The K3P Knowledge Protection Model
To Separate Know-Why from Know-How in Engineering Contexts

Authors: Lars Lindberg
         Gunilla Sivard
ABSTRACT

Today, much emphasis is put on globalization, collaboration, integration, and standardization in the engineering world. While this generally is both necessary and desirable, it does raise one important question: How does a company protect its key competence against imitators, copycats, and plagiarisms? In a project we have studied this question from a product data perspective, i.e., we leave out legal-, commercial-, and IT-aspects. The traditional approach has been to rely on the physical protection of the drawing archive combined with simple security classification on the documents themselves. We found this to be too simplistic, and instead we have developed a more elaborate, but still tentative model to identify the data to be protected. This model starts with product characteristics, e.g., product type, product concept maturity, expected life, and degree of customization. Further, the business strategy, e.g., customer focus vs. technical leadership vs. operational excellence is described. These product and business properties are mapped to the relevant processes and methods within the company, e.g., product development or order handling, as well as to the supporting systems and information structures. Finally, mapping is made to the individual data elements which are to be protected. In this model we do not only identify the data elements themselves, but also the context within which they need protection. This context can thus be a customer order, a development project, or a certain time frame. We illustrate the use of our model with industrial test cases and conclude with a proposal for future work.

INTRODUCTION

The discrete manufacturing industry is undergoing significant changes in the last decades. Globalization, extended enterprise, collaboration, voice of the customer are but a few of the buzzwords that tries to capture these changes. Paired with the ITC development of communication networks, digitally stored information and standards allowing simple and unambiguous communication, the basis for protection of product data related to the company’s key competencies is changing radically.

Lars Lindberg, Gunilla Sivard: KTH Royal Institute of technology, Industrial Technology and Management, 10044 Stockholm, Sweden
This paper reports the results of a prestudy primarily looking into the possibilities of identifying the parts of a company’s product data (as part of the company’s intellectual property) which are essential to protect. Being a prestudy there are several limitations:

- We look at the product data only. By product data we mean data directly related to the product or the parts building it up. Drawings, bills of material, 3D- and other models, analysis (e.g., FEM) are examples of product data, while, for instance, patents, copyrights, agreements and other legal data, as well as contracts and other commercial data are excluded, as is the physical product itself.
- The implementation in IT-systems of any protection or classification mechanisms is out of our current scope, e.g., firewalls, encryptions, user authentication and permissions, etc.
- We focus on mechanical products and data for these, and the number of industrial cases is limited.

Our main result is a proposed model for classification of product data. We will present this model in some detail, but before that we give the background, motivation, and approach for our work.

**BACKGROUND, MOTIVATION, AND APPROACH**

Even though intellectual property often is referred to in the literature concerning product data, e.g. [1], we have not found any method to classify the product data with respect to the intellectual property. In [2] an elaborate scheme for access control is presented, but this assumes that the data is properly classified. The same is true for commercial IT-tools, e.g., Enovia [3]. Product data classification is discussed in [4], but from the perspective of unambiguous data exchange. Technical and business aspects of intellectual property protection is discussed in [4] and [5] respectively, without any application to product data.

In figure 1 we show a high level, generic process view of a manufacturing company. The Product Planning process is responsible for the product portfolio and the scheduling of new products and product releases. In Product Development the most of the product design, process planning and sourcing take place. Sales to Order is the marketing and sales process, Order to Delivery the manufacturing and delivery process, while Delivery to Repurchase is responsible for aftermarket activities like support and service. We will discuss these processes from an intellectual property point of view in more detail below.

Traditionally, companies have done essentially all work in these processes in-house, with their own staff. The primary product data in this setting consists of the drawings and bills of material which, when properly approved, are released to the production processes, i.e., StO, OtD, and DtR. This data is generally documents, stored on paper or as individual files on a file server. Very little of this product data is ever distributed outside the company. Purchase drawings of detail parts are sent to suppliers, and dimension prints, installation drawings, and service manuals are sent to the customer, but that is (or rather was) the extent of it.
The mechanism used to protect product data is a simple classification, combined with equally simple rules for assignment and usage. The classification typically is: Open, Internal, Restricted, and Secret. The classification assignment is on the document level only, and the assignment is made on broad classes of documents, e.g., all drawings are ‘restricted’, except purchase drawings which are ‘internal’. Here ‘internal’ means that everyone inside the company can have access to it, as can select external suppliers. Restricted on the other hand, means that the document cannot be distributed outside the company.

