

1.- La experimentación en Ingeniería del Software

Los resultados que se obtendrán aplicando las tecnologías¹ para construcción de software son impredecibles [Wholin 00]. No existe evidencia alguna que apoye la mayoría de las creencias sobre las que basamos la construcción de software [Juristo 02]. Actualmente, las tecnologías que se emplean en el desarrollo de software carecen de evidencias sobre su adecuación, límites, cualidades, costes y riesgos [Jedlitschka 04]. El único modo para contrastar las creencias y las opiniones para convertirlas en hechos es la **experimentación**.

La Ingeniería del Software Experimental (ESE, en sus siglas inglesas) traslada a la Ingeniería del software (IS) el paradigma experimental que ha sido aplicado con éxito no ya en disciplinas científicas naturales clásicas (Física, Química, Medicina, etc.) sino que más recientemente ha permitido avanzar en disciplinas tradicionalmente no experimentales, tales como la Economía Experimental [Davis 92] o la Psicología Experimental [Myers 05].

El fin de la experimentación es identificar las causas por las que se producen determinados resultados. Un experimento modela en el laboratorio (es decir, en condiciones controladas) las principales características de una realidad (en nuestro caso el desarrollo de software) lo que permite estudiarla y comprenderla mejor. La fortaleza de la experimentación en laboratorio es que permite variar iterativamente aspectos de la realidad para estudiar el impacto que tienen tales manipulaciones.

La experimentación en IS hará posible la comprensión e identificación de las variables que entran en juego en la construcción de software y las conexiones que existen entre ellas. Experimentar con la construcción de software nos permitirá, en palabras de [Pfleeger 99] “Aumentar la comprensión de lo que hace al software bueno y cómo hacer software bien”². La meta de la ESE es hacer del desarrollo de software una actividad predecible científicamente gracias al conocimiento de las relaciones entre los procesos de producción de software y los productos que se obtienen. La National Science Foundation estadounidense ha establecido la experimentación como uno de los objetivos prioritarios en el que centrar parte de la investigación en IS: “Avanzar nuestra *comprensión* del proceso de IS mediante la *experimentación*”³ [NSF 99]

En la historia de todas las disciplinas ingenieriles se ha producido un paso similar al que la ESE propugna para la IS. La información usada para construir artefactos transita por etapas que van desde las creencias, especulaciones y aciertos casuales hasta el conocimiento científico por medio del cual la disciplina ingenieril alcanza resultados predecibles [Shaw 90]. La ESE establece los cimientos para **realizar investigaciones experimentales en IS** con el fin de llegar a conclusiones científicamente válidas sobre las que basar el desarrollo de software. Este conocimiento debe permitir identificar las condiciones de aplicabilidad, las debilidades y las fortalezas de las distintas tecnologías de desarrollo de software.

Un experimento es una investigación formal, rigurosa y controlada donde se manipulan variables bajo estudio. Las variables bajo estudio toman distintos valores y la investigación pretende averiguar los efectos que producen tales variaciones. Los efectos analizados son los cambios que, los distintos valores de las variables estudiadas, provocan (o no) en las variables respuesta. Ejemplos de factores estudiados en experimentos de IS son el método de inspección usado por los desarrolladores, la técnica de pruebas usada, la experiencia de los desarrolladores, etc. Ejemplos de variables respuesta en IS sobre las que se desea conocer el efecto de los factores son la eficacia, la eficiencia o la productividad. Dependiendo del experimento tales

¹ Usamos el término tecnología para referirnos a métodos, técnicas, metodologías, y herramientas que se utilizan en el desarrollo de software

² “Gain more understanding of what makes software good and how to make software well”

³ “Advance our *understanding* of the SE process by *experimenting*”.

variables respuesta pueden instanciarse como, por ejemplo, número de defectos encontrados, número de líneas de código programadas, etc.

2.- Líneas actuales de investigación en ESE

Para avanzar en la investigación experimental en IS no es suficiente con aplicar el diseño experimental y la estadística para el análisis de datos. Para aplicar el paradigma experimental a una disciplina es necesario construir su propia metodología experimental basada en los principios generales del experimentalismo. La metodología experimental específica para una disciplina no puede importarse directamente de otras.

Por ejemplo, el laboratorio en química está formado por matraces y pipetas y las variables que se controlan son temperaturas y presiones. En economía experimental, sin embargo, el laboratorio (en el sentido de realidad simplificada y controlable donde se realizan experimentos) se convierte en el trabajo con estudiantes en lugar de con mercados. Ese mismo concepto es aplicable a ESE, donde el laboratorio también consiste en el trabajo con estudiantes en lugar de con proyectos reales. Sin embargo, un aspecto central del control de los experimentos en economía es la utilización de una estructura de incentivos para los participantes con el fin de inducir en los sujetos características similares a las recompensas que existen en la realidad de los mercados [Ball 96] Obviamente, este principio director no es válido para ESE, donde las manipulaciones a realizar en los estudiantes para imitar la realidad de la construcción de software deben ser otras [Juristo 01].

