

Study ID	Data Characteristics	Notes
S2	Industry	<ul style="list-style-type: none"> • Single company • 5 Web based applications developed using an OO framework.
S3	Academia	<ul style="list-style-type: none"> • Undergraduate students. • 45 Web applications, 6 of which were discarded due to incomplete information, leaving 39 observations.
S4	Academia	<ul style="list-style-type: none"> • Undergraduate students • 76 Web applications, 6 of which were discarded because of incomplete information leaving 70 observations.
S5	Academia	<ul style="list-style-type: none"> • Undergraduate students. • 45 Web applications, 6 of which were discarded due to incomplete information leaving 39 observations.
S6	Academia	<ul style="list-style-type: none"> • Final year undergraduate, or MSc students. • 43 Mediums sized Web applications of which 6 were removed due either to incomplete information or unjustifiably high effort leaving 37 observations.
S7	Academia	<ul style="list-style-type: none"> • Undergraduate students • 76 Web applications, 6 of which were discarded because of incomplete information leaving 70 observations.
S8	Academia	<ul style="list-style-type: none"> • Honours/MSc students. • 43 Hypermedia applications, 6 later removed due to insufficient information
S9	Academia	<ul style="list-style-type: none"> • Final year undergraduate, or MSc students. • 43 Mediums sized Web applications of which 6 were removed due either to incomplete information or unjustifiably high effort leaving 37 observations.
S10	Academia	<ul style="list-style-type: none"> • Final year undergraduate, or MSc students. • 43 Mediums sized Web applications of which 6 were removed due either to incomplete information or unjustifiably high effort leaving 37 observations.
S11	Academia	<ul style="list-style-type: none"> • Final year undergraduate, or MSc students. • 43 Mediums sized Web applications of which 6 were removed due either to incomplete information or unjustifiably high effort leaving 37 observations.
S14	Academia	<ul style="list-style-type: none"> • University of Auckland honors/post-grad Web hypermedia projects. • Web hypermedia applications • 2 case studies (34 applications and 25 applications respectively).

S15	Academia	<ul style="list-style-type: none"> • Final year undergraduate, or MSc students. • 43 Mediums sized Web applications of which 6 were removed due either to incomplete information or unjustifiably high effort leaving 37 observations.
S16	Academia	<ul style="list-style-type: none"> • Honours/Postgraduate students • 25 hypermedia applications
S17	Academia	<ul style="list-style-type: none"> • Postgraduate students • Web Hypermedia Applications • 37 applications • 3 outliers (observations where total effort was unrealistic compared to duration) were removed leaving 34 observations.
S19	Academia	<ul style="list-style-type: none"> • 4th and 5th year students. • 56 Web applications. • W2000 used as the design notation for all applications.
S21	Academia	<ul style="list-style-type: none"> • Postgraduate students • Web Hypermedia Applications • 37 applications • 3 outliers (observations where total effort was unrealistic compared to duration) were removed leaving 34 observations.
S22	Industry	<ul style="list-style-type: none"> • Cross-company data on Web software and hypermedia projects from the Tukutuku database. • 2 datasets evaluated from this database: <ul style="list-style-type: none"> ○ 12 projects from a single-company. ○ 37 projects from several companies.
S24	Industry	<ul style="list-style-type: none"> • Cross-company data on Web software and hypermedia projects from the Tukutuku database. • 2 datasets evaluated from this database: <ul style="list-style-type: none"> ○ 12 projects from a single-company. ○ 24 projects from several companies.
S25	Academia	<ul style="list-style-type: none"> • Postgraduate students • Web Hypermedia Applications • 37 applications • 3 outliers (observations where total effort was unrealistic compared to duration) were removed leaving 34 observations.
S26	Academia	<ul style="list-style-type: none"> • Graduate/undergraduate computer science students. • 22 Web projects. • “Real projects and clients” suggesting an industry link.
S27	Industry	<ul style="list-style-type: none"> • Single-company data. • 12 Web applications.
S28	Industry	<ul style="list-style-type: none"> • Single-company data.