There are several trends which together changes the rather idyllic picture we painted in the previous section. The industrial landscape is changing:

- Globalization is a strong force today, meaning that the many companies have not only customers everywhere, but also competitors everywhere. There is no such thing as a (safe) home market anymore.
- Various forms of partner networks, both regarding product design and production, are common to the point of being the rule today. This means that product data must be shared with many actors outside the company. These actors come from many cultures, with potentially different views on IP and related matters.
- The product complexity and variability are increasing, which means that more data must be managed. An example of the increasing complexity is that many products today contain a mix of mechanical parts, electronic circuits, and software. The increasing variability means that more elaborate processes are required to specify, sell, manufacture, and service these products.
- Regulatory requirements, e.g., regarding safety documentation or material declarations, increase the amount of public data about the product.

In addition we have several ITC trends:

- Development of networks in a broad sense, including for instance, internet, mobile communication, and cloud computing. This development means that in principle all computers, all data are accessible from anywhere in the world – and that with just a few clicks and keystrokes.
- International standards, from ASCII to STEP, are specifically designed to make sharing of digital data simple and unambiguous.
- Data is stored in well-structured formats, e.g., databases with standardized access methods. Some applications, most notably PLM-systems, boast that they provide a ‘single point of truth’. But that also means that they provide a ‘single point where I can get everything’.
One consequence of the above trends is, from an IP-protection point of view, that the traditional way of protecting product data, i.e., security classification of manufacturing documents, is no longer sufficient. First we distribute many of these documents to partners and customers and the classification is then often dependent on the recipient rather than the document itself. Second, it may very well be a part of a document that should be protected, not the complete document. It could, e.g., be a tolerance or a material specification. Third, there are several additional data types that may require protection, e.g., analysis programs and results, test methods and results, manufacturing processes, etc.

From the discussion in the previous section we identified three areas to address: how to handle more data on a more granular level than a document, how to handle the context where the data was used, and how to handle a wider set of product data than just drawings and bills of material. The first area requires detailed information models of product data. We can easily find this type of models; the standard ISO 10303 (a.k.a. STEP) provides several standards for our domain [7]. However, these standards do not provide any explicit security classification mechanisms, so some extensions to these standards are necessary. The two other areas, classification with respect to context and new types of data, are the focus for the reference model we propose below.

THE K3P KNOWLEDGE PROTECTION REFERENCE MODEL

Overview

As we suggested above, we need a framework where we can structure the information required to properly classify the product data from an IP-protection point of view. The model we introduce here is currently somewhat tentative, but we intend to elaborate and verify it into a full-fledged reference model.

In short, the levels represent:

- **Level 1: Product and Business Strategy.** These are the key concepts for any business, and we use them to capture the key characteristics of each from a product data perspective.
- **Level 2: Processes, methods, and Information Model.** This layer which contains the actual ways of working within the company and the corresponding information structures, which support the business in implementing the strategy and realize the products.
- **Level 3: Information systems.** Here we have the ITC systems and other tools and instructions implementing the information model and supporting the company processes.
- **Level 4: Data objects.** At the lowest level we have the objects that we somehow need to classify and possibly protect.

We elaborate each of these levels in the sections below. Our current assumptions are thus that first we can identify proper characteristics on each level and second that there are paths, links, or relationships between the characteristics of the different levels which we can identify and use to specify the security classification of the data objects. The model has been inspired by the MDM approach to modularization [8] and the system selection criteria outlined in [9].
Level 1: Products and Business Strategy

At the core of a manufacturing company lie the products it brings to the market, and the strategies it employs to do so. Below we list some characteristics for these concepts and how they could affect the data requiring protection.

PRODUCT

- Position on the scale from standard (catalogue) product – customized product – one-off product.
- Typical expected life of the product.
- The maturity of the product ‘type’, i.e., is it a newly introduced product (e.g., a tablet like iPad) or is it a mature product (e.g., a ball bearing).
- External rules, regulations and standards that the product and company must fulfill.

