly recommended. Conversely, the STC for Spark I would not bother with, as it is fully under the control of SQL DI anyway.







This is how DI looks as a whole:



https://www.ibm.com/docs/en/db2-for

-zos/13?topic=running-ai-queries-sql-data-insights

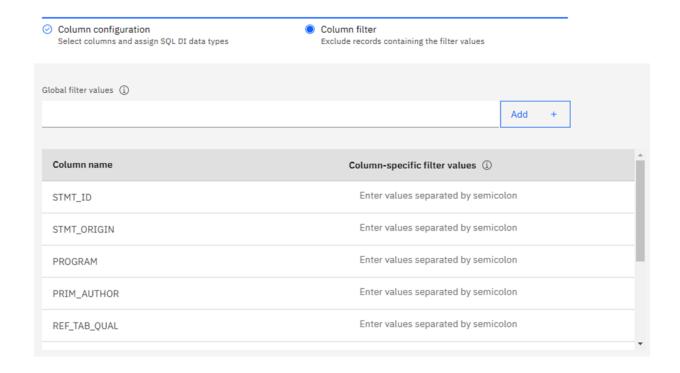




After having started the STC you can sign-on through the web interface:

Sign in to SQL Data Insights	FEF
Username	
Password	
Sign in	

The "tricky" part is choosing the "Usual Suspects". Which columns should be used as input to the Model? You only get one "key" column and so I chose STMT_ID and then 16 other columns from WLX:







Once you've done all that selecting you start the Model on its merry way...

											- 1	Inches
JOBNAME	StepName	ProcStep	JobID	Owner	С	Pos	DP	Real	Paging	SIO	CPU%	
SQLDSPKX	SQLDSPKX		STC09599	SQLDIID		ΙN	DE	12T	$\overline{0}$. $\overline{00}$	907.18	3.81	
SQLDSPKX	SQLDSPKX		STC09598	SQLDIID		ΙN	DE	13T	0.00	940.71	4.08	
SQLDSPKD	SQLDSPKD		STC09597	SQLDIID		IN	DE	39T	0.00	2177.2	64.75	
SQLDSPKX	SQLDSPKX		STC09596	SQLDIID		IN	DE	11T	0.00	892.00	4.08	
SQLDSPKX	SQLDSPKX		STC09595	SQLDIID		IN	DE	11T	0.00	940.73	4.20	
SQLDAPPS	SQLDAPPS	STEP1	STC08994	SQLDIID		LO	FF	407	0.00	0.00	0.00	
SQLDAPPS	SQLDAPPS		STC08972	SQLDIID		IN	DE	85T	0.00	2.56	5.37	
SQLDAPPS	SQLDAPPS		STC08877	SQLDIID		LO	FF	3401	0.00	0.00	0.00	
SQLDSPKM	SQLDSPKM		STC07682	SQLDIID		IN	DE	35T	0.00	0.85	0.26	
SQLDAPPS	SQLDAPPS		STC07665	SQLDIID		LO	FF	409	0.00	0.00	0.00	
SQLDSPKW	SQLDSPKW		STC07659	SQLDIID		IN	DE	32T	0.00	110.48	1.15	

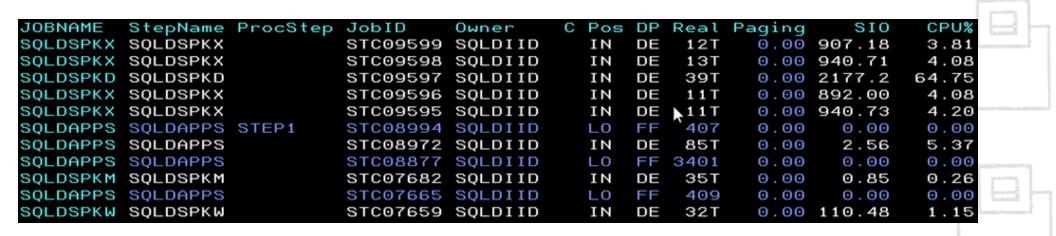
Yes, you can get multiple cups of coffee as you break all local records for SIO, CPU and zIIP CPU usage at your site!

