

EVOLUCIÓN RECIENTE DEL BANCO DE DATOS ISBSG (INTERNATIONAL SOFTWARE BENCHMARKING STANDARDS GROUP) DE PROYECTOS DE SOFTWARE

José-de-Paula Maciel

Marta Fernández-Diego

María-Dolores Sanz-Berzosa

José-María Torralba-Martínez

Universidad Politécnica de Valencia

Abstract

The evolution produced in the ISBSG dataset based on the new release 11 is presented in this paper, referred to the following characteristics: overall data quality, functional size data quality, measurement methods, development type, effort and project elapsed time.

Relevant changes have been produced with the new release, so future research based on the new release must take this into account in order to properly compare the results obtained with earlier releases.

Keywords: *software projects; ISBSG dataset; dataset evolution*

Resumen

Se presenta la evolución producida en el banco de datos ISBSG, basada en la nueva versión 11, referida a las características: calidad general de la información; calidad de la medición del tamaño funcional; métodos de medida; tipos de desarrollo; esfuerzo; y plazo de los proyectos.

Se han producido cambios relevantes con ocasión de la nueva versión, por lo que en las investigaciones a realizar con la nueva versión, hay que tenerlas en cuenta a efectos de la comparación con resultados basados en otras versiones.

Palabras clave: *proyectos de software; banco de datos ISBSG; evolución de bancos de datos*

1. Introducción

Es una oportunidad la aparición de la nueva versión 11 del repositorio ISBSG, con la incorporación de nuevos proyectos de software de reciente realización, pero plantea la necesidad de conocer cómo evolucionan las características generales del banco de datos, así como de los proyectos, a la hora de comparar análisis realizados con distintas versiones del mismo repositorio (Premraj et al., 2005).

Es posible que la distribución de alguna de las características de los proyectos insertados en la última versión altere las conclusiones de un momento dado basadas en versiones

anteriores. Es decir, en la versión 11 puede haber una parte importante de proyectos que ha variado respecto de la versión 10, en relación con algunas características del proyecto, como pueden ser las relativas a variables como las siguientes: tamaño, tipo de proyecto, esfuerzo, duración, sector del cliente,...

1.1 Objeto

Se presenta la evolución producida en el banco de datos ISBSG, basada en la nueva versión 11, referida a las características: calidad general de la información; calidad de la medición del tamaño funcional; métodos de medida del tamaño; tipos de desarrollo; esfuerzo; y plazo de los proyectos.

1.2 Justificación

Un primer problema de los bancos de datos que crecen con el tiempo, como es el caso del repositorio ISBSG, es que los resultados obtenidos con una versión no son reproducibles, si se utiliza otra versión posterior.

Otro problema aparece si no se explicitan los criterios de selección de proyectos utilizados en la investigación para, entre otras cosas, resolver el problema de los datos perdidos (Pickard, Kitchenham & Linkman, 1999).

En definitiva la no reproducibilidad de los resultados supone una violación de los principios científicos básicos (Kitchenham y Mendes, 2009). Aunque esto se resolvería simplemente congelando las distintas versiones.

Según Cheikhi, Abran y Buglione (2006), "... repositorios de datos como el de ISBSG pueden ayudar a la comunidad de Ingeniería del Software a entender mejor algunas de las relaciones causa-efecto,...".

Por todo ello, se considera que el trabajo que se presenta, tiene interés para la investigación basada en el banco de datos y la necesaria comparación con investigaciones realizadas con versiones previas.

2. Breve descripción general del banco de datos ISBSG

En la Web de ISBSG (www.isbsg.org¹) se presenta como una organización sin fines de lucro que estableció y ahora va haciendo crecer, mantiene y explota repositorios de datos históricos de Tecnologías de la Información (TI).

Tiene como misión mejorar la gestión de los recursos de TI de las empresas y de los gobiernos, mediante la provisión y explotación de repositorios públicos de conocimiento de ingeniería de software que están estandarizados, verificados, y son recientes y representativos de las tecnologías actuales.

La formación de ISBSG en 1997 fue realizada tras varios años de previa cooperación por un grupo de asociaciones nacionales de métricas de software. Los miembros actuales de ISBSG representan a 12 Asociaciones de TI y de Métricas de 11 países (Australia, Alemania, China, España, EE.UU., Finlandia, Holanda, India, Italia, Japón y Suiza).

