Università
Ca'Foscari
Venezia

Corso di Laurea magistrale in
Informatica – Computer Science

Tesi di Laurea

—
Ca' Foscari
Dorsoduro 3246
30123 Venezia

Business Meta Model: A Structured Framework for Business Models Representation

Relatore
Ch. Prof. Andrea Albarelli

Laureando
Alessandro Popaiz
Matricola 817757

Anno Accademico
2011 / 2012

Abstract

Given the central role of Business Models in company characterization, it is not surprising that a great deal of effort has been spent in studying suitable representations for them. Most of the proposed models, however, pursue a semi-formal human-readable graphical paradigm that is mainly meant to be discussed among stakeholders, rather than to be easily handled by information systems. In this thesis we introduce a formal meta model that aspires to be general enough to capture the expressiveness of most currently adopted paradigms. At the same time, each produced Business Model instance is regular and structured enough to be processed through automated algorithms. Specifically, data are organized as a structured graph, allowing for the adoption of well-known graph-based mining techniques. The ability of the frame- work to deal with real-world scenarios is assessed by modeling several actual companies. Further, some examples of data processing are given, specifically with the aim of spotting common patterns within a database of Business Models.

Contents

1. Introduction

Formal representations are widely adopted within enterprises to model different breeds of relevant information. In fact, the acceptance of a common and standardized syntax is key to many critical tasks, such as the effective sharing of knowledge, the design of software or the reasoning over business processes. In order to better serve the many purposes of information representation, different paradigms have been introduced during the last few decades. It is safe to state that the historical roots of information and processes formal modeling blend with the massive introduction of computers in business.

Early formalism were of course derived from mathematical notations, such as those adopted in axiomatic set theory, lambda calculus, and first-order predicate logic. Some of them were more akin programming languages. For instance, APL (Iverson, 1962) was a seminal concise symbolic language that offered a high level of abstraction and supported recursion for the modeling of both hardware and software systems. The Z notation, introduced more than a decade later (Abrial, 1974), introduced a semi-graphical formal specification language loosely based on set theory that could be used to describe information systems in general.

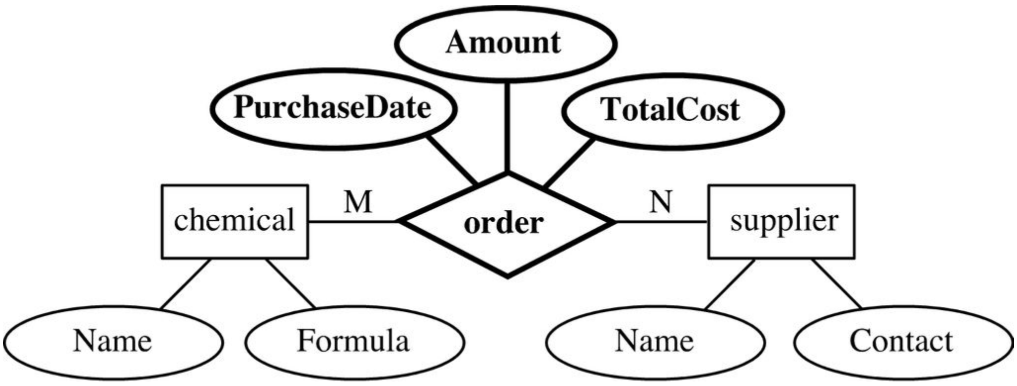

Figure 1 - An example of an order data structure represented with the original Entity-Relationship graphical formalism introduced by Chen. Entities are drawn as boxes, relations as diamonds and attributes as ellipses. Entities are connected through relations by lines showing the cardinality of the relation (in this case, M to N).

At roughly the same time period, however, it was possible to witness the flourishing of what can undoubtedly considered the formalism that set the standard for most modern time structured data representations, the E- R (Entity-Relationship) model (Brown, 1975, Chen, 1983, 1976). Within this model, entities are instances of a specific entity type, which corresponds to object classes, and can be linked by means of relations, that capture how entities are connected to one another through a role that is similar to the usage of verbs in natural language. Indeed, the analogy between the structure of the E-R model and plain English sentences has been a leading feature of its design (Hartmann and Link, 2007). The original model proposed by Chen pro- vided graphical formalism to define entities, relations and attributes in order to give a bird's eye view of the information model (Fig. 1). It is first and most successful usage of the E-R is related to the modeling of relational databases, however, during its forty years long life, its field of application has proven to be very broad and many derived models have been proposed. In (Hardwick, 1984) the limitations of classical E-R with respect to design applications are discussed and an extended model adopting an enhanced algebra is introduced. In (Elmasri and Wuu, 1990, Ferg, 1985, Gregersen, 2005) (and many others) extensions to the basic E-R model that accounts for temporal information are proposed. Business requirements and constraint are added to the classical conceptual modeling in (Khan et al., 2004), where they are modelled as additional attributes. In (Liu et al., 2005) the authors introduce Security Extended ER (SEER), combining security control mechanisms and standard Entity Relation models.

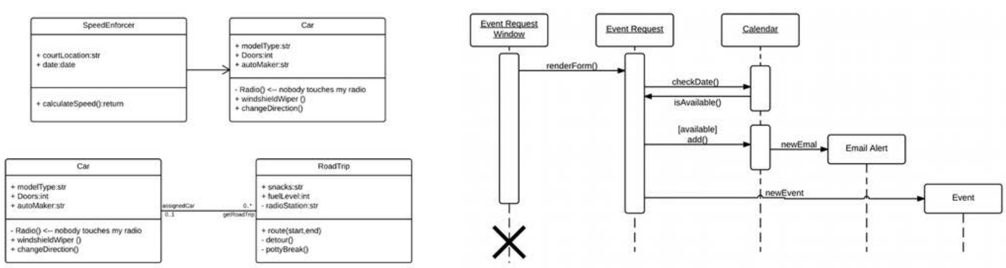

Figure 2 - Some examples of UML diagrams. On the left side a class diagram with both one to one and many to one relations. On the right side a sequence diagram that represents the activities to be performed in order to set and notify an event in a calendar.

The E-R paradigm was so successful that many of its features were embedded in what can be considered the next big step with structured representation formalism: UML, the Unified Modeling Language (Jacobson et al., 1999). As for its predecessors, UML has been initially designed to address the modeling of software and complex computer systems. Its standard comprises many different structure diagrams, such as class or component diagrams (see the right part of Fig. 2), that can be used to model the structural entities belonging to the system, and are very similar, both in spirit and in appearance, to the formalism used with E-R (De Lucia et al., 2010, Neal, 2000). However, this new model aims to cover a broader range of information as it also includes behavior diagrams (such as activities and use cases) to extensively to describe the functionality of a system, and interaction schemas (such as communication or sequence diagrams) to capture the flow of control and data among the things in the system being modeled (see the left part of Fig. 2). Especially these latter parts of the UML standard have shown to be suitable to be adapted to model far more information flavor than those directly related to software development tasks. For instance, activity diagrams can be used directly to model a business processes, or uses cases can be adopted to represent customer relations. For these reasons, many specialized applications have been proposed in the literature since the first inception of the standard. SysML (Weilkiens, 2008) extends the base standard to handle general-purpose modeling of systems engineering applications. In (Chabrol and Sarramia, 2000) the authors introduce an object-oriented modeling methodology for urban traffic systems based on the UTS modeling approach, also using UML. A representation formalism specific to the domain of earthquakes is presented in (Babaie and Babaei, 2005). The relation between business processes and data warehouse information is handled through an extension of the activity diagram in (Ste- fanov et al., 2005).

Many other formalizations however, mainly in the field of business mod- eling, adopt custom schemas in order to exploit a higher degree of flexibility. The ORE (Organizational Requirement Engineering) (Flynn and Jazi, 1994), is a diagrammatic method designed to build a representation that includes both organizational and information system modeling. The Business Reference Model (Fettke and Loos, 2007), can be used to create a graphical representation of the functional and organizational aspects of the core business of an enterprise, service organization or government agency. Business processes are described together with business operations with the BOPM (Business Operations Management environment) framework, which deal with multiple dimensions of artifacts across the extent of a company, such as people, information, processes, and systems (Lin et al., 2007).

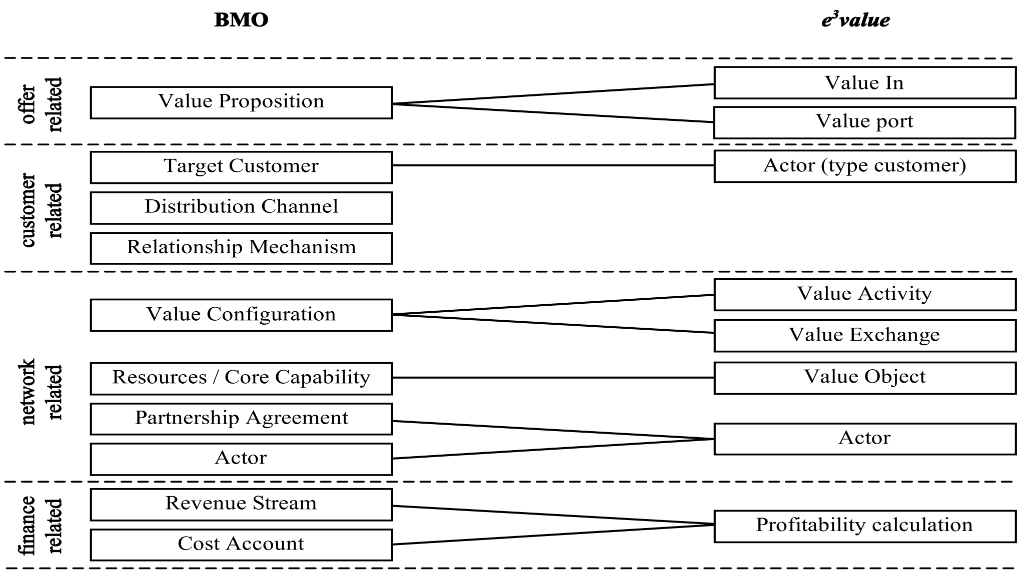

Figure 3 - The comparison of the ontology concepts offered by the BMO model and by the e^{3}value

The particular business object we are most interested with is however the Business Model (Burkhart et al., 2011, Zott et al., 2010). That is the description of the rationale of how an organization creates, delivers, and captures value. In other words, a business model portraits in detail from an organizational and economic point of view the core products or services offered by an entity, the resources that are needed to produce such offering, the customer segments that it addressed and distribution channels that are to be adopted to satisfy them. Specific business models have been extensively studied as single taxonomies. For instance in (Bambury, 1998) a taxonomy specially crafted for companies that primarily perform on-line commerce is presented. Other taxonomies related to the electronic markets include (Dai and Kauffman, 2002) and (Timmers, 1998).