After 40 minutes elapsed, the LOAD Utility got triggered and I was finished!





Once you've done all that selecting you start the Model on its merry way...



Yes, you can get multiple cups of coffee as you break all local records for SIO, CPU and zIIP CPU usage at your site!

After 40 minutes elapsed, the LOAD Utility got triggered and I was finished! This was with only 64.390 rows...



Db2 13 and Data Insights – AI for free?

You may now look into the SYSAIDB.* tables to review what it has saved but it is not really required. As long as the LOAD has completed you are good to go!



Now you can use the RUN QUERY button in the WEB UI or you can simply write SQL in SPUFI to see what the AI BiFs have given you.



One last performance point: 500,000 rows took five hours to model before being cancelled as it had taken 99% of local paging datasets and our machine was about to die. Looking in the logs it was only half way through the analysis...



Db2 13 and Data Insights – AI for free?



Naturally the question hanging in the air is:







Db2 13 and Data Insights – AI for free?



Naturally the question hanging in the air is:



Was that really all free?





Agenda

• What is Artificial Intelligence?

Db2 13 and Data Insights - Al for free?

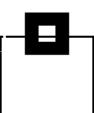


SEG and AI



Q&A





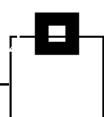
Using our WLX data I wrote and tested a bunch of AI SQL. First AI_ANALOGY:

```
SELECT AI ANALOGY ('BOXWEL3' USING MODEL COLUMN PRIM AUTHOR,
                  'IQATW001' USING MODEL COLUMN REF_TABLE ,
                  'KKKKKKK' USING MODEL COLUMN PRIM AUTHOR,
                  REF TABLE ) AS AI VALUE
      , A.WLX TIMESTAMP
      ,A.STMT ID
      , A.STMT TIMESTAMP
      , SUBSTR(A.PRIM_AUTHOR , 1 , 8 ) AS PRIM_AUTHOR
      ,SUBSTR(A.PROGRAM , 1 , 8 ) AS PROGRAM
      ,SUBSTR(A.REF_TABLE , 1 , 18) AS REF_TABLE
      , A. EXECUTIONS
      ,A.GETP OPERATIONS
      ,A.ELAPSE_TIME
      ,A.CPU_TIME
      , A.STMT_TEXT
FROM IQA061QB.IQATW001 A
WHERE A.PRIM_AUTHOR = 'KKKKKKK'
 AND AI_ANALOGY('BOXWEL3'
                             USING MODEL COLUMN PRIM AUTHOR,
                  'IQATW001'
                             USING MODEL COLUMN REF TABLE ,
                  'KKKKKKK'
                              USING MODEL COLUMN PRIM AUTHOR,
                 REF TABLE ) > 0.5
ORDER BY 1 DESC -- SHOW BEST FIRST
--ORDER BY 1
               -- SHOW WORST FIRST
FETCH FIRST 2000 ROWS ONLY;
```









Then AI_SEMANTIC_CLUSTER:

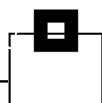
```
SELECT AI SEMANTIC CLUSTER (PROGRAM, 'SYSLN200', 'DSNATYPU')
                                   AS SIMILARITY
     ,A.STMT ORIGIN
     ,SUBSTR(A.STMT_TEXT , 1 , 60) AS STMT_TEXT
     ,SUBSTR(A.PROGRAM , 1 , 8) AS PROGRAM
     ,A.STMT ID
     , A. EXECUTIONS
     ,A.GETP_OPERATIONS
     ,A.ELAPSE_TIME
     ,A.CPU_TIME
FROM DAIN0610.IQATW001 A
WHERE 1 = 1
                  NOT IN ('SYSLN200', 'DSNATYPU')
 AND A.PROGRAM
 AND AI SEMANTIC CLUSTER (PROGRAM, 'SYSLN200', 'DSNATYPU')
                        > 0.3
ORDER BY SIMILARITY DESC
FETCH FIRST 20 ROWS ONLY;
```