		<ul style="list-style-type: none"> • 12 Web applications.
S30	Industry	<ul style="list-style-type: none"> • 20 single company Web applications.
S31	Academia	<ul style="list-style-type: none"> • 32 Web projects developed by Web engineering students.
S32	Industry	<ul style="list-style-type: none"> • Cross-company data on Web software and hypermedia projects from the Tukutuku database. • 53 projects were evaluated collectively and separately as 2 datasets: <ul style="list-style-type: none"> ○ 13 single-company projects ○ 40 cross-company projects
S34	Industry	<ul style="list-style-type: none"> • Cross-company data on Web software and hypermedia projects from the Tukutuku database. • 2 data sets evaluated: <ul style="list-style-type: none"> ○ The entire Tukutuku database of 67 projects. ○ 14 single-company projects.
S35	Industry	<ul style="list-style-type: none"> • Web application suite consisting of 3 related applications: <ul style="list-style-type: none"> ○ An end-user application, an end-user administrator application and a system administrator application.
S37	Industry	<ul style="list-style-type: none"> • 10 small cross-company Web applications.
S39	Academia	<ul style="list-style-type: none"> • Honours/Postgraduate students • 2 datasets: <ul style="list-style-type: none"> ○ Dataset 1 of 34 hypermedia applications ○ Dataset 2 of 25 hypermedia applications
S40	Industry	<ul style="list-style-type: none"> • Cross-company data on Web software and hypermedia projects from the Tukutuku database. • The entire database of 67 projects was evaluated.
S41	Academia	<ul style="list-style-type: none"> • Group projects (5 undergraduate students and a postgraduate project manager). • 44 Web software applications.
S42	Industry	<ul style="list-style-type: none"> • 15 single-company Web applications
S43	Industry	<ul style="list-style-type: none"> • Cross-company data on Web software and hypermedia projects from the Tukutuku database. • 54 projects evaluated.
S44	Industry	<ul style="list-style-type: none"> • Cross-company data on Web software and hypermedia projects from the Tukutuku database. • All 87 projects evaluated.
S45	Industry	<ul style="list-style-type: none"> • 15 maintenance projects on a single Web application.
S46	Industry	<ul style="list-style-type: none"> • 12 single-company Web projects. • All projects developed using the OO-H

		(Object-oriented hypermedia method).
S47	Academia	<ul style="list-style-type: none"> • 3 case studies <ul style="list-style-type: none"> ○ The first two involved 4th and 5th year students. The third involved groups of 2nd year students. ○ Web application developed (number of projects for each study not clearly specified). ○ W2000 used as the design notation for all applications.
S48	Industry	<ul style="list-style-type: none"> • 15 single company Web projects.
S49	Industry	<ul style="list-style-type: none"> • Single application – material purchasing system. • Divided into 2 systems depending on type of user. • System A for intranet users. General software application using client/server paradigm. • System B for Internet users. Web-based application. • Maintenance activities recorded for both systems. 112 maintenance activities in total.
S50	Industry	<ul style="list-style-type: none"> • Cross-company data on Web software and hypermedia projects from the Tukutuku database. • 53 projects evaluated.
S51	Industry	<ul style="list-style-type: none"> • Cross-company data on Web software and hypermedia projects from the Tukutuku database. • All 150 projects evaluated: <ul style="list-style-type: none"> ○ Training set of 120 projects, with remaining 30 projects used as a test set.
S52	Industry	<ul style="list-style-type: none"> • 150 Web applications from Tukutuku database. • Cross-company (international) data. • Each Web project characterized by 25 variables, including length and functionality size measures and cost drivers.
S53	Industry	<ul style="list-style-type: none"> • Cross-company data on Web software and hypermedia projects from the Tukutuku database. • All 150 projects evaluated: <ul style="list-style-type: none"> ○ Training set of 120 projects, with remaining 30 projects used as a test set.
S54	Industry	<ul style="list-style-type: none"> • Cross-company data on Web software and hypermedia projects from the Tukutuku database. • All 150 projects evaluated: <ul style="list-style-type: none"> ○ Training set of 120 projects, with

