oi.uchicago.edu

DIYALA PROJECT

Clemens Reichel

The year 2001/2002 proved to be a crucial one for the Diyala Project, which now brings us close to our final goal, the publication of all objects from the Diyala excavations.

It has been — and still remains — a gigantic task, but what has ever been less than gigantic about the Diyala expedition? Between 1930 and 1936 members of an Oriental Institute team of archaeologists, headed by Henri Frankfort, excavated four major sites — Tell Agrab, Tell Asmar, Ishchali, and Khafaje — in the Diyala region east of Baghdad (fig. 1). The scope and size of this project was immense — at Tell Asmar alone, the ancient city of Eshnunna, some

oi.uchicago.edu

Figure 1. THE EYE OF THE ARCHAEOLOGIST: Thorkild Jacobsen, archaeologist and epigrapher of the Diyala expedition, taking notes during the clearance of a large residential quarter at Tell Asmar, Iraq, in 1931/32. Field notes like the ones taken by Jacobsen (see also fig. 3) are our primary source for re-evaluating archaeological data from these excavations

70,000 square meters, over 25% of the total site, were investigated. Over seven seasons, the Oriental Institute systematically excavated temples, palaces, domestic quarters, and workshops. This exposure provided a cross section through most aspects of urban life in ancient Mesopotamia, including political organization, cult and religion, social stratification, and economic interactions. Several soundings reached a depth of 15 meters, providing a cultural sequence from 3000 to 1800 BC, a time span during which the first territorial states in Mesopotamia emerged, large urban centers developed, and writing was invented. Needless to say, excavating at such a large scale resulted in the recovery of many finds, of which 15,500 were registered. The systematic recording of artifacts by level and locus, unparalleled at that time, helped to establish a chronological sequence upon which much of Mesopotamian chronology has rested to the present. The scale of these excavations and interceding World War II are mainly to blame for the fact that, over sixty years later, the publication of this material still has not been completed. Most of the architecture, including temples, palaces, houses, and city walls, has been presented in volumes of the Oriental Institute Publications series. Some 5,000 more "spectacular" finds, such as sculpture, relief art, and cylinder seals, as well as pottery types, also have appeared in the several volumes of the same series. Over 10,000 objects, however, have so far remained unpublished. In the original publication plan devised by the excavators, they were meant to appear in a volume called "Miscellaneous Objects." This title, however, is somewhat of an understatement and misnomer, considering that these objects include stone vessels, tools, weapons, jewelry, cosmetic sets, weights, figurines, stone and metal vessels, inlays, stamp seals, clay sealings, and some 1,200 cuneiform tablets. In the past decades several publication attempts were started, but the large number of objects to work with has made data management difficult if not altogether impossible.

By 1994, the year when the Diyala Project was launched by McGuire Gibson with the support of an NEH grant, desktop computers had helped to make data of this size and magnitude more manageable. As a first step, all of the over 15,000 catalog cards and field register entries, such as the ones shown in figure 2, were entered into database files. Moving data from paper and catalog cards into a computer database was a significant step, yet only gradually did we learn to utilize its potentials. Paper records, by definition, are two-dimensional. Cross-references between different catalog entries on paper have to be added and updated by hand. The relationship between datasets, however, is often not linear. How, for example, does one