Es decir, todas las disciplinas experimentales necesitan adaptar los principios del experimentalismo a su propio contexto. Es precisamente ese desarrollo de una metodología experimental específica para la IS en lo que consiste la investigación en ESE.

En sus inicios, la ESE se centró en estudiar la aplicación a IS de los principios del laboratorio y el experimento. Hoy en día la realización de experimentos en ESE se ha convertido en una tarea bastante comprendida. Hemos realizado progresos en la realización de experimentos aislados, pero éste es sólo un primer paso en la secuencia de actividades del paradigma experimental. Queda pendiente el análisis y adaptación de otros principios del experimentalismo.

Actualmente, la investigación en ESE se centra, no ya en la realización de experimentos de laboratorio a pequeña escala (aislados) sino pasar a la experimentación a gran escala (concatenando experimentos) [Pfleeger 05]. Tal migración dejará atrás los resultados poco generalizables para producir un cuerpo de conocimientos contrastados experimentalmente sobre construcción de software. Para ello, se hace necesario mejorar aspectos como:

- El reporte de experimentos. Para lo cual la comunidad está haciendo esfuerzos por generar guidelines de amplio uso para las publicaciones experimentales. Trabajan en esta línea, entre otros, Jedlitschka [Jedlitschka 05], Kitchenham [Kitchenham 02] y Singer [Singer 99]
- La comprensión del factor humano en los experimentos en IS. Para ello, la investigación se centra en la incorporación de métodos cualitativos provenientes de las ciencias sociales a los experimentos clásicos cuantitativos. La promotora principal de esta línea fue Seaman [Seaman 99].
- La realización de experimentos, no ya en el laboratorio, sino a nivel clínico (por utilizar un paralelismo con la medicina experimental). Es decir, involucrar más a la industria en la realización de sus propios estudios experimentales que comprueben los hallazgos en el laboratorio. Esta área ha sido liderada, principalmente por Rombach [Laitenberger 03] y la red ESERNET que él lideró.
- La generación de evidencias a partir de la suma de varios experimentos. Para ello la investigación se centra en mejorar la replicación de experimentos, la realización de revisiones sistemáticas de experimentos y estudiar cómo agregar los resultados para combinar los hallazgos de diversos experimentos.
 - La investigación en mejora de las replications se subdivide en dos aproximaciones complementarias. La línea que tiene por objetivo comprender mejor el concepto de replicación para adaptarlo a la realidad de ESE y en la que trabaja el grupo de la Universidad de Stratchlyde [Brooks 08] y James Miller, ahora en Alberta [Miller 05].

La otra línea pretende mejorar la transmisión de información mediante los paquetes de replicación. El pionero en estos trabajos fue Victor Basili [Basili 96] al que se han unido posteriormente Conradi [Conradi 01] y Travassos [Shull 02]. También nuestro grupo trabaja en esta línea [Vegas 06]

- Sobre revisiones sistemáticas, la pionera en establecer un proceso formalizado fue Barbara Kitchenham [Kitchenham 04]. Tras lo cual otros investigadores lo han llevado a la práctica, como Dyba [Dyba 06] y nuestro propio grupo [Davis 06].
- Finalmente el concepto de agregación fue inicialmente propuesto en ESE por Basili [Basili 99] y Miller [Miller 99]. Uno de los primeros intentos prácticos de combinar resultados de diferentes experimentos fue nuestro trabajo [Juristo 02].

3. Bibliografía

Ball & Cech (1996) *Subject pool choice and treatment effects in economic laboratory research* En Isaac (ed.) **Research in Experimental Economics** vol. 6. JAI Press

Basili, Green, Laitenberger, Lanubile, Shull, Sorumgaard, Zelkowitz (1996) *Packaging researcher experience to assist replication of experiments* **ISERN Meeting**

Basili, Shull, Lanubile (1999) *Building knowledge through families of experiments* **IEEE Transactions on SE** 25 (4)

Brooks, Roper, Wood, Daly, Miller (2008) *Replication's Role in SE* En Shull, Singer, Sjoberg (Eds.) **Guide to Advanced Empirical SE** Springer-Science

Conradi, Basili, Carver, Shull, Travassos (2001) *A pragmatic documents standard for an experience library: Roles, documents, contents and structure* **University of Maryland CS-TR-4235**

Cooper, Hedges (1994) **The Handbook of Research Synthesis** Russell Sage Foundation

Davis, Dieste, Hickey, Juristo, Moreno (2006) *Effectiveness of requirements elicitation techniques: Empirical results derived from a systematic review* **Proc. of the IEEE Int. Conf. on Requirements Eng.**