This only worked with > 0.3 while it found none with > 0.5









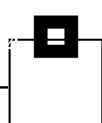
The semantic cluster did find just Dynamic SQL in the cluster of those two Dynamic SQL Packages:

SIMILARI	ry STMT_ORIGIN	+	PROGRAM
+0.4471843242645264E+0		SELECT DISTINCT PG.COLLID FROM SYSIBM.SYSPACKLIST PL, SYSIBM	DSNCLINF
+0.4471843242645264E+0	90 D	SELECT * FROM SYSIBM.SYSDBRM WHERE NAME = " AND PLNAM	DSNCLINF
+0.4471843242645264E+0	90 D	SELECT STMT, STMTNO FROM SYSIBM.SYSPACKSTMT WHERE LOCATION =	DSNCLINF
+0.4471843242645264E+0	90 D	SELECT COLLID, OWNER, CREATOR, VERSION, PDSNAME, QUALIFIER,	DSNCLINF
+0.4471843242645264E+0	90 D	SELECT STMT, STMTNO FROM SYSIBM.SYSPACKSTMT WHERE LOCATION =	DSNCLINF
+0.4471843242645264E+0	90 D	SELECT COLLID, OWNER, CREATOR, VERSION, PDSNAME, QUALIFIER,	DSNCLINF
+0.4471843242645264E+0	90 D	SELECT DISTINCT PG.COLLID FROM SYSIBM.SYSPACKLIST PL, SYSIBM	DSNCLINF
+0.4471843242645264E+0	90 D	SELECT * FROM SYSIBM.SYSDBRM WHERE NAME = AND PLNAM	DSNCLINF
+0.4471843242645264E+0	90 D	SELECT * FROM SYSIBM.SYSPLAN WHERE NAME = ' FOR FET	DSNCLINF
+0.4471843242645264E+0	00 D	SELECT * FROM SYSIBM.SYSPLAN WHERE NAME = ''''' ' FOR FET	DSNCLINF



However, if the program, in this case, was not in the training data you get back NULL...





Then with AI_SIMILARITY:

```
SELECT AI SIMILARITY ( PROGRAM,
                     'IQADBACP') AS AI_VALUE
      ,A.STMT ORIGIN
      ,SUBSTR(A.STMT_TEXT , 1 , 60) AS STMT_TEXT
      ,SUBSTR(A.PROGRAM , 1 , 8) AS PROGRAM
      ,A.WLX_TIMESTAMP
      ,A.STMT_ID
      ,A.STMT_TIMESTAMP
      , A. EXECUTIONS
      ,A.GETP_OPERATIONS
      ,A.ELAPSE_TIME
      ,A.CPU_TIME
FROM DAIN0610.IQATW001 A
WHERE 1 = 1
 AND NOT A.PROGRAM = 'IQADBACP'
 AND AI SIMILARITY ( PROGRAM,
                    'IQADBACP') IS NOT NULL
ORDER BY 1 DESC -- SHOW BEST FIRST
--ORDER BY 1 -- SHOW WORST FIRST
FETCH FIRST 10 ROWS ONLY;
```







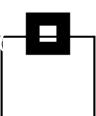
This is the result of looking for similar programs to our main dynamic SQL program:











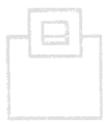
Finally using a not similar version of AI_SIMILARITY:

```
SELECT * FROM
    (SELECT AI_SIMILARITY(PROGRAM, 'SYSLN200') AS SIMILARITY
      ,C.STMT_ORIGIN
      ,SUBSTR(C.STMT_TEXT , 1 , 60) AS STMT_TEXT
      ,SUBSTR(C.PROGRAM , 1 , 8) AS PROGRAM
      ,C.STMT ID
      , C.END_USERID
      ,C.STMT_TIMESTAMP
      , C. EXECUTIONS
      ,C.GETP_OPERATIONS
      , C.ELAPSE_TIME
      , C.CPU_TIME
    FROM DAIN0610.IQATW001 C
    WHERE PROGRAM <> 'SYSLN200')
  WHERE SIMILARITY < 0.5
  ORDER BY SIMILARITY ASC
FETCH FIRST 200 ROWS ONLY
```

Notice that this one is doing dissimilar!