oi.uchicago.edu

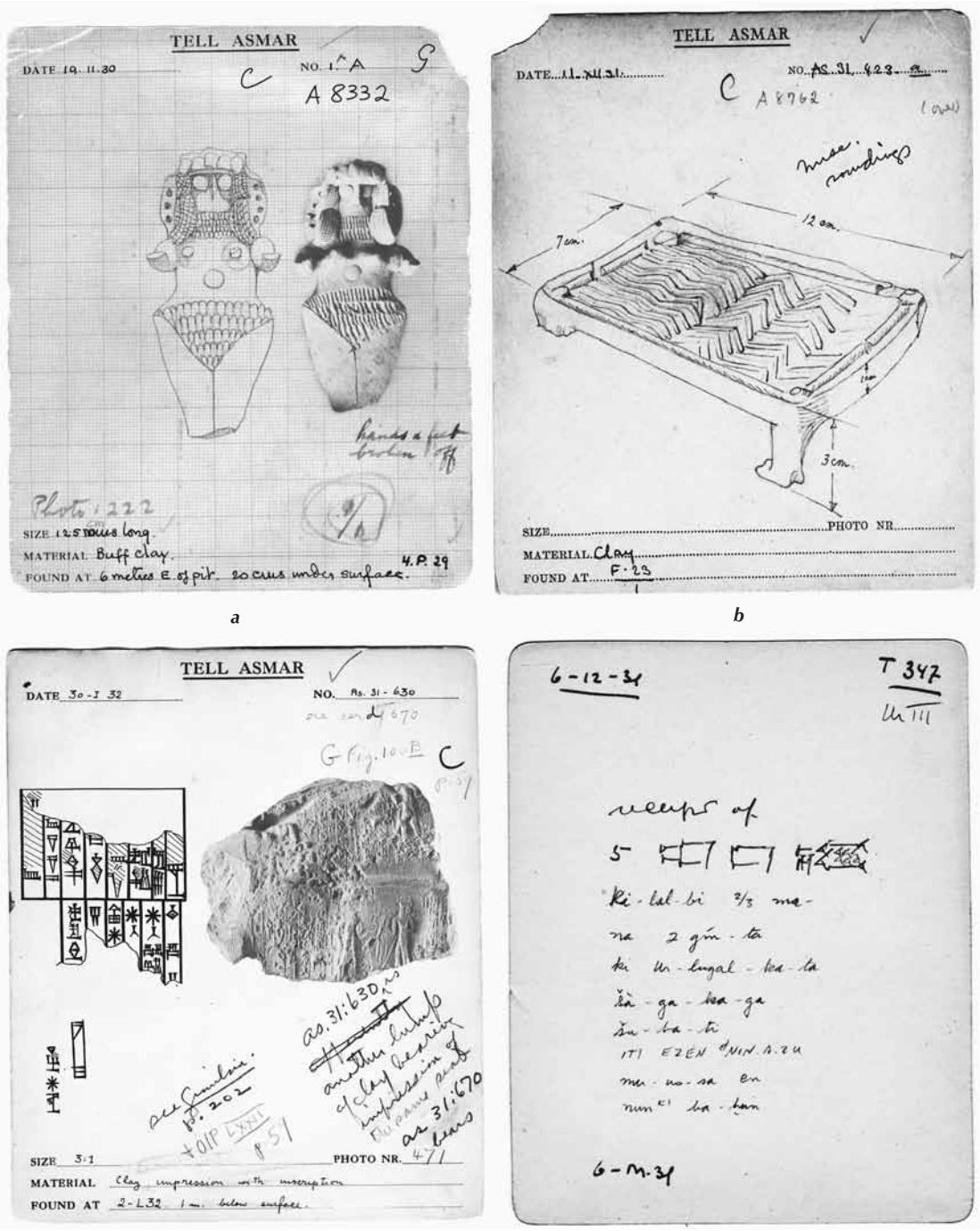


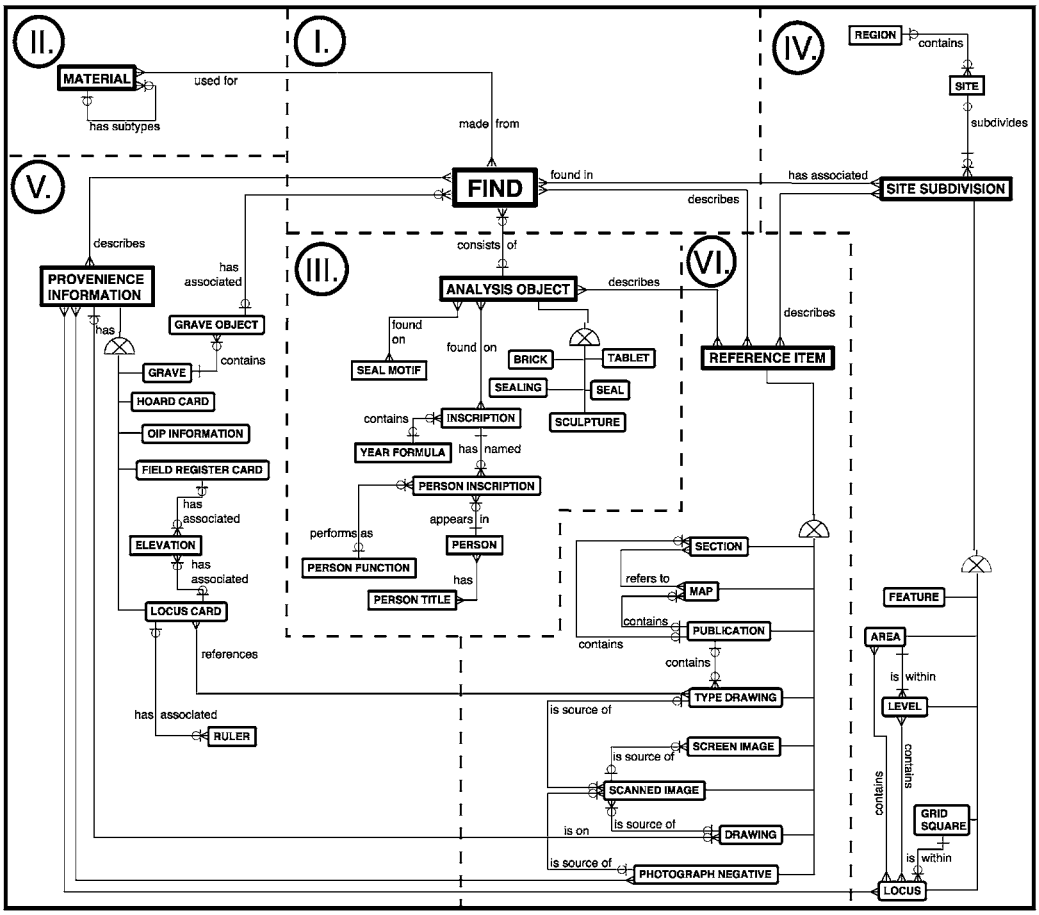
Figure 2. CATALOGING OLD RECORDS: Four object cards from the Diyala excavations, showing varying types of object descriptions: (a) terra-cotta figurine of a naked woman (ca. 1900 BC); (b) clay model bed (ca. 1900 BC); (c) clay sealing with two impressions of the seal of King Shuiliya (ca. 2020 BC); (d) Ur III tablet: (transliteration of text) receipt for five bronze mortars, received by an individual named Shagaduga from an official named Ur-lugal, in month 5 of 2038 BC (= Amar-Su'en, year 9); the signs for "mortars" were not transliterated but copied as cuneiform signs. Find numbers are listed in top right corner, with Oriental Institute Museum numbers below them in (a) and (b)

oi.uchicago.edu

best catalog a tablet with a cuneiform text that contains the names of thirty people? All thirty names could be added to the catalog card with the description of the tablet or the translation of its text, but that makes it impossible to sort these cards alphabetically for a prosopographic (name-based) studies — which name would be the decisive one, and how would the other twenty-nine be found? To handle this data properly one also has to file thirty separate cards for each personal name in addition to the entries on the tablet description card. Double bookkeeping, however, is high maintenance and subject to errors. A computer database program with relational capabilities makes this work considerably easier. To stay with our example, the description or translation of the cuneiform text would go into one table, the personal names into a different one; one entry in the first table corresponds to thirty entries in the second one. When these two tables are linked ("related") to each other, a user can obtain answers in seconds that would otherwise take hours, if not days, of painful manual searches. In our example he may want to obtain descriptions of all or a chosen set of cuneiform texts, each of them followed by a list of personal names found in them. Then he may want an alphabetically sorted list of personal names from all texts, text numbers in which these names occur, and even see the text lines excerpted in which these names occur. Such results can be returned within seconds and without duplicating any primary data entry.