Davis, Holt (1992) **Experimental Economics** Princeton University Press

Dyba, Kampenes, Sjoberg (2006) *A systematic review of statistical power in SE experiments.* **Information and Software Technology** 48(8)

Jedlitschka, Ciolkowski (2004b) *Towards evidence in SE* **Proc. of ACM/IEEE Int. Symp. on Empirical SE**

Jedlitschka, Pfahl, (2005) *Reporting experiments in SE* **Proc. of the ACM/IEEE Int. Symp. on Empirical SE**

Judd, Smith, Kidder (1991) **Research Methods in Social Relations** Jovanovich College Publishers

Juristo, Moreno (2001) **Basics of Software Engineering Experimentation** Kluwer

Juristo, Moreno, Vegas (2002) *A survey on testing technique empirical studies: How limited is our knowledge* **Proc. of the ACM/IEEE Int. Symp. on Empirical SE**

Juristo, Moreno (2002) *Reliable knowledge for software development* **IEEE Software** 19(5)

Jørgensen (2004) *A review of studies on expert estimation of software development effort* **J. of Systems and Software** 70(1-2)

Kamsties, Lott (2005) *An empirical evaluation of three defect-detection techniques* **Proc. of the IEEE European SE Conf.**

Kitchenham, Pflieger, Pickard, Jones, Hoaglin, El Emam, Rosenberg (2002) *Preliminary guidelines for empirical research in SE.* **IEEE Transactions on SE** 28(8)

Kitchenham (2004) *Procedures for performing systematic reviews.* **Keele University TR/SE-0401**

Laitenberger, Rombach (2003) *Quasi-experimental studies in industrial settings* En Juristo and Moreno (eds.) **Lecture Notes on Empirical SE** World Scientific

Lázaro, Marcos (2006) *Experiences in integrating qualitative and quantitative methods* **Proc. of the ACM/IEEE Int. Symp. on Empirical SE**

Miller (1999) *Can results from SE experiments be safely combined?* **Int. Software Metrics Symp.**

- Miller (2000) *Applying meta-analytical procedures to SE experiments* **J. of Systems and Software** 54(1)
- Miller (2005) *Replicating SE experiments: A poisoned chalice or the holy grail* **Information and Software Technology** 47
- Myers, Hansen (2005) **Experimental Psychology** Wadsworth Publishing
- Nestle (2007) *Eating made simple: How do you cope with a mountain of conflicting diet advice?* **Scientific American**
- NSF (1999) *Final report National Science Foundation Workshop on a software research program for the 21st century* **Software Engineering Notes** 24(3)
- Pfleeger (1999) *Albert Einstein and empirical SE* **IEEE Computer** 32(10)
- Pfleeger (2005) *Soup or Art? The role of evidential force in empirical SE* **IEEE Software** 22(1)
- Pickard, Kitchenham, Jones (1998) *Combining empirical results in SE* **Information and Software Technology** 40(14)
- Runeson, Thelin (2004) *Prospects and limitations for cross-study analyses: A study on an experiment series* **Proc. of the Int. Workshop on Empirical SE**
- Seaman (1999) *Qualitative methods in empirical studies of SE* **IEEE Transactions on SE** 25
- Singer (1999) *Using the American Psychological Association (APA) style guidelines to report experimental results* **Proc. of the IEEE Workshop on Empirical Studies in Software Maintenance**
- Shaw (1990) *Prospects for an engineering discipline of software* **IEEE Software** 7(6)
- Shull, Basili, Carver, Maldonado, Travassos, Mendonca, Fabbri (2002) *Replicating SE experiments: Addressing the tacit knowledge problem* **Proc. of the ACM/IEEE Int. Symp. on Empirical SE**
- Shull, Carver, Travassos, Maldonado, Conradi, Basili (2003) *Replicated studies: Building a body of knowledge about software reading techniques* En Juristo y Moreno (Eds.) **Lecture Notes on Empirical SE** World Scientific
- Sjoeberg, Hannay, Hansen, Kampenes, Karahasanovic, Liborg, Rekdal (2005) *A survey of controlled experiments in SE* **IEEE Transactions on SE** 31
- Straus, Richardson, Glasziou, Haynes (2005) **Evidence-based Medicine. How to practice and teach EBM** Elsevier
- Vegas, Juristo, Moreno, Solari, Letelier (2006) *Analysis of the influence of communication between researchers on experiment replication* **Proc. of the ACM/IEEE Int. Symp. on Empirical SE**
- Wohlin, Runeson, Höst, Ohlsson, Regnell, Wesslén (2000) **Experimentation in Software Engineering: An Introduction** Kluwer
- Yin, Heald, (1975) *Using the case survey method to analyze policy studies* **Administrative Science Quarterly** 20(3)