Actualmente este conocimiento está siendo usado principalmente por varios grupos de interés: clientes, proveedores y académicos. Para los profesionales del sector, este banco de datos es de interés para llevar a cabo una planificación y gestión más realista de dichos

¹ Última consulta realizada el 20 de marzo de 2010.

proyectos. Por otro lado, los investigadores hacen uso de este banco de datos como base empírica para la investigación sobre distintos aspectos relacionados con las TI (Maciel et al., 2009), contrastando hipótesis y modelos que mejoren esta industria en el futuro (Villanueva, 2005; Alba, 2008).

El banco de datos se organiza por versiones, estando actualmente—2010—con la versión 11, que incluye más de 5.000 proyectos de software, distribuidos en el intervalo temporal 1989-2009.

3. Metodología

3.1 ISBSG crece con el tiempo

De una versión a otra del banco de datos, el número de proyectos se incrementa. La distribución temporal de proyectos para las versiones 10 y 11 permite mostrar la filosofía de inserción de proyectos de una versión a otra. Como cabía esperar, en la versión más reciente (versión 11) únicamente se añaden proyectos en los últimos años. Es decir que entre 1989 y 2002, ambas versiones cuentan con el mismo número de proyectos por año, un total de 2853. A partir del 2003, ya difieren, ya que en la versión 11 aparecen proyectos que no están en la versión 10. Pero en definitiva la versión 11 comparte los 4106 proyectos de la versión anterior y tan sólo 946 proyectos nuevos han sido añadidos entre 2003 y 2009.

A continuación, la tabla 1 muestra la distribución temporal de los proyectos que han sido añadidos entre 2003 y 2009.

Tabla 1: Distribución temporal de los proyectos añadidos en la versión 11

Año	Frecuencia	%	% acumulado
s.d.	16	1,7	1,7
2003	2	,2	1,9
2004	107	11,3	13,2
2005	130	13,7	27,0
2006	536	56,7	83,6
2007	95	10,0	93,7
2008	49	5,2	98,8
2009	11	1,2	100,0
Total	946	100,0	

Nota: s.d. significa sin dato.

Este análisis y todo el tratamiento estadístico de los datos del apartado de resultados se ha realizado utilizando SPSS versión 17.

3.2 Selección de proyectos que se han incorporado a la versión 11

Con SPSS el proceso de selección de proyectos que se han incorporado a la versión 11 no se podía automatizar fácilmente. Aunque los proyectos comunes comparten el identificador

de proyecto, la versión 11 ofrece nuevas variables. Esto no es un problema en sí, la dificultad viene porque no se ha guardado coherencia en el nombre de las variables. Así, un mismo nombre de variable puede corresponder a características distintas según la versión. Por esta razón el proceso de selección de los nuevos proyectos pasaba por procesar en SPSS de alguna manera todas las variables.

Como solución, se ha utilizado otra herramienta, en concreto la programación Shell en Linux, para realizar automáticamente este proceso de selección. El código empleado aparece en la figura 1.

Figura 1: Código del proceso de selección de los nuevos proyectos incorporados a la versión

11

```
sed 's/^.....;/&0;/' R10.csv > x.R10.csv
sed 's/^.....;/&1;/' R11.csv > x.R11.csv
sort x.* > all.csv

cat all.csv | awk -F';' '{if ($1==id) print
id=$1 }' > repes.csv

cat all.csv | awk -F';' '{if (($1!=id)&&($2==1)) print
id=$1 }' > nuevos.csv
```

Partiendo del fichero de la versión 10 y del fichero de la versión 11, la filosofía consiste en marcar los proyectos añadiendo un segundo campo en ambos ficheros que identifica la versión. Así, en el fichero de la versión 10 se ha añadido un campo tras el identificador de proyecto con valor 0, y en el fichero de la versión 11, con el valor 1. Una vez hecho esto, se combinan ambos ficheros en uno, ordenados los proyectos por identificador y asegurándose que para los proyectos comunes aparecerá en primer lugar el proyecto de la versión 10 y en segundo lugar el proyecto de la versión 11.

A continuación a partir de este nuevo fichero generado, para identificar los proyectos repetidos simplemente se tiene que verificar si dos identificadores de proyecto en líneas consecutivas son iguales, y en ese caso se selecciona el proyecto repetido de la versión 11. Para los proyectos nuevos se verifica que el identificador no corresponda con el de la línea anterior, y sólo en el caso que se esté ante un proyecto de la versión 11, éste es incluido en el fichero de proyectos nuevos.