This is only one possible example of how splitting up data into separate, linked tables can refine the search capabilities within a computer database. Archaeological objects can be composites made of several materials. A user may not only request the composition of an individual object, but also all occurrences of one specific material — a job easily done if all material components of an object are entered separately into a related table. A detailed description of an archaeological findspot, which in a printed catalog either has to be repeated for each object from the same context or to be cross-referenced manually ("for archeological provenience see …"), only has to be entered once into a table and displayed for each object from the same context, making corrections very easy. Drawings, plans, and pictures of objects can be linked to text displays, conveniently merging data from multiple sources on the screen for further evaluation and interpretation.

The benefits of a relational database were not lost on us. Between 1995 and 1997, the number of tables increased to fifteen. Further data entry on objects, notably on tablets and sealings done by Clemens Reichel since 1997, on weights by Colleen Coyle in 1998 and 1999, and on stone vessels and inlays by Claudia Suter between 1996 to 1999, as well as further refinement of the data layout increased the number to over sixty tables. Impressive as this number seems, it is dwarfed by the number of programs, procedures, and screens, which exceeds 2,000 by now. Seven years of work have shown clear rewards. In June 2001 Clemens Reichel, who acted as project coordinator since 1999, submitted his dissertation on the Palace of the Rulers at Eshnunna. His work, which incorporates data gathered from cuneiform texts and clay sealings in its analysis, showed us that relational databases can support both archaeological and philological research. As reported in previous *Annual Reports* (see *2000–2001 Annual Report*), our changes in data management had also modified our outlook on publication. We dispensed with the idea of a book publication and instead opted for a web-based dissemination of the data. But we also began to realize our limitations in the world of computer databases. None of us had any previous experience with electronic data management. We were learning as we moved along, but our lack of experience had resulted in flaws within the data structure and in inconsistencies in data entry. We also desperately needed to change the supporting software. For years we had been using Microsoft's FoxPro, a relational database application that had served us well, but gradually we had outgrown it and its limitations — such as its