Subsequently, more general meta-models have been proposed as tools for modeling more than one specific business model. Most notably, the Resource- Event-Agent (REA) ontology (Geerts and McCarthy, 1999), the e3value ontology (Akkermans et al., 2004), based on UML and RDF, and the Business Model Ontology (Osterwalder et al., 2004), based on OWL (Dean and Schreiber, 2004). The latter two formalisms have been compared in (Gordijn et al., 2005) and they have shown to aim at largerly parallel purposes. However the BMO ontology is more developed on some aspects, such as distribution channels and customer related areas, while the e^3value ontology exhibits a better modeling of value exchanges (see Fig. 3).

The conclusion of this latter review paper is that there would be an advantage in merging different ontologies to build a common formalism able to capture a more complex set of information. To this end, a possible merge of the two models is proposed. Of course, the main benefit of a bigger model is that it can accommodate different flavors of business models, overcoming the limitation of industry-specific ontologies and allowing for an analysis of the data that encompasses similar models from non-homogeneous markets. However, this comes with the drawback of more complex definitions, with the danger of being unable of representing at the best detail the many specific aspects that can characterize different vertical segments.

The approach that we are suggesting with this paper points towards the opposite direction. In fact, we are seeking for a general meta-model formalism that sets the rules to define different business model structures which can be tailored with features specific to different market segments and still being quite comparable, by virtue of common general rules and sub-ontologies.

The remaining of this paper is organized as following: section 2 will introduce the general motivation and the rules that we adopted to define our business meta model language, section 3 will define the business meta model language itself, section 4 will present an example of a business model instance that loosely resemble the Business Model Canvas (Osterwalder and Pigneur, 2010), section 5 will suggest some applications of the formalism, especially with respect to mining an pattern recognition activities.

2. Toward a General Formal Business Model Representation

As stated in the previous section, we are aiming to create a general meta model language that can be used to define several business model formalisms akin to the many usage needs. Such requisites are usually related to different market segments, however, different formalisms could be necessary to capture different perspectives of the same company, for instance by putting more focus on the customer relations or in the value production. To this end, the main goal of the meta language that we are proposing is to be general enough to provide great flexibility. At the same time, it should fix some general concepts that are common to all the types of business models, allowing to compare business models of different type.

The design of the meta language that we are proposing, that we called the Business Meta Model Language (and that will be described in detail in Sec. 3), has been guided by three main principles.

2.1. Business Models are graphs

We think that, generally speaking, there are two main concepts that contribute to the definition of a business model.

The first one is embodied by the involved entities. That is the physical or intellectual facts that provides for the actuation of the business plan. We call such entities Building Blocks. Examples of building blocks are the partners that supply the raw materials, the products to be offered, the associated costs or the intended customer segments.

The second concept is that building blocks are related with each another. Such relations, that we call Enabling Relations can express the concept of causality, as with the cost associated to the acquisition of a resource, or just correlation such as the connection between a product and a customer segment.

From a topological point of view building blocks can be considered vertices and enabling relations edges of a graph. If an enabling relation connects mor than two building blocks it can be regarded as an hyperedge with the cardinality of the connected blocks. This kind of interpretation opens the way for the application of the vast graph matching literature as a tool for analyzing business model data.

2.2. Typing defines specific Business Meta Models

The Business Meta Model Language can be used to defne specific Business Meta Models. This happens by declaring a set of building blocks and the associated enabling relation. The very nature of each entity depends on the characteristics of the Business Meta Model that is to be produced, however, some common features are available for each Meta Model.
Specifically, attributes of both building blocks and enabling relations are typed through predefined native types, that are part of the meta language, or through hierarchical ontologies, that can be shared through different meta models (and thus used for comparisons). Further, an intrinsic feature of both building blocks and enabling relation is a weight value that expresses the importance of each entity with respect to all the entities of the same type in the business model. This weight can also be used to obtain more meaningful comparison between different business model.

2.3. Mapping enable comparison between different Meta Models

The adoption of common base ontologies, as well as common native types, enables the ability of creating mappings between different Business Meta Model. A mapping should be a simple injection between entities (and their attributes) that preserves the type of the entity (building block or enabling relation) and the cardinality of the involved enabling relation. Once a mapping is defined the business meta model induce by their restriction to the domain and codomain of the injection can be deemed as isomorphic from both a topological and semantical point of view, and thus all the analysis techniques adopted between business models belonging to the same meta model can be applied.

3. The Business Meta Model Approach

In this chapter we discuss the graphical syntax which we use to describe the business meta model and business model instance. As described in the previous chapter we want to be flexible enough to describe any business model canvas and then to instantiate a business model from any canvas. To do this we need two languages, one goal to formalize the business model and another to formalize the business model. After presenting the graphics syntax of both we will discuss the choice of XML to represent models.

3.1. Graphical Syntax

The graphical syntax is necessary to describe the structure, behavior, and the architecture of the model. In our case we define a meta model to describe the business and one other to describe the business model. The business meta model consists of three main objects: building block, enabling relation and ontology.

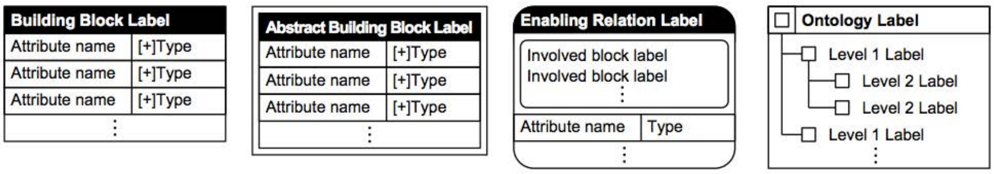

Figure 4 - The graphical syntax proposed for the Business Meta Model Language

The first object can be also abstract, this means that this may be extended to other building blocks. The object that extends an abstract object inherits all the features of the first object adding your own. A building block is formed by a label, which we have called *Building Block Label*, and by a series of attributes, which have a name, *Attribute name*, and a type, *Type*. We have defined the type of an attribute can be a *number*, *string*, *percentage* or an *ontology*. The symbol "+" next to the type of an attribute indicates whether this attribute can be instantiated only once, in case of absence of the symbol, or it can be instantiated multiple times, otherwise. The second block represents the enabling relation, this object is composed of a label, the Enabling Relation Label, a set of names of building block that corresponds to the blocks that connects the relation and finally a set of attributes. The latter set is structured in the same way in which it is present in the building block. Finally, the last object that can be defined is the ontology. This is defined as a tree structure, and may be as many levels are preferred. This object plays a very important role; in fact, two different business meta model using the same set of ontologies can be compared.

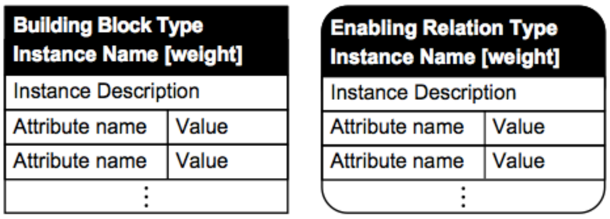

Figure 5 - The graphical syntax proposed for the Business Models

The second language is that we have defined to represent business models. Two objects, building blocks and enabling relations, represent a business model. The first is represented by the type of building block, an instance name and a weight. The weight of the building blocks of the same type must be equal to one hundred. Finally contains a set of attributes that correspond to those of the business meta model corresponding to that block type. Similarly defined is the object that corresponds to the enabling relation.

3.2. XML Serialization

We now describe the next step, which is to pass from the graphic representation to a structured representation. We decided to use XML to represent both business meta model that the business model. Even though the models defined by us are very flexible we decided that the information structure must be rigid in how it was structured. This decision is motivated by one of our starting points; all of the business model must be comparable. So we have defined an XML schema that determines the structure of the XML for both business meta model for the business model.

3.2.1. Business Meta Model Serialization

The objects to be represented in the business meta model are three building blocks, enabling relation and ontology. In Table 1 is represented the xml code relative to a building block of example. It is composed of an element "BuildingBlock" which contains three attributes, Label, Extends and Abstract and finally contains a set of elements, called "Attribute", containing three attributes: Label, Type and Multiple.

```
<BuildingBlock Label="BuildingBlockLabel" Extends="BuildingBlockLabel"
Abstract="false">
  <Attribute Label="AttributeLabel" Type="AttributeType" Multiple="false"/>
  ...
</BuildingBlock>
```

Table 1 - XML of Building Block

To verify the structure of this object is required the following XML schema:

```
<xs:complexType name="BuildingBlockType">
    <xs:sequence>
      <xs:element name="Extends" minOccurs="0" maxOccurs="1">
        <xs:complexType>
          <xs:attribute name="Label" type="xs:string" use="required"/>
        </xs:complexType>
      </xs:element>
      <xs:element ref="Attribute" minOccurs="0" maxOccurs="unbounded"/>
    </xs:sequence>
    <xs:attribute name="Label" type="xs:string" use="required"/>
    <xs:attribute name="Abstract" type="xs:boolean" use="required"/>
  </xs:complexType>
```

Table 2 - XML schema of Building Block

The xml code that describes the enabling relation is shown in Table 3. This consists of a tag "EnablingRelation" that contains a single attribute and two subelements Label: Involves and Attribute. The first element contains the attribute label, while the second contains three: Label, Type and Multiple. Both subelements are repeatable.

```
<EnablingRelation Label="EnablingRelationLabel">
  <Involves Label=”BuildingBlockLabel”>
  ...
  <Attribute Label="AttributeLabel" Type="AttributeType" Multiple="false"/>
  ...
</EnablingRelation>
```