De esta manera se ha podido dividir los proyectos de la nueva versión 11 en dos grupos, los comunes a la versión anterior, y los nuevos de esta versión.

4. Resultados

4.1 Evolución de la calidad de los datos

Cada proyecto recibe dos tipos de valoraciones, realizadas por ISBSG, sobre la calidad de su información:

- La calidad en general de los datos, que está categorizada en cuatro niveles decrecientes de calidad (A, B, C y D);
- La calidad de la medición funcional del proyecto. La categorización de esta variable es la misma que lo planteado para la calidad en general. Sin embargo esta variable hace referencia únicamente a la calidad de la medición del tamaño funcional del proyecto.

En primer lugar de los 2853 proyectos comunes del periodo comprendido entre 1989 y 2002, ambas versiones comparten el mismo número de proyectos clasificados como calidad C (134 proyectos) y D (105 proyectos). En la versión 10, 717 proyectos eran clasificados según la máxima calidad (A) y 1.897 como calidad B. Sin embargo, 12 de los proyectos considerados de calidad A en la versión 10 aparecen como de calidad B en la versión 11, es decir que se ha producido una pequeña recalificación.

Aunque esto sólo ocurre para la variable calidad de los datos en general, porque para la variable calidad de los datos de tamaño funcional, la distribución de frecuencias es idéntica en ambas versiones para este periodo.

A partir de este resultado sería interesante verificar si la única variable que cambia, y tampoco en gran medida, es la de calidad general de los datos. También se podría verificar en futuros estudios la hipótesis de que se habrá recalificado proyectos relativamente recientes.

La tabla 2 presenta ya los resultados obtenidos de la comparación entre los proyectos comunes a ambas versiones y los proyectos nuevos incorporados a la versión 11 en lo que a calidad en general de los datos se refiere.

Tabla 2: Calidad general de los datos y del tamaño funcional

	Calidad general de los datos				Calidad del tamaño funcional			
	Proyectos comunes		Proyectos nuevos		Proyectos comunes		Proyectos nuevos	
	Frecuencia	%	Frecuencia	%	Frecuencia	%	Frecuencia	%
<i>s.d.</i>					272	6,6	268	28,3
A	835	20,3	93	9,8	2265	55,2	63	6,7
B	2976	72,5	840	88,8	852	20,8	189	20,0
C	154	3,8	9	1,0	702	17,1	426	45,0
D	141	3,4	4	,4	15	,4		
Total	4106	100,0	946	100,0	4106	100,0	946	100,0

Nota: *s.d.* significa sin dato.

En cuanto a la calidad de los datos en general, el porcentaje de proyectos de calidad A ha disminuido de un 20% a un 10% aproximadamente. Pensamos que esto no significa una

pérdida de calidad real, sino que ha cambiado el procedimiento de valoración de la calidad de algunos proyectos, que habrá que tener en cuenta.

En cuanto a la calidad de los datos de tamaño funcional, dos observaciones resultan interesantes a propósito de la tabla:

- En primer lugar el porcentaje de proyectos de calidad A en relación con el tamaño funcional ha disminuido de un 55% a un 7% aproximadamente. Por lo que procede hacer el mismo comentario del párrafo anterior. En los análisis en que el tamaño funcional de los proyectos es clave, habría que preguntarse hasta qué punto es interesante contar con este subconjunto de proyectos nuevos. Por un lado resolvemos el problema de la obsolescencia, pero por otro lado, la fiabilidad nominal–aunque no la real–de las mediciones parece que no es comparable;
- En segundo lugar, tenemos un porcentaje muy elevado de proyectos sin valoración de la calidad de los datos del tamaño funcional (28%) con respecto a los proyectos de la versión 10 (7%). En el siguiente punto encontraremos la explicación.

4.2 Evolución de los métodos de medida del tamaño funcional

La tabla 3 presenta la evolución de los métodos de medida del tamaño funcional en el repositorio ISBSG en la versión 11.