Relationships between database files are expressed with the following symbols:

A ——— B : ONE-TO-ONE RELATIONSHIP: Single lines connecting Table A and B indicate that an item in one table has a relationship to ONE item in the other table.

A ——<< B : ONE-TO-MANY RELATIONSHIP: "Crowfeet" on the right side indicate that an item in Table A can have relationships to ONE OR SEVERAL items in Table B, while an item in Table B may only have a relationship to ONE item in Table A.

A >——<< B : MANY-TO-MANY RELATIONSHIP: "Crowfeet" on both sides indicate that items in either table have relationships to either ONE OR SEVERAL items in the other table.

A >o|——o| B : MANDATORY / OPTIONAL RELATIONSHIP: Circles indicate optional relationships between tables. An item inTable A MAY have a relationship to ONE item in Table B. Added "crowfeet" on the left side indicate that an item in Table B MAY have relationships to ONE OR SEVERAL items in Table A.

A —— B C D : SUB-TYPES: Tables B, C, D are SUB-TYPES to Table A.

Figure 3. STRUCTURING ARCHAEOLOGICAL DATA FOR THE COMPUTER: Just like a file cabinet, a computer database requires careful planning in its layout and internal organization in order to be fully functional. This model shows the layout for the main tables of the Diyala database. Connecting lines indicate relationships and file hierarchies. The significance of "Zones I–IV" is described in the main text

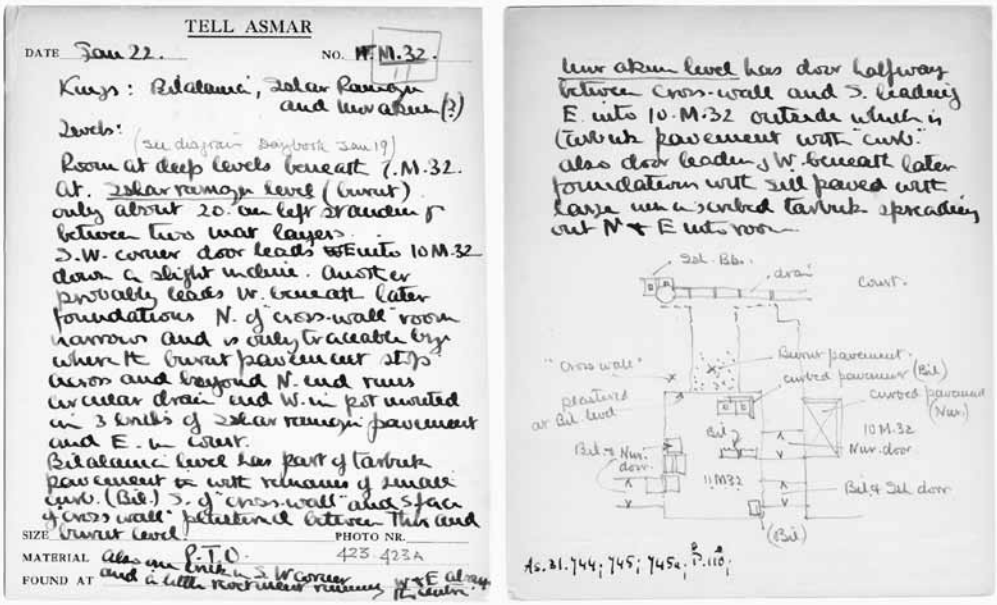
oi.uchicago.edu

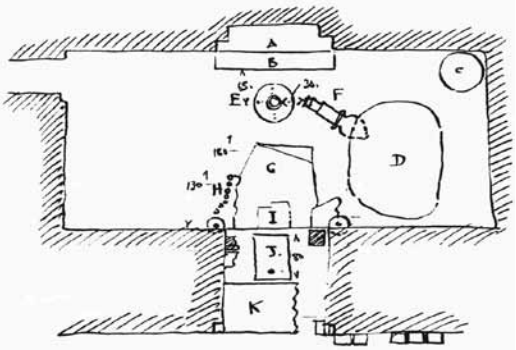
incompatibility with a web-based publication — were becoming obvious. By 2000 we decided to switch to Oracle, the market leader for relational databases in the business world. What we needed now, however, was the help of a professional.