Table 3 - XML of Enabling Relation

To verify the structure of this object is required the following XML schema:

```
<xs:complexType name="EnablingRelationType">
    <xs:sequence>
      <xs:element name="Involves" minOccurs="2" maxOccurs="unbounded">
        <xs:complexType>
          <xs:attribute name="Label" type="xs:string" use="required"/>
        </xs:complexType>
      </xs:element>
      <xs:element ref="Attribute" minOccurs="0" maxOccurs="unbounded">
      </xs:element>
    </xs:sequence>
    <xs:attribute name="Label" type="xs:string" use="required"/>
  </xs:complexType>
```

Table 4 - XML schema of Enabling Relation

The last object to represent the ontology is shown in Table 5. is formed by a tag "Ontology" which has an attribute Label and a set of subelements "OntologyItem". These are described in the same way ontology and are repeatable.

```
<Ontology Label="OntologyLabel">
  <OntologyItem Label="OntologyItemLabel”>
    <OntologyItem Label="OntologyItemLabel"/>
    <OntologyItem Label="OntologyItemLabel"/>
    <OntologyItem Label="OntologyItemLabel"/>
    ...
  </OntologyItem>
  ...
</Ontology>
```

Table 5 - XML of Ontology

To verify the structure of this object is required the following XML schema:

```
<xs:complexType name="OntologyType">
    <xs:sequence minOccurs="0" maxOccurs="unbounded">
      <xs:element name="OntologyItem" type="OntologyType"/>
    </xs:sequence>
    <xs:attribute name="Label" type="xs:string" use="required"/>
  </xs:complexType>
```

Table 6 - XML schema of Ontology

```
<xs:element name="Attribute">
    <xs:complexType>
      <xs:attribute name="AttributeLabel" type="xs:string" use="required"/>
      <xs:attribute name="AttributeType" type="xs:string" use="required"/>
      <xs:attribute name="Multiple" type="xs:boolean" use="required"/>
    </xs:complexType>
  </xs:element>
```

Table 7 – XML schema of Attribute

For completeness of information we show how the elements "Attribute" are checked with XML schemas in table 7.

3.2.2. Business Model Instance Serialization

In the framework that describes the business model instance we need two objects: building blocks and enabling relations. To avoid confusion with the previous model we call the first building block instance and the second enabling relation instance. The first object therefore has a tag "BuiildingBlockInstance" with attributes Type, which is a label of the previous model, Label and Weight. This element also can have a subset of items "Attribute" that are specified with a Label and a Value.

```
<BuildingBlockInstance Type="BuildingBlockLabel"
Label="buildingBlockInstanceLabel" Weight="Weight">
  <Attribute Label="AttributeLabel" Value="AttributeValue"/>
</BuildingBlockInstance>
```

Table 8 - XML of Building Block Instance

To verify the structure of this object is required the following XML schema:

```
<xs:complexType name="BuildingBlockInstanceType">
    <xs:sequence>
       <xs:element name="Attribute">
          <xs:complexType>
             <xs:attribute name="Label" type="xs:string" use="required"/>
             <xs:attribute name="Value" type="xs:int" use="required"/>
          </xs:complexType>
       </xs:element>
    </xs:sequence>
    <xs:attribute name="Type" type="xs:string" use="required"/>
    <xs:attribute name="Label" type="xs:string" use="required"/>
    <xs:attribute name="Weight" type="xs:int" use="required"/>
  </xs:complexType>
```

Table 9 - XML schema of Building Block Instance

The tag “EnablingRelationInstance” represents the second object, this has three attributes: type, which is a label of the previous model, Label and Weight. This object has a subset of items "Attribute" also has an additional feature compared to the previous object. This also has the subelement

"Involves" indicates that all the building blocks instance that the relationship connects.

```
<EnablingRelationInstance Type="EnablingRelationLabel"
Label="EnablingRelationInstanceLabel" Weight="Weight">
    <Involves="buildingBlockInstanceLabel”/>
   …
   <Attribute Label="AttributeLabel" Value="AttributeValue"/>
   …
</EnablingRelationInstance>
```

Table 10 - XML of Enabling Relation Instance

To verify the structure of this object is required the following XML schema:

```
<xs:complexType name="EnablingRelationInstanceType">
      <xs:sequence>
         <xs:element name="Attribute">
            <xs:complexType>
               <xs:attribute name="Label" type="xs:string" use="required"/>
               <xs:attribute name="Value" type="xs:int" use="required"/>
            </xs:complexType>
         </xs:element>
      </xs:sequence>
      <xs:attribute name="Type" type="xs:string" use="required"/>
      <xs:attribute name="Label" type="xs:string" use="required"/>
      <xs:attribute name="Weight" type="xs:int" use="required"/>
   </xs:complexType>
```

Table 11 - XML schema of Enabling Relation Instance

To understand how the two are connected to each other we need to describe the parent tag of each of the two. The first model is described by the root element “BusinessMetaModel” that has the attribute “Label” to differentiate it from others. The second has a root element named "BusinessModelInstance" that has an attribute "Label" and an attribute "Template", the latter is used to specify the "Label" of the business meta model referred to.

4. A Simple Meta Model Application

4.1. A (more) Formal Representation of the Osterwalder's Canvas

In this section we will define a possible representation of a business model template that contains all the possible objects that can be defined according to our XML formalization. The initial idea was to define the template canvas proposed by Osterwalder, but this is not designed for a structured representation. The model proposed by Osterwalder is based on four main blocks that represent all the content that a business model should contain. The first is the block of *product innovation*, which describes the value proposition of a company. The second block is the *customer relationship*, which describes how a company gets in touch with its customers and what kind of relationship established with them. The third block is the *infrastructure management* that describes what partners, resources and activities are necessary to start up the first two blocks. The last block is the *financial aspects* that describe how a company makes money through the first three blocks. Then proposes a further subdivision of these blocks up to the extended representation shown in [*Figure 6*].

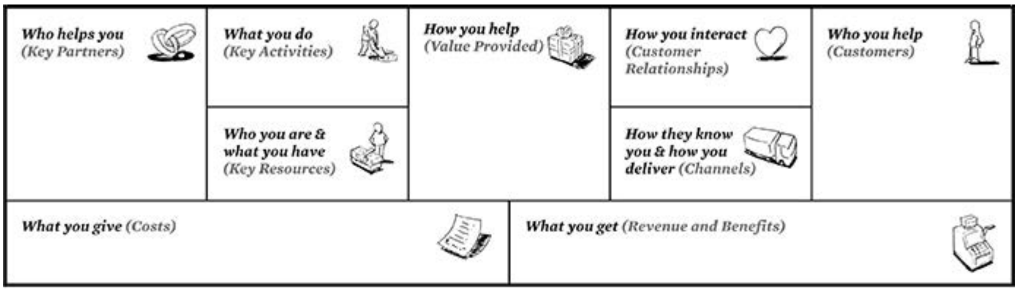

Figure 6 - The Business Model Canvas graphical sketchboard suggested by Osterwalder

The first block, *product innovation*, it contains only the block value proposition, which represents an overall view of the products and services that are value for a specific customer segment. Moreover, it proposes a subdivision of any *value proposition* in parts, called *offering*, which describe the different aspects of each of them. The second main block, the *customer relationship*, including all aspects of customers and can be divided into three blocks: *Target Customer*, *Distribution Channel* and *Customer Relationship*. The first block defines what type of customer the company has decided to propose its value proposition. Good customer segmentation allows the company to allocate the correct investments to its clients so that they are more attracted to their value proposition. The most important distinction exists between the target customer business and individual customer. Other criteria of subdivision proposed, refer to geographical or socio-demographic nature. The second block concerns the channels that connect the company's value proposition and the corresponding target customer. A channel can be direct, for example, using its website, or indirectly through intermediaries, such as resellers, brokers or cybermediaries. Finally, the third block analyzes the nature of the interactions that a company manages its clients. The purpose of this block is to describe the mechanisms for improving the acquisition, retention and add-on selling. Infrastructure Management, the third main block, it can be divided into the following three groups: *key resource*, *key activity* and *key partner*. The first block, *key resource*, describes everything that the company owns and what it can do. In order to create value for the company needs resources that can be of three main types: tangible, intangible and human resources. The tangible resources include plant and equipment; intangible assets include those patents, copyrights, reputation, brands and trade secret. Finally, the human resources are the people that the company needs in order to create value using the tangible and intangible resources. The second block describes all the activities that are necessary for a company to create value for the customer. Key activity can be of three main types, which are value chain, value shop and value network. The activities of the value chain framework include inbound logistics, operations, outbound logistics, and sales and service marketing. The value shop instead differentiates into problem finding and acquisition, problem solving, choice execution and control and evaluation. The last, the value network, includes the activities of network promotion and contract management, service provisioning and infrastructure operations. The last block of the infrastructure management is Key Partner. This block describes the partnership, a voluntarily initiated cooperative agreement formed between two or more independent companies in order to carry out a project or a specific activity jointly by coordinating the necessary Key Resource and Key Activity. The specifications of this block relate to the terms and conditions of a particular partnership, which can be reasoning, degree of integration and degree of dependency. A representation of this kind is suitable for human during a discussion between representatives of business and management experts but is not suitable for computer processing. The limits of the structure just described consist in too much generalization of the data, i.e. provide too many attributes discursive useless in comparison phase, and the implicit relations that link the blocks. In fact Osterwalder, in its arrangement of the blocks, suggests that for each adjacent block there is a relationship but this is never defined. With our business meta model language, described in chapter three, we want to define a structure that is complete both in blocks both in their connections. Looking at the specifications proposed by Osterwalder we found that some blocks proposed by him can be seen as relations between blocks. The building blocks that we have defined are as follows:

- **Key Partner**: represents the partner of the company, is specified by the attributes *Business Relations* and *Business Sector*. The first describes whether the partner is a supplier or an associate, while the second describes the international classification of the type of industry partners.
- **Cost Source**: represents the costs that take the company to bring its value to customers. Its attributes are *Yearly Amount* and *Currency*. The first is numerical and indicates the annual cost of a good or service, while the second indicates the type of currency that is associated with the previous numerical value.
- **Key Resource**: represents the key resources of the company, we decided to define abstract and without attributes because the building blocks that extend not share the same attributes.
- **Physical Resource**: is the first building block that extends Key Resource. It represents all of the company's premises from construction to machines up to information systems.