		remaining 30 projects used as a test set.
S55	Industry	<ul style="list-style-type: none"> • Cross-company data on Web software and hypermedia projects from the Tuketuku database. • Of the 150 projects available: <ul style="list-style-type: none"> ○ 68 were used as a cross-company dataset. ○ 15 were used as a single-company dataset.
S56	Industry	<ul style="list-style-type: none"> • 3 cross-company Web projects
S57	Industry	<ul style="list-style-type: none"> • Single Web project.
S58	Industry and academia	<ul style="list-style-type: none"> • 12 single-company Web content management systems. • 70 Web content management systems developed by postgraduate students. • Estimation model built using industry data, and calibrated using student data. • Evaluated on entire dataset of 82 projects.
S62	Industry	<ul style="list-style-type: none"> • 2 datasets from a single-company: <ul style="list-style-type: none"> ○ 15 Web applications were used to build estimation models. ○ 4 Web applications developed at a later date used to evaluate estimation models.
S63	Industry	<ul style="list-style-type: none"> • 16 cross-company projects.
S64	Industry	<ul style="list-style-type: none"> • Cross-company data on Web software and hypermedia projects from the Tuketuku database. • 53 projects evaluated.
S65	Industry	<ul style="list-style-type: none"> • Cross-company data on Web software and hypermedia projects from the Tuketuku database. • All 195 projects were used to randomly create 2 datasets. Each dataset consisted of: <ul style="list-style-type: none"> ○ A 130 project training set. ○ A 65 project test set.
S66	Industry	<ul style="list-style-type: none"> • Cross-company data on Web software and hypermedia projects from the Tuketuku database. • All 195 projects were used to randomly create 2 datasets. Each dataset consisted of: <ul style="list-style-type: none"> ○ A 130 project training set. ○ A 65 project test set.
S67	Industry	<ul style="list-style-type: none"> • Cross-company data on Web software and hypermedia projects from the Tuketuku database. • Of the 150 projects available: <ul style="list-style-type: none"> ○ 68 were used as a cross-company dataset.

		<ul style="list-style-type: none"> ○ 15 were used as a single-company dataset.
S69	Industry and academia	<ul style="list-style-type: none"> ● 12 single company Web content management systems. ● 70 Web content management systems developed by postgraduate students. ● Estimation model built using industry data, and calibrated using student data. ● Evaluated on entire dataset of 82 projects.
S70	Industry	<ul style="list-style-type: none"> ● Single company – Italian software developer ● Similar information centered Web based systems. ● 10 projects in all
S71	Industry	<ul style="list-style-type: none"> ● Cross-company data on Web software and hypermedia projects from the Tuketuku database. ● All 195 projects were used to randomly create a dataset consisting of: <ul style="list-style-type: none"> ○ A 130 project training set. ○ A 65 project test set.
S72	Industry	<ul style="list-style-type: none"> ● Cross-company data on Web software and hypermedia projects from the Tuketuku database. ● All 195 projects were used to randomly create 2 datasets. Each dataset consisted of: <ul style="list-style-type: none"> ○ A 130 project training set. ○ A 65 project test set.
S74	Industry	<ul style="list-style-type: none"> ● 15 single-company Web applications
S75	Industry	<ul style="list-style-type: none"> ● 15 single-company Web applications.
S76	Industry	<ul style="list-style-type: none"> ● 15 single-company Web applications
S77	Industry	<ul style="list-style-type: none"> ● 15 single-company Web applications
S80	Not specified	<ul style="list-style-type: none"> ● Single Web application.
S81	Industry	<ul style="list-style-type: none"> ● Industry – cross company ● 87 projects from the Tuketuku database
S82	Industry	<ul style="list-style-type: none"> ● 31 single company Web applications. ● All projects were new developments developed using the OO-H (object-oriented hypermedia) method.
S83	Industry	<ul style="list-style-type: none"> ● Data from 6 different Web companies was used
S84	Industry	<ul style="list-style-type: none"> ● Cross-company data on Web software and hypermedia projects from the Tuketuku database. ● All 195 projects were used to randomly create 2 datasets. Each dataset consisted of: <ul style="list-style-type: none"> ○ A 130 project training set. ○ A 65 project test set.