We found him at exactly the right time. In November 2000, George Sundell, who had just retired from his job as data architect at Ameritech, joined the Diyala Project as a volunteer. Creating structural layouts for database projects was George's bread and butter. He also has a background in American archaeology and has participated in excavations; quite clearly we could not have found a better person to help us. Clemens and George started working on a new and improved layout for the Diyala Project that would use Oracle 9-i as supporting application and that would enable eventual web publication. This was easier said than done because George had to disentangle years of work on a database that had grown organically without a real preconceived layout. Many ideas were good, but some led nowhere, and there were obvious flaws and shortcomings. Over the last two years George has been creating more and more refined fundamental, logical, and physical models of how the data should be structured.

George's present fundamental model, which shows many but not all of the tables, is shown in figure 3. Names of tables are shown in square boxes, the principal names being rendered in larger type font. Relations between tables are expressed by connecting lines; some added symbols, explained in the legend below the chart, further specify the nature of their relationships. The nature of the data allows this layout to be divided into six zones (I–VI), which are marked in the diagram. Zone I contains the final, revised, information on each artifact; this is the nodal point of the layout to which every table is linked directly or indirectly. Materials are kept in separate tables (Zone II), which allows the listing of all composites as well as the ranking of primary and secondary materials. With over 15,000 objects we are dealing with a large number of very different object categories, each of them having their own characteristics. It would be impossible to squeeze all data for, say, bone tools and cuneiform texts into the same data fields. Information that is specific to such object categories is therefore accommodated in a separate zone (Zone III). These tables contain "dynamic data," data that will develop as individual researchers work on certain object classes. Personalized log-ins will allow authorized scholars to add or modify data and be credited for it. Varying interpretations from different scholars will be stored and displayed separately, a feature that we believe to be vital in ongoing research. The tables listed at present for Zone III reflect Clemens's work on tablets and sealings, but more tables for other object categories can be added as needed. Information on the objects' archaeological provenience is contained in Zones IV and V. The allocation of this data into two different zones is due to its nature. While existing descriptions of an object can always be updated, corrected, and improved as long as the object itself is available for analysis, we cannot add to an excavator's observations for archaeological provenience; every piece of information is therefore vital. The sources for archaeological contexts (fig. 4) consist of notebooks entries, locus cards, field plans, sketches, grave and hoard descriptions, and published reports, forming a data pool that grew over the years of excavations with often inconsistent recording conventions. To complicate matters further, different excavators occasionally interpreted the same evidence differently, resulting in contradictory assignments of level or locus for some artifacts and different dates for archaeological contexts. The quality of record-keeping often made re-evaluations possible and allowed us to prepare stratigraphic work sequences for areas, levels, and loci on all Diyala sites; this sequence is presented in tables within Zone IV of our diagram. In order to give scholars outside of Chicago the chance to evaluate our assignments and potentially to disagree with them, we decided to also make the field records that we had at our disposal for the stratigraphic reanalysis available to the

oi.uchicago.edu





c

Figure 4. (a) Obverse and (b) reverse of catalog card with detailed description of an archaeological context (Locus M32:11 at Tell Asmar), including an annotated sketch plan. (c) Sketch plan from an excavator's field notebook, identifying certain installations in a room. To a layperson these notes may seem insignificant, but to archaeologists they are primary sources to re-evaluate archaeological data, often full of unpublished information, and therefore as valuable artifacts. There are several thousand locus cards and dozens of notebooks. Eventually all of this data will be part of the publicly available Diyala archive on the Internet

rest of the world; they comprise the tables of Zone V. These table are linked to the ones of Zone IV, providing the user with the full array of published as well as unpublished records for each archaeological context. Zone VI contains the reference items necessary for analysis, which includes publication references as well as photographs and drawings of objects.