- **Intellectual resource**: is the second building block that extends Key Resource. Represents all the intellectuals' ownership of the company.
- **Value Proposition**: represents the value that the company offers to its customers. It has two attributes: type of product and Value. The first differentiate between whether the proposed value is a product or service while the second indicates the level of innovation of the product or service.
- **Customer Segment**: represents the company's customers. has two attributes, customer type and geographic area. The first indicates the type of customer segment in which the company wants to offer its value, while the second is a multiple attribute that indicates where the specific type of customer.
- **Revenue Source**: is the revenue of the company that is where the company benefits. It has the same attributes as the block Resource Cost

The enabling relation, which we have defined between the building blocks described above, are the following:

- **Resource Cost** [*Cost Source – Key Resource*]: is the relation that associates each key resource, both physical and intellectual, to its cost.
- **Partner Brokered Resource** [*Key Partner – Physical Resource – Value Proposition*]: This report connects three blocks, in our language is a n-ary relation, describing the annual cost of a partner as a supplier of goods and services.
- **Production Resource** [*Value Proposition – Key Resource*]: Unlike the previous this report describes the activities performed by the company itself by combining the resourcesto the value for which they are used.
- **Distribution Channel** [*Value Proposition – Customer Segment – Revenue Source*]: Represents the distribution channel used by the company to reach a certain type of customer. Has an attribute that

describes the type of channel used, whether directly or indirectly. This is also a ternary relationship.

- **Customer Relationship** [*Value Proposition – Customer Segment*]: This report forms the agreements that have been made with a specific type of customer to promote / improve the sales of their value.

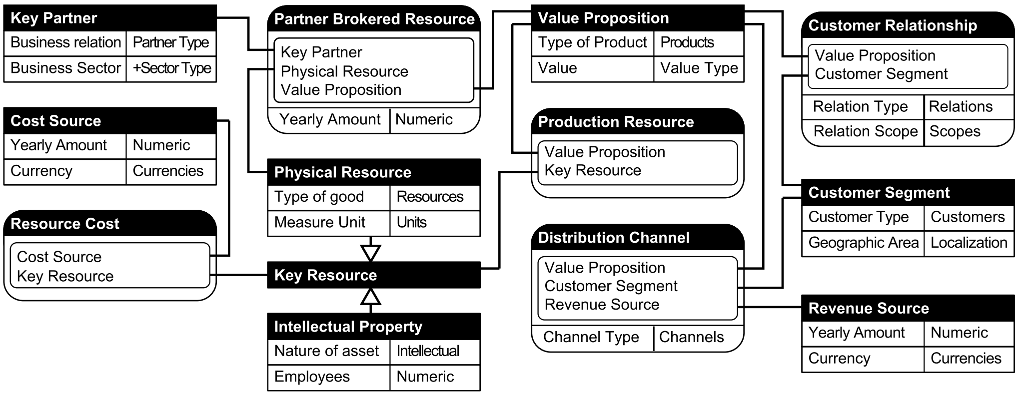

Figure 7 - The toy example of Business Meta Model loosely inspired to the Business Model Canvas

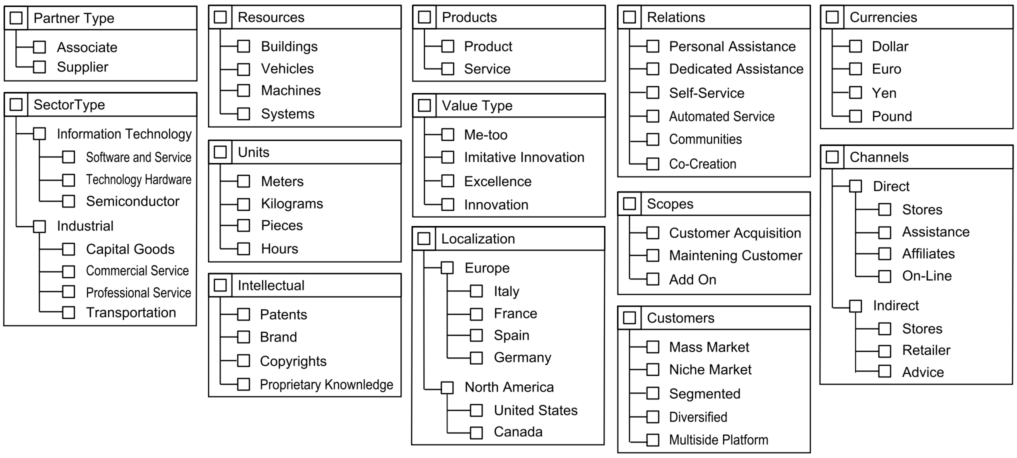

Figure 8 - A reduced set of ontologies to be used with our toy example

The attributes of which we spoke to describe the building blocks and the Enabling relation with the structure we have defined ontology. In Figure 7 and Figure 8 show the toy business meta model and the set of ontologies used. The business meta model we propose is a smart version of a true business goal model that describes every single detail of a company, but it is a good test to see if you can instantiate any type of business and whether it is possible to apply algorithms mining structures instantiated.

4.2. Some real-World Examples

In this chapter we show how our model can be adopted to describe the business model of any company and we see, through the application of graph matching algorithms, how we can find common subgraphs between multiple business model and how this can be useful in order to improve the company's business of certain companies. To do this we decided to consider a sample of forty companies that adopt different systems of production and distribution. To choose these companies, we decide to rely on three parameters: type of product, distribution type and customer segment. Each of these can be seen as an axis of a three-dimensional table, where each cell represents a company. We taken into consideration six types of product: books, fashion, footwear, music, electronics and travel. The type of distribution has been chosen to vary from the classic brick and mortar to e-commerce passing from the large retail chains. Finally, the customer segment was chosen to vary from products directed to a niche up to those addressed to the mass. Given the difficulty to represent a table on three dimensions on a sheet of paper, we will show a table for each selected product. The table will have the horizontal axis as the type of distribution and the vertical axis have the customer segment. Now we will show, in the order indicated above, the tables related to each type of product chosen, each with appropriate explanations of the choices made.

PRODUCT: BOOK

Niche	Lulu	Diamond Comic Distributors, Inc.	BookNiche
	The Book Depository		CAFOSCARINA
Mass	amazon	BARNES&NOBLE BOOKSELLERS	LIBRERIA LOVAT
	Web		*Brick and Mortar*

Table 12 - Product Book

We now proceed to see Table 12 for rows, then starting from the companies that are addressed to a niche market. *Lulu* is a company that promotes the publication of books of unknown authors through its e-commerce site in either printed and e-book. Their business is aimed at a niche market, as it only sells books of new authors and not classics or bestsellers. Its position in the first column is due to the fact that his only method of distribution to customers is its e-commerce site. In the next cell, then in the large retail chains, we find the Diamond Comic Distributors. This company is responsible for the distribution of American and Japanese comics in more than twenty member stores. The fact that both niche is due to the fact that deals with only comics. It addresses to a niche customer segment because it focuses exclusively on comic books, so one part of the product considered books. BookNiche display of last company is the first line in that only sells antique books and does so in a single store in Colorado in America. Moving on to the next line we find bookdepository, a company that offers only the sale of books on the web. It is located between niche and mass market for the fact that it uses a distribution channel but does not sell eBooks online. Ca Foscarina is a university bookstore that mainly deals with the sale of textbooks for students, but also handles commercial books. For this reason, is in the middle axis of the customer segment. The distribution channel is based on two shops near the University of Venice. Looking at now the last row we find in the cell that represents a company that uses an online distribution and caters to a customer type mass Amazon. In fact, this company sells both books and eBooks of all genres and does so only through its popular ecommerce site. Barnes and Nobles is an American company that sells all kinds of books at all its points of sale throughout the United States. It also started to sell eBooks through its own channel ecommerce. For these reasons it is placed in large deployments and has mass-market as customer segment. We find that the last cell library Lovat as Barnes and Nobles sells all kinds of books but has a distribution limited to a few stores in the province of Treviso.

Product: Fashion			
Niche	Prenatal	BODY GLOVE	SARTORIA BONAPARTE
	Stradivarius	NIKE	BREMA SPORT ESPERIENZA TECNICA
Mass	zalando	UNITED COLORS OF BENETTON.	BUSNATTO MAGLIERIA
	Web		*Brick and Mortar*

Table 13 - Product Fashion

The first company that appears in the table 13 is Prenatal, this produces and sells clothing for children and pregnant women, then in fashion is in a niche sector. For the distribution uses both its ecommerce site both a large chain of shops. The next one is Body Glove, a company that is specialize in clothing for watersport and is considered an excellence in this field. Distributes with direct mode on its website and in indirect mode by all affiliates. The last of the first line is Sartoria Bonaparte, a manufacturer of custom-made clothes for men or women. The distribution channel is limited to the store located in the province of Venice. Continuing to the next line we find Stradivarius, a company that produces clothes for ladies only. For the distribution uses its website ecommerce and owns a chain of stores throughout Europe. Soon after we find Nike, a company that produces sportswear for all major sports. The distribution is in indirect mode by shops and retailers of sports clothing. Brema Sport is a store that deals exclusively with the sale of sports equipment of the most famous brands. Being a shop the distribution channel is obviously brick and mortar. Turning to companies that have the product fashion on a mass market find first Zalando. This company is an online reseller of all kinds of clothes and shoes of all brands. For these reasons is located in the position that intersects mass and web. In the next cell there is a company that produces Benetton clothing under its own brand and uses a distribution channel of indirect affiliates with its brand.