BUSINESS STRATEGY

Here we use the model from “The Discipline of Market Leaders” [10], which introduces three value disciplines which a company could focus on: Operational Excellence, i.e., optimize processes to deliver lowest cost, Product leadership, i.e., focus on being first-to-market with new products to achieve price premium, Customer intimacy, i.e., tailor and shape the products exactly to the customer’s need to build long term loyalty.
Level 2: Processes and Information Model

This is the level where the actual job is done in the company. Here everything from the top level process map to the workflow for releasing a document should be included. However, here we limit ourselves to the top levels. We have already introduced a generic process model, see figure 2 above. We outline the role of each of the above processes from our IP-perspective below:

- **Product Planning** – This process is clearly highly sensitive from an IP perspective. Here the company defines the performance characteristics and release schedule of its products. However, from a data protection point of view, it is probably less critical, of a couple of reasons: First, the product planning is generally an internal process, with few and well controlled interactions with external customers or suppliers. Second, the staff awareness is high concerning the sensitive nature of the information in this process. Third, not much product data is produced, which makes it easy to control. Clearly there are strong dependencies on Level 1; it is after all here that the product and the strategies are coming together. However, since we do not believe that Product Planning is crucial from our point of view, we do not elaborate these dependencies.

- **Product Development** – This is the key process from our point of view:
  - This is the primary source of virtually all product data; also product data created in downstream processes generally uses Product Development data as a starting point.
  - The data volumes are generally big, often with high complexity and many interdependencies.
  - In this process there are often many interactions: both with staff from other processes and with external partners. Of special interest from our point of view is the use of external design partners. This type of co-operation poses special problems: The partner must get sufficient information to do the job, but not more than that.
  - Since this is where much of the product data is created, we believe that it is here that much of the identification of ‘know-why’ data must take place.

- **StO** – In this process there are obviously many interactions with external actors. Here we have a clear dependency on the product characteristics and also the business strategy:
  - For catalogue products, little unpublished product data is used in this process, but for customized products technical specifications and similar documents are critically important for the business.
  - Depending on the business strategy, more or less product data may be managed in this process. For example, a business focusing on ‘customer intimacy’ is likely to provide more specific data to its customers than a business focusing on ‘operational excellence.’
  - One issue, which arises if the ‘know-why’ data is not clearly specified, is that a sales engineer may be tempted to provide the customer with ‘too much’ additional information, in order to win the business.
 OtD – Again, this is a process where there are many interactions with different types of partners, and again there are dependencies on the product characteristics and the business strategy:
  o A specific manufacturing process can give the company a competitive edge and be just as important as, e.g., a specific design solution. Examples are heat treatments, welding methods, etc.
  o The sourcing strategy (use of manufacturing partners and component providers) is clearly linked to the business strategy and also to the type of product.
  o Suppliers and other partners must be provided with sufficient data to manufacture and deliver components meeting the specifications and quality requirements.
  o The sourcing data contains (and must do so) quite detailed information, e.g., dimensions, tolerances, material data, surface treatments, etc.
  o As part of the delivery, the customer can get (depending on the product characteristics) more or less extensive information about the product. There is a potential conflict between what the customer wants and what the company wants to provide. An example could be calculation programs.

 DtR – This is a process where there are significant differences depending on the product characteristics. For a mobile phone manufacturer, the aftermarket is more of a nuisance (we should buy a new one…), while it for a manufacturer of construction equipment could provide significant profit.
  o The relationship to partners is heavily dependent on business strategy and product characteristics. Spare part lists may for instance be freely available, or carefully guarded depending on the strategy.
  o As a general remark, we should emphasize that protecting data is more difficult in this process than the other: the product is now physically available and can be scrutinized; many people are aware of the product, its operation and functions; and the product life increases the exposure of it.

 Information is ‘the ghost in the machine’, i.e., the virtually invisible thing that makes the wheels turn in a manufacturing company. Information is ubiquitous and a commodity in the sense that its absence is more noted than its presence.