While our immediate goal is publication of the objects, our long-term plan extends far beyond that. We plan to scan all archival material, published and unpublished, such as plans, sections, excavation photographs, notebooks, locus cards, and sketches, and include it in the database (fig. 4). In other words, we will add an on-line Diyala "archive" to the on-line object publication. Based on the original sources, the user will be able to evaluate the excavators as well as our conclusions. What currently requires a trip to Chicago and days, if not weeks, of work in the Oriental Institute archives will be available worldwide with little more than a few mouse clicks. One may ask whether this is necessary or whether we are overburdening the user with this amount of data. As of today, relatively few people may be interested in this amount of data, but we believe that this situation may change very soon. In the past twenty years more and more scholars, some of them from Chicago, have re-analyzed important architectural complexes from sites such as Nippur, Ashur, and Ur, sites that had

oi.uchicago.edu

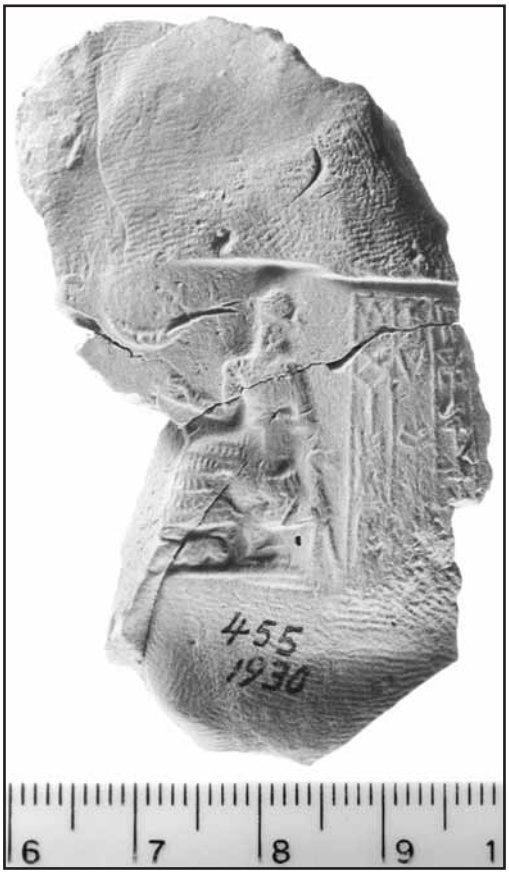


Figure 5. Clay sealing from Tell Asmar (Eshnunna) with impression of a seal. The seal legend identifies the seal owner as Idam-arshi son of Buṭṭutu. The preserved part of the seal scene shows a seated god with horned cap and flounced garment holding a cup in front of a moon crescent. The throne upon which the god is seated has the shape of a snake-dragon, attribute of Tishpak, the city of Eshnunna (ca. 1900 BC)

been excavated and published decades earlier. By re-analyzing original field data from the excavations they were able to make substantial additions and corrections to the published reports. Since these sites are of key importance to our understanding of Mesopotamian archaeology and history, the fundamental importance of "re-excavating," on paper, old excavations can hardly be overstated. As Clemens's dissertation has already shown, the sites in the Diyala have similar potentials. But at this time, other scholars would have to come to Chicago to study the Diyala field notes, and they would have to be very familiar with the nature of these records to even know what questions to ask. Making all of the archival data from all Diyala sites available via the World Wide Web will make their work much easier and hopefully stimulate more studies on the Diyala and other key sites in Mesopotamia.

Our vision is gradually turning into reality. At this point, George has created all of the tables in Oracle on a PC in the Oriental Institute's computer laboratory, and by late September we hope to have completed the data transferal. Designing an initial set of screen layouts and programs will probably take us until the end of this year. While developing the database we are operating with a single user license for Oracle; once we will have obtained a site or multi-user license we are ready to go worldwide. The help and support of John Sanders and of the computer lab is gratefully acknowledged here.

So much for our future, but no longer futuristic, plans. In the present, and on a more basic level, work has been continuing as well. Clemens has continued his work on Diyala tablets and sealings and is currently preparing his dissertation for publication. We are glad to report that Betsy Kremers, our object photographer for the Diyala Project since 1998, was able to rejoin us in June after a long break due to hip surgery. Her exceptional skills and care in taking photographs had been sorely missed. Despite such adverse circumstances she was able to add another 1,200 new object photographs in 2001/2002, mostly of tablets and sealings (fig. 5), to over 4,100 photos taken between 1998 and 2001. Thanks to her efforts we now have 15,000 photographs of about 7,700 objects, an enormous achievement by everyone's standard. Pointing out that so far we have photographs of just about 50% of all Diyala objects (many of them in Baghdad) by no means diminishes Betsy's achievement; it underscores the enormous task of getting this data published. As we said in our introductory statement: nothing about the Diyala seems to be less than gigantic.

— — — — — — — — — —