The last cell contains Maglieria Busatto, a company in the fashion world that produces and sells in the province of Venice. Distribution is via a store and is aimed at a mass market as customer segment.

PRODUCT: FOOTWEAR			
Niche	Finish Line	Foot Locker	Ortopedia & Sanitaria MARRA
	heels.com	DECATHLON	
Mass	zalando	SCARPE & SCARPE	wL newLions
	Web		*Brick and Mortar*

Table 14 - Product Footwear

FinishLine is a company that is positioned in a niche market, in the table 14, because sells only sports shoes of all the most famous brands, as a distribution channel uses only his ecommerce site. Similarly Footlocker is a niche market because it sells the same products as the previous company. The difference is in the distribution channel, this fact relies for its stores located throughout the world. Display of last to be in a niche sector is Marra, for the fact that is responsible for selling sanitary footwear tailored. Its distribution is done solely in a single store, for this is in the last column. Another company that uses the web as the sole distribution channel and is located between a customer and a niche segment of mass heels. This takes care only of women's shoes of all kinds. In the same row are Decathlon, a manufacturer of sports footwear and sells both its own and those of other brands to give customers more choice. It is located in the middle as it uses as a distribution channel its stores located throughout Europe. Looking finally the last row, and then in the mass market, we find Zalando that is responsible for selling all kinds of shoes of all brands and does so only through its ecommerce site. Then there's Scarpe & Scarpe that as the previous company deals with the sale of all kinds of shoes of various brands. The difference is in the distribution channel, Scarpe & Scarpe relies primarily in its stores.

Product: Music			
Niche	MSOL.biz www.musicshoponline.com		Metropolis dischi
	cubomusica		Autogrill
Mass	iTunes	Feltrinelli	CDLand
	Web		*Brick and Mortar*

Table 15 - Product Music

In the music world (table 15) there are two companies that occupy a niche position in the customer segment. The first is musicshoponline, which aims to sell only cd of an independent record label and does so only through its ecommerce site. The second is a store that Metropolis is located in Milan, which is mainly engaged in the sale of only vinyl records. Its position is justified by the fact that distributes exclusively through his shop. On the next line there are two other companies, both mainly sell only cd more commercial but they do differently. The first Cubomusica, sells individual tracks through its online channel ecommerce, instead Autogrill sells albums through its stores located throughout the Italian motorway network. Finally in the mass market we find iTunes with regard to the web. In fact, the iTunes Store allows anyone to buy any type of music, album and single track that the online store. The company is next Feltrinelli, is responsible for sales of music albums, and does so using as a distribution channel its stores located throughout the Italian territory. Finally, the last cell is CDLand a store that sells any type of musical album with his single store.

PRODUCT: ELECTRONICS

Niche	Mr.pc	PC CITY Computer Superstore	INTERPUTER hardware & software
Mass	e-Key SHOP ON LINE	MediaWorld	MARCO POLO
	Web		*Brick and Mortar*

Table 16 - Product Electronics

Let us now consider the products in the electronics industry. In the first line, where we find companies that link to a niche market, we find misterpcpoint company that only sells PC components, so a niche electronics, and does so through its online site. In the next cell we find PC City, a chain that is responsible for selling only computers or components of it. The distribution channel is assigned to its stores throughout Italy. In last place are Interputer, a company that uses as a distribution channel his shop and is responsible only for the sale of computer and its components. In the next line we find only Apple positioned in the center, which sells computers, music players and phones of its production. The distribution channel that is used primarily entrusted to the store spread all over the world. In the last line we find e-key, a company that sells electronic components of all types and makes him only through its ecommerce site. In the last line there are three companies that are involved in selling all kinds of electronic components, from large to small appliances. The first is ekey that relies on a channel of distribution exclusively online. The second is that Mediaworld as a distribution channel uses its shops present in the most important European countries. The latest is Marcopolo that as MediaWorld uses its stores for distribution to the customer, but is limited in a local area of Italy.

Product: Travel

Niche	VIAGGI AVVENTURE NEL MONDO www.viaggiavventurenelmondo.it		Viaggi à la carte MANUREVA TOUR OPERATOR
	BOOKING.COM online hotel reservations		
Mass	Bluvacanze	alpitour	TOP TOURS
	Web		*Brick and Mortar*

Table 17 - Product Travel

The latest product that we have considered are the voyages as can be seen in Table 17. In the first row there are two companies that offer trips for a niche sector clients. The first is viaggiavventuranelmondo offering only adventure travel and does so by relying completely on your Web site. The second is Manureva, a small shop that specializes in travel in French Polynesia and offers them to customers through a single store. In the second row we find only booking, the only company that offers booking of hotels and it does so through its channel ecommerce online. Finally, the last line we find those companies that point to a mass market. The first is Bluvacanze that enable the customer to choose any type of journey from its online site. Another company that offers the same service to customers and Alpitour but does so with a different distribution channel, through the shops of travel arrangement. To we find last Top Tours, a single travel agency based in Venice, which organizes all types of travel to its customers but does so in its unique selling point.

4.2.1. Examples of Business Model

The companies chosen in the previous chapter are been insert in our toy model of business model template. This analysis should be done together to the economists that know every detail of all the company descripted, but the scope of this chapter is to demonstrate that every type of company can be structured in our model, also the full business model of a company are to big to be represent in a thesis. Then the choice to utilize the toy template is adequate to better understand the behavior of all the company, also we limit to a small analysis of the mainly mechanism of every company. As descripted in the previous chapter we considerer seven building block and five enabling relation that we have taken as fundamental. The seven building blocks are the following:

- *Key Partner*;
- *Cost Source*;
- *Physical Resource* that specializes *Key Resource*;
- *Intellectual Resource* that specializes *Key Resource*;
- *Value Proposition*;
- *Customer Segment*;
- *Revenue Source*.

While the five enabling relations are the following:

- *Resource Cost*
- *Partner Brokered Resource*
- *Production Resource*
- *Distribution Channel*
- *Customer Relationship*

The schemas that link the relation of all the enabling relation with the building block are illustrated in Figure 1.

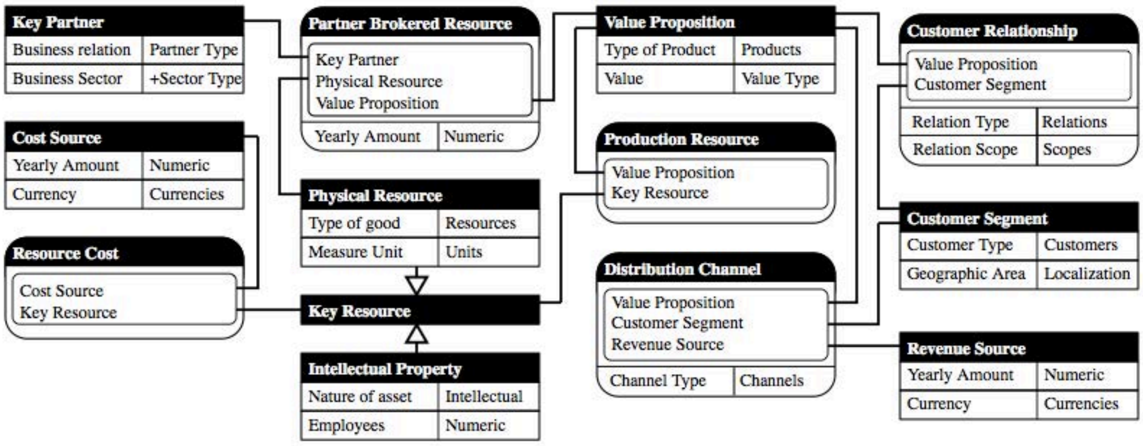

Figure 9 - The toy example of Business Meta Model loosely inspired to the Business Model Canvas

Now we explain the structure of twenty instance of business model to better understand as is possible formulate every type of business model with our model. For each table we have selected the most representative company for this scope. The companies that we have choose in the previous tables are the following:

- *Lulu (book);*
- *Diamond Comic Distributors (book);*
- *Amazon (book and electronics);*
- *Barnes and Nobles (book);*
- *Body Glove (fashion);*
- *Brema Sport (fashion);*
- *Stradivarius (fashion);*
- *Benetton (fashion);*
- *Heels (footwear);*
- *Decathlon (fashion and footwear);*
- *New Lion (footwear);*
- *Metropolis Dischi (music);*
- *iTunes (music);*
- *PC City (electronics);*
- *Mediaworld (electronics).*

For each of these we illustrate the business model instance and explain all the building blocks and the enabling relations that we have instanced.

4.2.1.1. Case lulu.com

Lulu.com is a demonstration of how the Web has changed the book publishing industry over the past years. The authors, in the old model, they had to send their work to publishing houses and hope that this be approved and then be printed. This was because the publishers were only interested in books that could be printed in massive quantities to be sold to a wide audience. Instead lulu has developed a new business model in publishing that enables anyone to publish. The idea is based in helping amateur and niche authors in bringing their work in their online shop. This system works because the books are printed only when they reach a predetermined number of orders. The failure of a particular book to sell is irrelevant to lulu, because failure does not involve any cost. Now let's see how this business model we presented in our data structure. We have instantiated the following building blocks:

Key Partner		
	---	---
Cost Source		
	Platform development	Represents the cost of development and maintenance of the ecommerce site
Physical Resource		
	Platform	Is the marketplace of lulu.com
	Print on Demand infrastructure	Is the infrastructure used by lulu to print the books that reach a sufficient number of sales
Intellectual Resource		
	---	---
Value Proposition		
	Self Publishing Services	Is the service that allows any author to publish their books in the marketplace
	Marketplace for niche content	Is the ecommerce site that allows users to buy all the books available
Customer Segment		
	Niche Audiences	Represents all the public interested in books on the site
	Niche Authors	Representing authors for lulu are seen as customers, not as key partners
Revenue Source		
	Sales Commission	Revenue from the sale of books
	Publishing Service Fees	Revenues from the publishing service

Table 18 - Building Block [lulu]

The enabling relation instantiated to the corresponding building block are the following:

Resource Cost		
	Link platform	This relationship linking the ecommerce platform to its cost
Partner Brokered Resource		
	---	---
Production Resource		
	Platform Development	Connects the physical resource platform to the marketplace to indicate that this is necessary for the functioning of the value proposition
	Logistics	This relationship linking the publishing service to the physical resource of printing
Distribution Channel		
	lulu.com	Represents the distribution channel used by lulu to get to its customers. Is also connected to the revenue generated by using this channel
	Authors publishing	Is the channel used to allow authors to promote their work through the publishing service
Customer Relationship		
	Communities of interest	Is the service that enables lulu to enhance their loyalty through the discussion of common issues
	Online Profile	Is the way lulu maintains a close relationship with its authors.