Tabla 3: Evolución de los métodos de medida del tamaño funcional

	Proyectos comunes		Proyectos nuevos	
	Frecuencia	%	Frecuencia	%
Albrecht	2	,0		
Automated	4	,1		
Backfired	8	,2		
COSMIC	117	2,8	228	24,1
Dreger	10	,2		
Feature Points	2	,0		
FiSMA	340	8,3	156	16,5
Fuzzy Logic	2	,0		
IFPUG	3281	79,9	518	54,8
In-house	1	,0		
LOC	146	3,6	40	4,2
Mark II	35	,9	1	,1
NESMA	152	3,7	3	,3
Retrofitted	2	,0		
System Components	1	,0		
Unknown	3	,1		
Total	4106	100,0	946	100,0

Las principales características a destacar del análisis son las siguientes:

- La proporción de la métrica COSMIC en nuevos proyectos es de un 24% siendo del 3% en proyectos comunes. Para estos proyectos COSMIC, la variable calidad de los datos de tamaño funcional está sin dato. Este hecho explica la mayor proporción de blancos observada en el punto anterior. Además, la proporción de proyectos FiSMA ha aumentado también de un 8% a un 16%;
- Sin embargo, la proporción de la métrica IFPUG ha disminuido en una proporción muy considerable, de un 80% en proyectos comunes a un 55% en proyectos nuevos.

Siendo COSMIC el método más empleado en Europa, podemos pensar en que la escasez de proyectos europeos pudiera estar subsanándose en este banco de datos, con sus ventajas e inconvenientes.

4.3 Evolución de algunas características de los proyectos

1. El tipo de desarrollo

Para el tipo de desarrollo, las diferencias en porcentajes entre los distintos proyectos, comunes o nuevos, no parecen significativas. Sin embargo, en este caso parece más interesante la evolución temporal, hasta el año 2002 y a partir del año 2003, que muestra la tabla 4.

Tabla 4: Evolución temporal del tipo de desarrollo de los proyectos

	1989-2002		2003-2009	
	Frecuencia	%	Frecuencia	%
Mantenimiento	1459	51,1	1501	68,8
Nuevo desarrollo	1318	46,2	645	29,5
Re-desarrollo	59	2,1	30	1,4
Otros	17	0,6	7	0,3
Total	2853	100,0	2183	100,0

De esta tabla se observa un dato interesante. Así como hasta el 2002, prácticamente la proporción de proyectos de nuevos desarrollos y de mejora es la misma, desde el 2003 estamos ante una situación en que se suele mejorar el software ya existente y los proyectos de nuevos desarrollos son más escasos.

2. El esfuerzo

A la hora de estudiar la influencia que tiene el factor proyectos nuevos en la variable esfuerzo, lo que tratamos de estimar es el carácter significativo del efecto de los proyectos añadidos en la versión 11 con respecto a la versión 10, gracias al análisis de la varianza (ANOVA).

Teniendo en cuenta que este esfuerzo lo tenemos desglosado en las diferentes fases del ciclo de desarrollo de un proyecto (planificación, especificación, diseño, construcción, pruebas, implantación), la tabla 5 muestra los resultados para el esfuerzo global y para cada una de estas variables.

Tabla 5: ANOVA del esfuerzo

	Esfuerzo	Esfuerzo desglosado					
		Planificación	Especificación	Diseño	Construcción	Pruebas	Implantación
F	5,920	,000	11,833	,981	2,002	1,353	,070
Sig.	,015	,989	,001	,322	,157	,245	,792

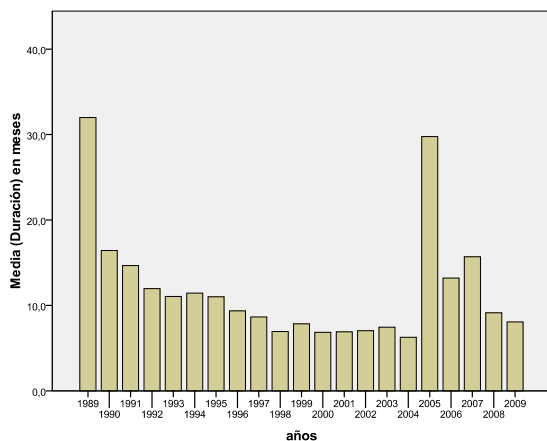
Según la tabla, aunque podemos considerar que los nuevos proyectos son significativamente distintos con respecto a los otros en cuanto al esfuerzo global, aparece especialmente significativa la diferencia en la fase de especificación (Maxwell, 2002; Romero & Zúñica, 2010).