And it works! It truly finds in this case *just* Static SQL doing stuff nowhere like the Dynamic SQL Package we used in the previous example:







This is actually pretty neat!



As you have seen these queries actually *do* find "hidden patterns" as long as you have an idea about what is *in* the data.

All three Al BiFs come with nasty little notes like:

"The result can be null; if any argument is null, the result is the null value. If the arguments to the function contain values that <u>were not seen during model training</u>, and the model column is trained as categorical, the result is the null value."

This is, for me, a bit of a killer...







The next problem is that to find an abnormal result you must know the abnormal result first.

cannot have any abnormal data within otherwise the Model will not "see" that as being abnormal in a query. How can you possibly know this?



This then begets the question:

What are we doing this for?





```
<PROFILE>
                                         VALUE="IQA061QC"
   KALIAS-CREATOR
  <MIN-TOT-CPU</pre>
                                         VALUE="1000" />
                                         VALUE="25" />
  KTHRS-AVG-CPU
  KLESS-MORE
                                          VALUE="BOTH" />
  KWTO-MSG
                                         VALUE="WTO" />
  KALARM-RC
                                         VALUE="N" />
  (XPLNSUPPR-PRAUTHID
(PROCESS)
                                         VALUE="WEEK" />
   <compare-against</pre>
  KCOMPARE-HOURS
                                         VALUE="12" />
/WLX>
```







```
<PROFILE>
                                         VALUE="IQA061QC"
   KALIAS-CREATOR
  KMIN-TOT-CPU
                                         VALUE="1000" />
   KTHRS-AVG-CPU
                                         VALUE="25" />
  KLESS-MORE
                                         VALUE="BOTH" />
  KWTO-MSG
                                         VALUE="WTO" />
  KALARM-RC
                                         VALUE="N" />
  (XPLNSUPPR-PRAUTHID
(PROCESS)
                                         VALUE="WEEK" />
  <compare-against</pre>
  KCOMPARE-HOURS
                                         VALUE="12" />
/WLX>
```







```
<PROFILE>
                                         VALUE="IQA061QC"
   KALIAS-CREATOR
  KMIN-TOT-CPU
                                         VALUE="1000" />
   KTHRS-AVG-CPU
                                         VALUE="25" />
  KLESS-MORE
                                         VALUE="BOTH" />
  KWTO-MSG
                                         VALUE="WTO" />
  KALARM-RC
                                         VALUE="N" />
  (XPLNSUPPR-PRAUTHID
(PROCESS)
                                         VALUE="WEEK" />
  <compare-against</pre>
  KCOMPARE-HOURS
                                         VALUE="12" />
/WLX>
```







```
<PROFILE>
                                         VALUE="IQA061QC"
   KALIAS-CREATOR
   KMIN-TOT-CPU
                                         VALUE="1000" />
   KTHRS-AVG-CPU
                                         VALUE="25" />
   KLESS-MORE
                                          VALUE="BOTH" />
   <WTO-MSG</pre>
                                         VALUE="WTO" />
   KALARM-RC
                                         VALUE="N" />
  (XPLNSUPPR-PRAUTHID
(PROCESS)
                                         VALUE="WEEK" />
   <compare-against</pre>
  KCOMPARE-HOURS
                                         VALUE="12" />
/WLX>
```