Table 19 - Enabling Relation [lulu]

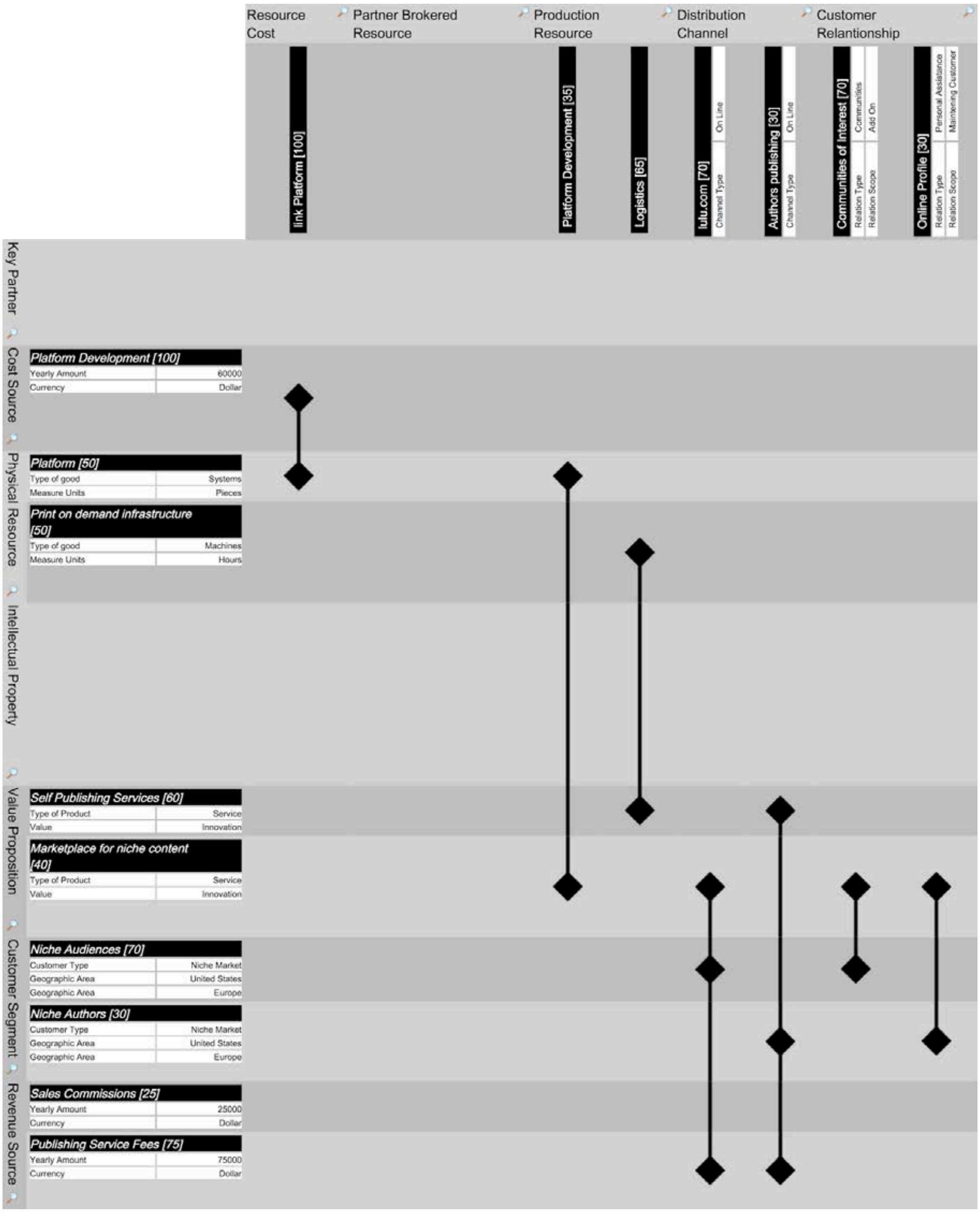

Figure 10 - Business Model Instance of lulu

4.2.1.2. *Case Diamond Comic Distributors*

Diamond Comic Distributors is a wholesaler of comic operates in North America. His business is limited to specialized libraries in this area, which therefore represent a niche market. Diamond Comic Distributors, through the comic publishers, fills its stores and then organizing shipments to their entire network of stores distributed between the United States and Canada. We have instantiated the following building blocks:

Key Partner		
	Comic Supplier	Represents all the publishers who provide the comic book to Diamond Comic Distributors
	Distribution Supplier	Is the service provider for the distribution of comics. Diamond Comic Distributors leaves in outsourcing this service
Cost Source		
	Comic Management	Corresponds to the cost of managing warehouses run by the company
Physical Resource		
	Distribution Network	Distribution, even if outsourced, is a physical resource used by the company
	Warehouses	Represents all the warehouses managed by Diamond Comic Distributors
Intellectual Resource		
	---	---
Value Proposition		
	Comic Distribution	Is the service offered by the wholesaler involved
Customer Segment		
	Niche Library	Represents the bookstores that are customers of Diamond Comic Distributors
Revenue Source		
	Comic Sales	Is the revenue that comes from selling to libraries

Table 20 - Building Block [Diamond Comic Distributors]

The enabling relation instantiated to the corresponding building block are the following:

Resource Cost		
	Link comic	Connects the stores to their cost of managing
Partner Brokered Resource		
	Distribution Management	Represents the outsourcing activities carried out by the company providing the service distribution. Then connect the service provider, network distribution and the corresponding service.
Production Resource		
	Logistics	Corresponds to the activity logistics of sorting to specific stores.
Distribution Channel		
	Retail Network	Is the distribution channel used by the company to reach its customers, which is made from the warehouses that distribute to bookstores customers.
Customer Relationship		
	---	---

Table 21 - Enabling Relation [Diamond Comic Distributors]

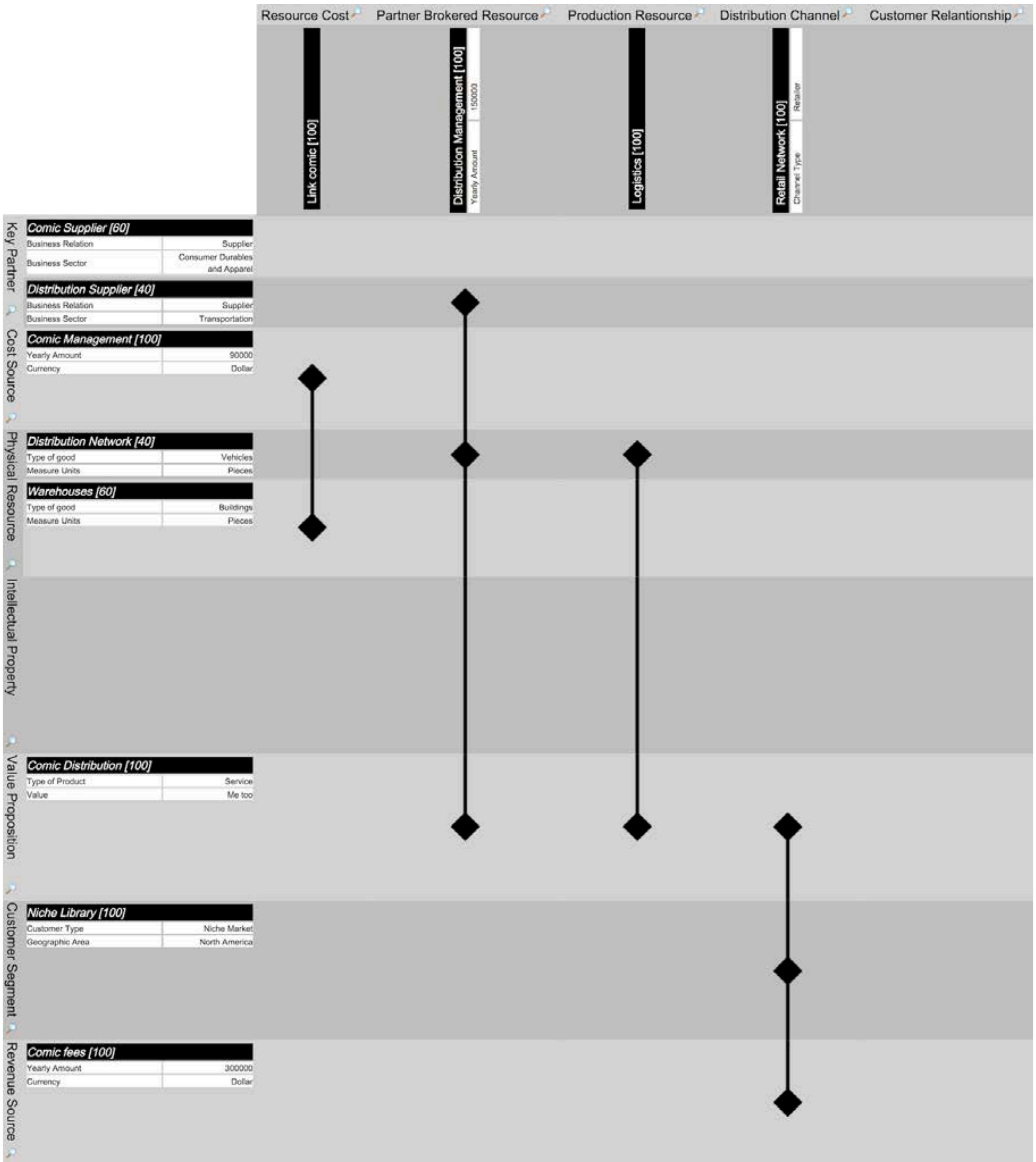

Figure 11 - Business Model Instance of Diamond Comic Distributors

4.2.1.3. Case Amazon

Amazon was one of the first companies to operate exclusively over the Web Its business consists of selling books and electronics products at lower prices than traditional retail chains. A component that has allowed us to achieve this success is the personalized suggestion engine that is based on user experience while browsing the website. This allows introducing products to the customer who may not have ever tried but which may be helpful. In addition, the company focuses on customer retention through strong relationship through premium account. We have instantiated the following building blocks:

Key Partner		
	Electronics Supplier	Represents the suppliers of the components of electronic
	Books Supplier	Representing book publishers supplying the holding
	Distribution Supplier	Represents the service providers for distribution to customers
Cost Source		
	Warehouse Cost	Is the cost of management of the warehouses owned by the company
	Platform Management	Represents the cost of development and maintenance of the marketplace
	Suggestion Cost	Is the cost of staff dedicated to the development of the engine of suggestion
Physical Resource		
	Platform	Is the marketplace used by Amazon
	Warehouse	Represents the warehouses used to sort the products by the company
	Distribution Network	Represents the way of distribution of Amazon
Intellectual Resource		
	Suggestion Staff	Is staff development and maintenance of the engine of suggestion
Value Proposition		
	Books a very low price	
	Electronics a very low price	
	Marketplace	
Customer Segment		
	Web Surfer	Are customers who may have access to the marketplace amazon
Revenue Source		
	Revenue of books	Represents the proceeds from the sale of books
	Revenue of electronics	Represents the proceeds from the sale of electronic components

Table 22 - Building Block [Amazon]

The enabling relation instantiated to the corresponding building block are the following:

Resource Cost		
	Warehouse management	Connects the costs of the warehouses to the warehouses themselves
	Platform management	Linking the cost of development and maintenance of the marketplace to the resource marketplace
	Suggestion link	Connects the costs of staff in charge to improve the motor of suggestion the corresponding intellectual resource
Partner Brokered Resource		
	Electronics to warehouse	Represents the mode of delivery of the electronic components, connecting the suppliers, to the warehouses until the product supply
	Book to warehouse	Is the mode of delivery of books, linking suppliers, and warehouses until the product supply
Production Resource		
	Suggestion improvement	Is the link between staff committed to improving of the engine of suggestion the marketplace itself
	Platform develop	Is the link between the developments of the Web platform with the service offered to customers
Distribution Channel		
	Book on amazon.com	Is the distribution channel between the company and customers interested in books
	Electronics on amazon.com	Is the distribution channel between the company and customers interested in electronics
Customer Relationship		
	Premium Account	Is the way in which the company tries to retain its customers

Table 23 - Enabling Relation [Amazon]

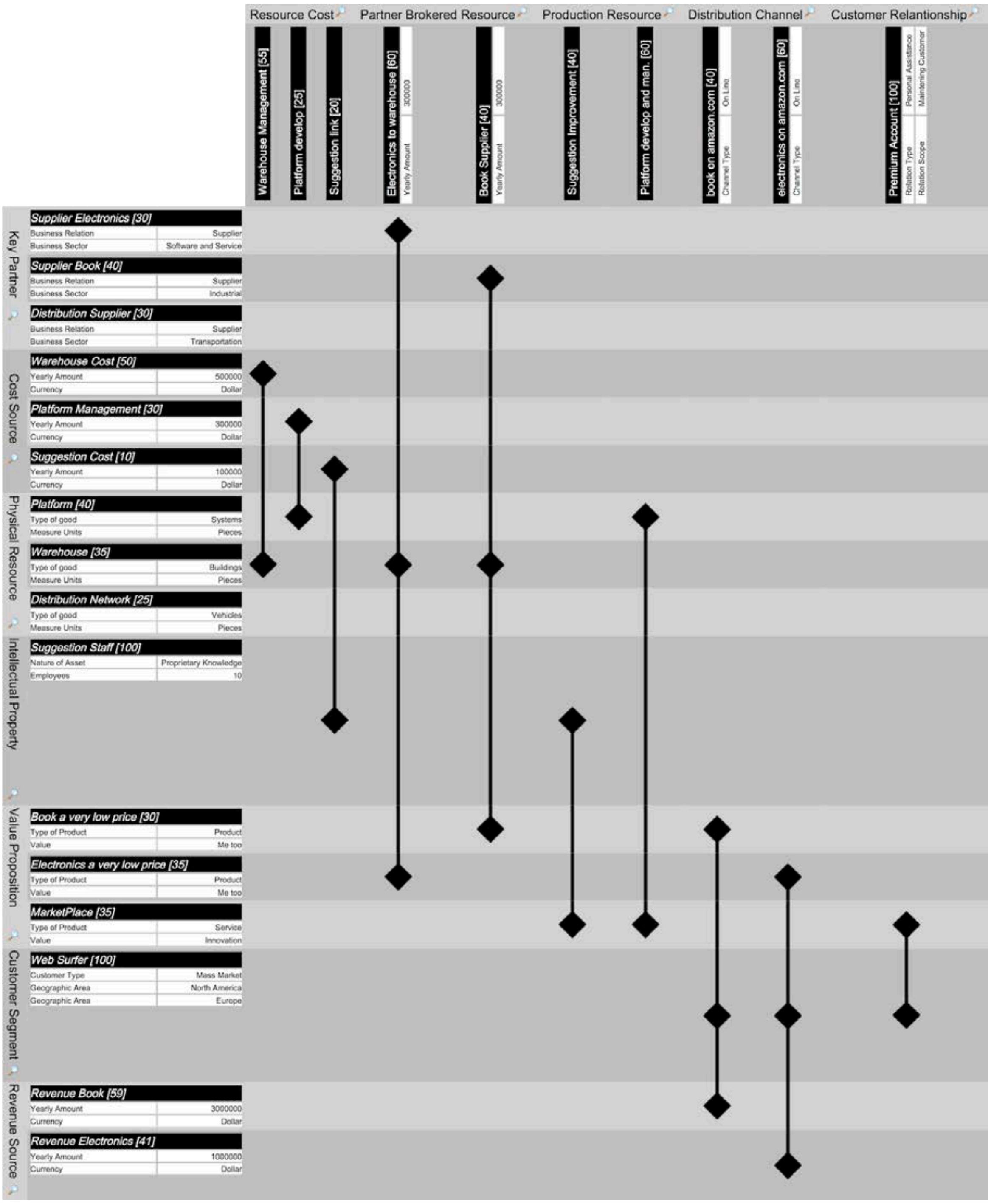

Figure 12 - Business Model Instance of Amazon

4.2.1.4. Case Barnes and Noble

Barnes and Noble is the largest retailer of books in the United States. His business is addressed to a mass market, using as a distribution channel more than seven hundred stores located throughout the Americas. The company is buying books from the largest publishing houses in the world, and manages the entire distribution to retail sales through its stores.

We have instantiated the following building blocks:

Key Partner		
	Books Supplier	Representing book publishers supplying the holding
Cost Source		
	Warehouse Cost	Is the cost of management of the warehouses owned by the company
	Cost Shop	Is the cost of management of the shops owned by the company
	Cost Distribution	Represents the cost of distribution of the product
Physical Resource		
	Retail Network	Represents the entire chain of stores owned by the company
	Warehouse	Represents the warehouses used to sort the products by the company
	Distribution Network	Represents the way of distribution of Barnes and Nobles
Intellectual Resource		
Value Proposition		
	Books	Is the product on which the company aims
Customer Segment		
	Book readers	Are the customers interested in the company's product
Revenue Source		
	Revenue of books	Represents the proceeds from the sale of books

Table 24 - Building Block [Barnes and Noble]

The enabling relation instantiated to the corresponding building block are the following:

Resource Cost		
	Warehouse management	Connects the costs of the warehouses to the warehouses themselves
	Store management	Connects the costs of the stores to the stores themselves
	Distribution management	Is the link between the cost of distribution and the network available
Partner Brokered Resource		
	Orders book	It is the business of providing books that connects publishers, stores and the company's product
Production Resource		
	---	---
Distribution Channel		
	Sell on shops	the company uses its stores as a distribution channel to the end customer
Customer Relationship		
	Loyalty card	Is the way in which the company tries to retain its customers

Table 25 - Enabling Relation [Barnes and Noble]