 A manufacturing company is today dependent on an often large set of IT-systems. However, it is quite rare to find an up to date and complete information model, even for the limited domain of product data. Models developed for specific projects cover only the project scope and are rarely maintained after the project finishes. Instead, information models are implicit, either in the IT-system or in product documentation rules and standards and similar documents. From our point of view this means that an important building block in the traceability from the top level (product and business strategy) to the bottom level (the information objects) often is missing.
Level 3: Information Systems

For our purposes here, we find the simple system architecture or system map outlined in figure 4 sufficient.

Here we first have a set of authoring (or desktop) applications, where most of the product data is created, used and modified. Sometimes there is an intermediate layer tailored to manage the files from one or several applications. Next we have the Product Lifecycle Management (PLM) system, where the product data, is coming together. Finally we have the ERP-system, which here is a catch-all for all systems used to handle the physical flows of components, parts, and products. A simplified view is that PLM manages the virtual product while ERP manages the physical one.

From our point of view we can identify a number of issues here. Overarching issues are the lack of common concepts, different security mechanisms, and that one information object can be represented in several systems. We summarize these below:

- The ways to protect data is different on the different layers of the architecture above: desktop applications typically rely on the file system security, while database applications generally have their own security mechanisms, sometimes using the OS functionality (e.g., Active Directory) for authentication. This means that it is difficult at best and impossible at worst to create a common security policy for all systems and information objects.
- One object, e.g., a part, is often represented in several applications. The part is e.g., typically represented with the 3D-model and the drawings in a CAD-system, the part structure in the PLM-system, and the sourcing data in the ERP-system. With the previous point it is clear that it is difficult to provide a consistent access control and security classification across systems.
- Since many of the systems use similar terms, e.g., state, project, lifecycle, product, but with sometimes widely different definitions, it is non-trivial to build a common information model.
- The granularity of the data represented and the possibilities to classify it varies substantially between systems. In a PLM-system it could, for instance,
be possible to assign a security classification to a property and its value (‘Rated Voltage = 150 kV’) while this data would be included in a data sheet in the ERP-system, where the security classification is on the document level.

- Traditional software development methods do not provide mechanisms to specify and verify security classification, especially not across a complete application portfolio.

We should also again stress, as we did in the introduction, that the progress being made in information technology means that we now have virtually all product data available in digital form, and often highly accessible through well-designed (sort of…) information structures. This means on the one hand that anyone getting access to the digital archives is able to quickly extract significant volumes of product data, which is not possible in a traditional, document oriented archive. On the other hand gives

This would not be possible in a traditional document oriented archive, where access is only granted to the very few, and it takes considerable time to find and copy all documents.

**Level 4: Data**

At the lowest level we have the data objects, some of which we want to be able to protect. Yesterday these were represented as documents, on paper, aperture cards or similar media. The physical access was thus limited, and so was the distribution, since much of the manufacturing was made in-house. Then a simple security classification, for instance ‘Open’, ‘Internal’, ‘Confidential, ‘Secret’, on the documents was adequate. Today the information objects are stored in digital formats, and represented as files or database records or some combination of these. Even though the old security classification still is in use, it is now inadequate, and the key challenge we are facing is how to address this. We have to:

- Identify the objects to protect.
- Find a mechanism of clustering them; it is not realistic to classify millions of records in a plethora of databases.
- Provide the proper protection mechanism, which must be possible to maintain.

One issue is, as we mentioned above, the granularity of the data. We can possibly protect a drawing, but what if the only data we need, must and want to protect is a tolerance value? Another important aspect is context: ‘Power = 120 kW’ does not say anything in isolation. That company X is planning to develop a new engine with ‘Power = 120 kW’ may be a vital piece of information. A picture of a car in a product pamphlet with caption ‘Engine power = 120 kW’ may be informative, but completely harmless from an IP point of view. A third aspect is time: the body of a new car is a kept secret until the new model is presented, but after that it is not only open information, but the car company want that everyone know it.
SUMMARY AND FUTURE WORK

Above we introduced a model to aid in the classification of product data from an IP-protection point of view. Our preliminary results so far are promising, but we need to further validate and elaborate the model based on industrial cases. We also need to verify the model with respect to other product development approaches, e.g., modularization and design patterns.

ACKNOWLEDGEMENTS

The authors wish to thank our industrial partners Cargotec Marine and Siemens Industrial Turbines for valuable inputs, Vinnova for financially supporting the project and the XPRES lab at KTH for a productive environment.

REFERENCES