3. La duración del proyecto

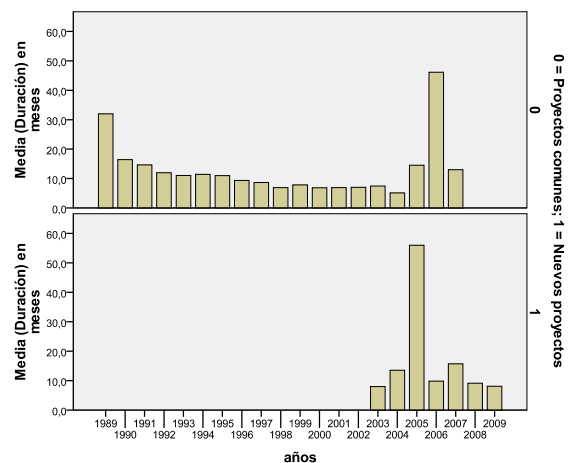
En cuanto a la duración del proyecto, se obtiene una $F = 21,251$ y un alto grado de significación $Sig. = 0,000$.

La figura 2 muestra la media de la duración de los proyectos en meses para los nuevos proyectos incluidos en la versión 11 y los proyectos comunes a ambas versiones.

Figura 2: Evolución de la duración de los proyectos



a) Para todos los proyectos



b) Para los proyectos comunes y los proyectos nuevos por separado

La gráfica nos muestra claramente cómo para el año 2005 los nuevos proyectos suben la media global de manera considerable.

5. Discusión de resultados

La calidad en general de los datos de los proyectos ha variado, en la versión 11 respecto de la anterior, de forma relevante. Y la calidad de la medida funcional ha variado debido, entre otras circunstancias, a la variación del tipo de métrica utilizada en los nuevos proyectos y el aumento de proyectos sin dato de valoración de dicha calidad.

El método de medida COSMIC ha aumentado, y ha disminuido IFPUG, lo que habrá que tener en cuenta cuando se tenga que utilizar el dato del tamaño.

Las variaciones producidas en la versión 11 exigen que, cuando se utilice para la investigación, y se quiera comparar con resultados basados en otras versiones, se tengan en cuenta las cuestiones referidas. En la literatura se insiste mucho en aspectos como éstos (Premraj et al., 2005; Kitchenham, 1998).

6. Conclusiones

Se han producido cambios relevantes con ocasión de la nueva versión 11 del banco de datos de proyectos de software de ISBSG, por lo que en las investigaciones a realizar con la nueva versión, hay que tenerlas en cuenta a efectos de la comparación con resultados basados en otras versiones.

Los autores agradecen a los revisores las oportunas recomendaciones que realizaron sobre la propuesta de comunicación.

7. Referencias

Alba, C. (2008). *Predicción y clasificación del nivel de riesgo en proyectos de sistemas de información*. Unpublished doctoral dissertation, Universidad de Oviedo, España.

Cheikhi, L., Abran, A., & Buglione, L. (2006). ISBSG Software Project Repository & ISO 9126: An Opportunity for Quality Benchmarking. *European Journal for the Informatics Professional*, (Vol. 7, No. 1, pp. 46-52).

International Software Benchmarking Standards Group (ISBSG). www.isbsg.org (27-03-2010).

Kitchenham, B., & Mendes, E. (2009). Why Comparative Effort Prediction Studies may be Invalid. *Proceedings of the 5th International Conference on Predictor Models in Software Engineering*, Vancouver, Canada.

Kitchenham, B. (1998). A Procedure for Analyzing Unbalanced Datasets. *IEEE Transaction on Software Engineering*, (Vol. 24, No. 4).

Maciel, J. P., Fernández-Diego, M., Vázquez-Barrachina, E., Neves-Silva, A., & Torralba-Martínez, J. M. (2009). Fuente de datos para investigación sobre Gestión de Proyectos Informáticos: Repositorio ISBSG - International Software Benchmarking Standards Group. / *Congreso Iberoamericano SOCOTE*, Universidad Politécnica de Valencia, España.

Maxwell, K. (2002). *Applied statistics for software managers*. Prentice-Hall.

Pickard, L., Kitchenham, B., & Linkman, S. (1999). An Investigation of Analysis Techniques for Software Datasets. *Sixth International Software Metrics Symposium (METRICS'99)*.

Romero, R., & Zúñica, L. (2010). *Métodos Estadísticos en Ingeniería*. En Editorial de la Universidad Politécnica de Valencia, España.

Villanueva, B. J. (2005). *Estimación de costes y plazos en proyectos de sistemas de información*. Unpublished doctoral dissertation, Universidad de Oviedo, España.

Correspondencia (Para más información contacte con):

Marta Fernández-Diego
Universidad Politécnica de Valencia
Departamento de Organización de Empresas
Camino de Vera s/n, 46022 Valencia, Spain
Phone: +34 96 387 76 85
E-mail : marferdi@omp.upv.es