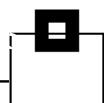


```
<PROFILE>
                                         VALUE="IQA061QC"
  KALIAS-CREATOR
  <MIN-TOT-CPU</pre>
                                         VALUE="1000" />
  KTHRS-AVG-CPU
                                         VALUE="25" />
  KLESS-MORE
                                         VALUE="BOTH" />
  KWTO-MSG
                                         VALUE="WTO" />
  KALARM-RC
                                         VALUE="N" />
  (XPLNSUPPR-PRAUTHID
(PROCESS)
                                         VALUE="WEEK" />
  <compare-against</pre>
  KCOMPARE-HOURS
                                         VALUE="12" />
WLX>
```









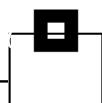
Output is just in batch at this time but it is written to a Db2 table which we will soon integrate into our Zowe.

```
SQL WorkloadExpert for Db2 z/OS Version 2.2 (22311)
                                                               Date: 2023-04-14
                                                               Time:
                                                                      09:57:34
ERRORLOG
                                                               Page:
WLX024E PRODUCT = WLX0610, PROGRAM = WLXKPROC, LOCATION = 0045
       Thresholds exceeded
       SQL length:
       SOL text : DELETE FROM IQAEQC01.PLAN TABLE WHERE 1=1
       Obj.creator: IOAE0C01
       Obj. name : PLAN TABLE
       Executions
                     Total
       Elapsed time Total
                             --- % average
       CPU time
                     Total
                             --- % average
                                             -39 %
       Getpage oper. Total
                             --- % average
       PrimAuthID : HOPPE
WLX024E PRODUCT = WLX0610, PROGRAM = WLXKPROC, LOCATION = 0045
       Thresholds exceeded
       SQL length:
                            117
       SQL text : SELECT SUBSYS , ID FROM IQA061QC.WLX_ADM_DSGROUP ORDER BY
       ID
       Obj.creator: IQA061QC
       Obj. name : IQATW042
       Executions
                     Total
       Elapsed time Total
                             --- % average
       CPU time
                                 % average
                                             102 %
                     Total
       Getpage open. Total
                             --- % average
       PrimAuthID : OPCC
```









Output is just in batch at this time but it is written to a Db2 table which we will soon integrate into our Zowe.

```
SQL WorkloadExpert for Db2 z/OS Version 2.2 (22311)
                                                              Date: 2023-04-14
                                                              Time:
                                                                      09:57:34
ERRORLOG
                                                              Page:
WLX024E PRODUCT = WLX0610, PROGRAM = WLXKPROC, LOCATION = 0045
       Thresholds exceeded
       SQL length:
       SOL text : DELETE FROM IQAEQC01.PLAN TABLE WHERE 1=1
       Obj.creator: IQAEQC01
       Obj. name : PLAN TABLE
       Executions
                     Total
       Elapsed time Total
                             --- % average
       CPU time
                     Total
                             --- % average
                                             -39 %
       Getpage oper. Total
                             --- % average
       PrimAuthID : HOPPE
WLX024E PRODUCT = WLX0610, PROGRAM = WLXKPROC, LOCATION = 0045
       Thresholds exceeded
       SQL length:
                            117
       SQL text : SELECT SUBSYS , ID FROM IQA061QC.WLX_ADM_DSGROUP ORDER BY
       ID
       Obj.creator: IQA061QC
       Obj. name : IQATW042
       Executions
                     Total
       Elapsed time Total
                             --- % average
       CPU time
                                 % average
                                             102 %
                     Total
       Getpage open. Total
                             --- % average
       PrimAuthID : OPCC
```







Sadly, it does not use any SQL DI BiFs, however, it works a treat and does not require half your machine to run!

We are still running and testing the DI BiFs but, as of this time, we see no real world chance of actually ever using them.



This can, and hopefully will, change of course!



Agenda

, — **—**

• What is Artificial Intelligence?

Db2 13 and Data Insights - Al for free?



SEG and Al



Q&A





Questions & Answers











Questions & Answers











PSP #11

Roy Boxwell SOFTWARE ENGINEERING GmbH

r.boxwell@seg.de