S85	Industry	<ul style="list-style-type: none"> • Cross-company data on Web software and hypermedia projects from the Tuketuku database. • All 195 projects were used to randomly create 2 datasets. Each dataset consisted of: <ul style="list-style-type: none"> ○ A 130 project training set. ○ A 65 project test set.
S87	Industry and academia	<ul style="list-style-type: none"> • 4 Web projects obtained from literature. • Project 1 involved data taken from a set of 12 different Web projects done by post-graduate students. • Project 2 involved data taken from a java based Web portal (industry) • Projects 3 and 4 looked at student data on developing eService applications.
S89	Industry	<ul style="list-style-type: none"> • Subset of Tuketuku database used. • Cross-company • 53 Web projects
S90	Industry	<ul style="list-style-type: none"> • Cross-company data on Web software and hypermedia projects from the Tuketuku database. • All 195 projects were used to randomly create 3 datasets. Each dataset consisted of: <ul style="list-style-type: none"> ○ A 130 project training set. ○ A 65 project test set.
S91	Industry	<ul style="list-style-type: none"> • Tuketuku database. • 8 subsets generated based on: <ul style="list-style-type: none"> ○ Whether or not projects followed a defined and documented process. ○ Enhancement vs. New projects. ○ Whether or not the dev team was part of a software metrics program. ○ Whether or not the dev team was part of a process improvement program. • Subset size ranged from 65 to 130 projects. • Note that databases of non-Web projects were also evaluated. Not considered as not relevant to this SLR
S92	Industry	<ul style="list-style-type: none"> • Single company • 10 Web applications
S93	Industry	<ul style="list-style-type: none"> • Single Italian software developer • Data from 25 Web applications
S96	Industry	<ul style="list-style-type: none"> • Single company • 24 projects
S97	Industry	<ul style="list-style-type: none"> • Single NZ Web company • 22 projects of different sizes and levels of complexity
S98	Industry	<ul style="list-style-type: none"> • Single company • 222 software maintenance projects for

		Web applications
E1	Industry	<ul style="list-style-type: none"> • First field study involved 15 maintenance projects on a single Web-based application (i.e. single company) in the linguistic domain. • Second field study did not involve Web projects so it was not considered.
E2	Academia	<ul style="list-style-type: none"> • Medium sized Web applications • 43 projects, 4 of which were removed due to incomplete information and 2 of which were removed because of unjustifiably high effort estimates. • 4th year students.
E3	Industry	<ul style="list-style-type: none"> • Single Web project from a large software company.
E4	Academia	<ul style="list-style-type: none"> • Web application (hypothetical e-commerce application). • Students worked on the same project; all that differed was the application domain (e.g. books, CDs). • All applications designed using W2000 notation. • Number of applications studied dependent on hypotheses being assessed. Projects with incomplete/incorrect data, as well as outliers removed.
E5	Academia	<ul style="list-style-type: none"> • 43 Web applications of which 6 were not considered due to missing information leaving 37 data points.
E6	Academia	<ul style="list-style-type: none"> • 76 hypermedia projects • Second-year Computer Science students. • 6 projects were removed due to incomplete data leaving 70 projects overall. • Projects divided into 2 groups based on experience of authors; 41 LEL projects (low experience), and 29 HEL (high experience) projects.
E7	Industry	<ul style="list-style-type: none"> • Single company BN created by expert