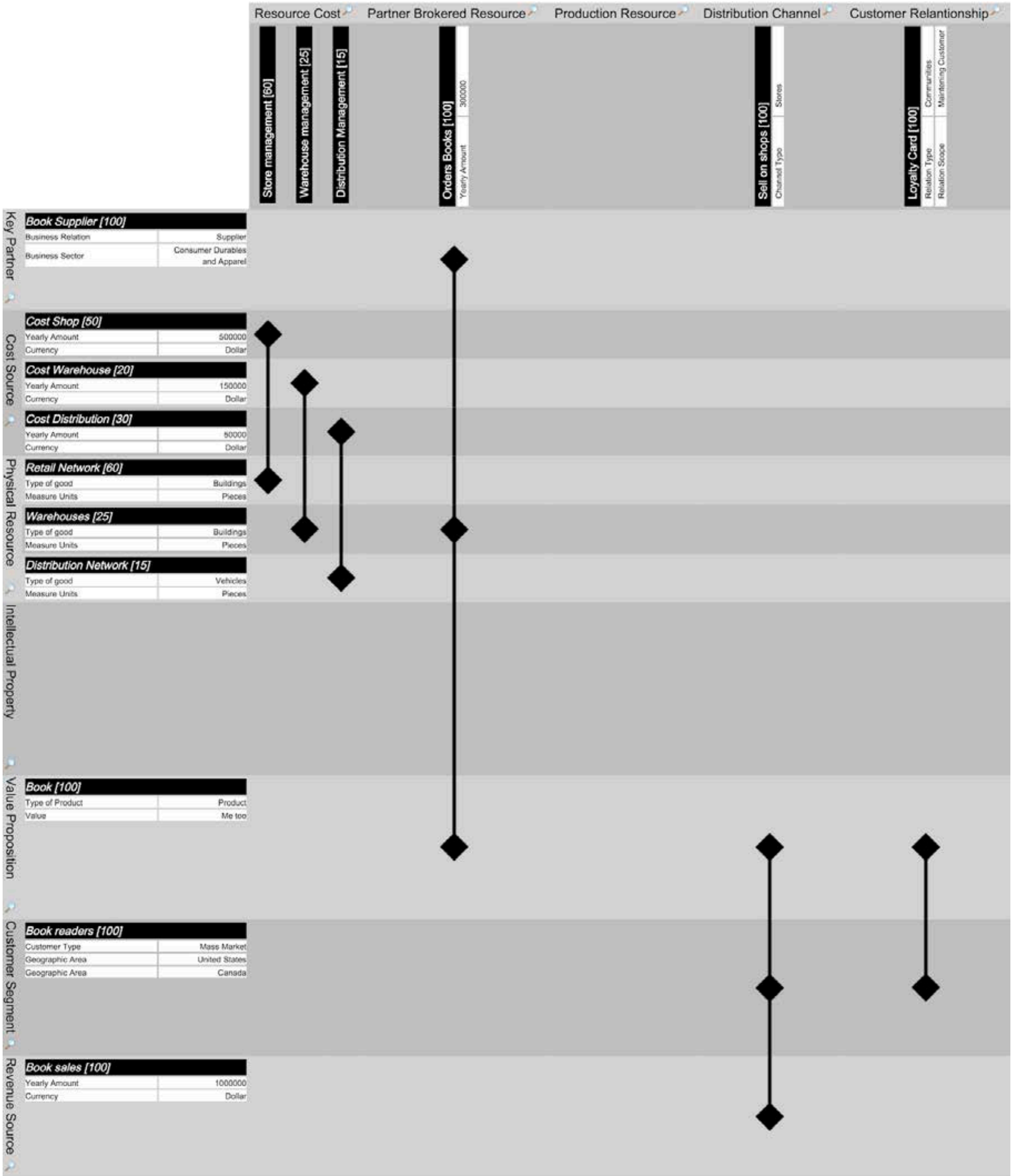

Figure 13 - Business Model Instance of Barner and Noble

4.2.1.5. Case Body Glove

Body Glove was one of the first companies to take care of the clothing market for water sports. This exclusive interest in this market leads the company to be targeted at a niche market. The company manufactures and distributes its products in specialty shops in the industry. In this case then there is also a staff, which is concerned with the design and characteristics of the new products.

We have instantiated the following building blocks:

Key Partner		
	Distribution Supplier	Represents the service providers for distribution to customers
Cost Source		
	Cost Warehouse	Is the cost of management of the warehouses owned by the company
	Cost Factory	Is the cost of management of the factory production
	Cost Research	Is the cost of staff that creates new models
Physical Resource		
	Factory	Represents the factories used by the company to create its products
	Warehouse	Represents the warehouses used to sort the products by the company
	Distribution Network	Represents the way of distribution of Body Glove
Intellectual Resource		
	Brand Body Glove	Is the brand of the company
	Research new product	Is the team that designs new products
Value Proposition		
	Wear WaterSport	Is the product that produces the company
Customer Segment		
	People interested WaterSport	Represents clients who are interested in the product provided by the company
Revenue Source		
	Revenue of WaterSport	Represents the proceeds from the sale of watersport's product

Table 26 - Building Block [Body Glove]

The enabling relation instantiated to the corresponding building block are the following:

Resource Cost		
	Warehouse management	Connects the costs of the warehouses to the warehouses themselves
	Factory management	Is the link between the cost of the factory and the factory itself
	Research management	Connects the costs of staff in charge to improve the motor of suggestion the corresponding intellectual resource
Partner Brokered Resource		
	Distribution Management	Represents the outsourcing activities carried out by the company providing the service distribution. Then connect the service provider, network distribution and the corresponding service.
Production Resource		
	Research	Represents the research staff with the product they are designing new models
	Make product	Is the activity that connects the production with the product in question
Distribution Channel		
	Distribution Retailer	Is the distribution channel used by the company to reach its customers, which is made from the warehouses that distribute their products
Customer Relationship		
	---	---

Table 27 - Enabling Relation [Body Glove]

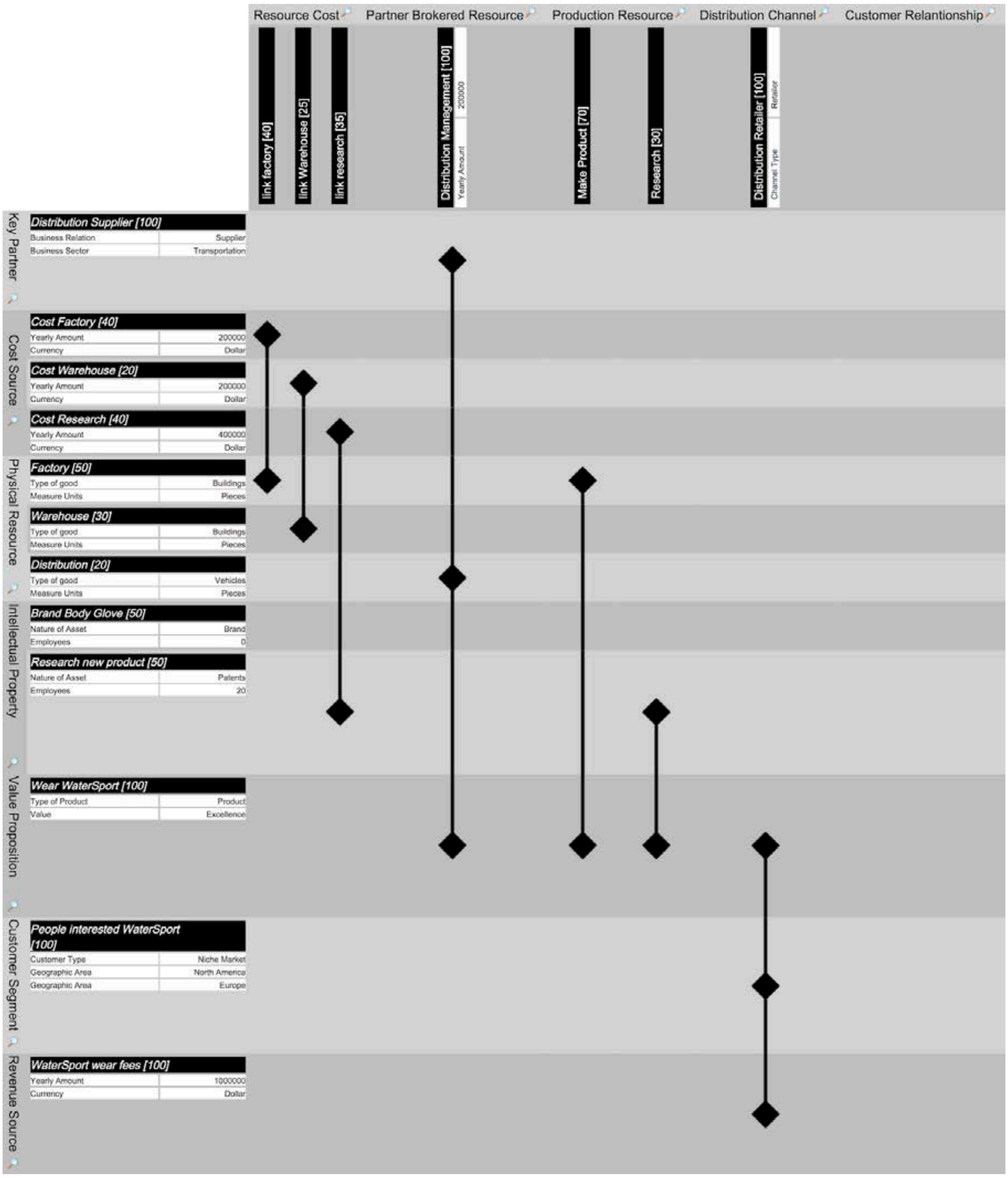

Figure 14 - Business Model Instance of Body Glove

4.2.1.6. Case Brema Sport

Brema Sport is a store that deals with sportswear and sports shoes. The products we sell are addressed only to the sports market than the clothing therefore making it in the middle between the niche and mass. The distribution channel is its unique shop located in the northeast Italian. The company is only a reseller, as does not produce any kind of clothing.

We have instantiated the following building blocks:

Key Partner		
	Sport wear supplier	Represents the suppliers of sports clothing
	Sport shoes supplier	Represents the suppliers of sports shoes
Cost Source		
	Cost warehouse	Is the cost of management of the warehouses owned by the company
	Cost store	Is the cost of management of the shops owned by the company
Physical Resource		
	Warehouse	Represents the warehouses used to sort the products by the company
	Stores	Represents the entire chain of stores owned by the company
Intellectual Resource		
	---	---
Value Proposition		
	Sport wear	Is the product on which the company aims
	Sport shoes	Is the product on which the company aims
Customer Segment		
	People interested sport	Are the customers interested in the company's product
Revenue Source		
	Wear sales	Represents the proceeds from the sale of wear
	Shoes sales	Represents the proceeds from the sale of shoes

Table 28 - Building Block [Brema Sport]

The enabling relation instantiated to the corresponding building block are the following:

Resource Cost		
	Warehouse management	Connects the costs of the warehouses to the warehouses themselves
	Store management	Connects the costs of the stores to the stores themselves
Partner Brokered Resource		
	Wear to warehouse	It is the business of providing wear that connects supplier, warehouse and the company product
	Shoes to warehouse	It is the business of providing shoes that connects supplier, warehouse and the company product
Production Resource		
	---	---
Distribution Channel		
	Distribution wear store	Company uses its stores as a distribution channel to the end customer
	Distribution shoes store	Company uses its stores as a distribution channel to the end customer
Customer Relationship		
	---	---

Table 29 - Enabling Relation [Brema Sport]

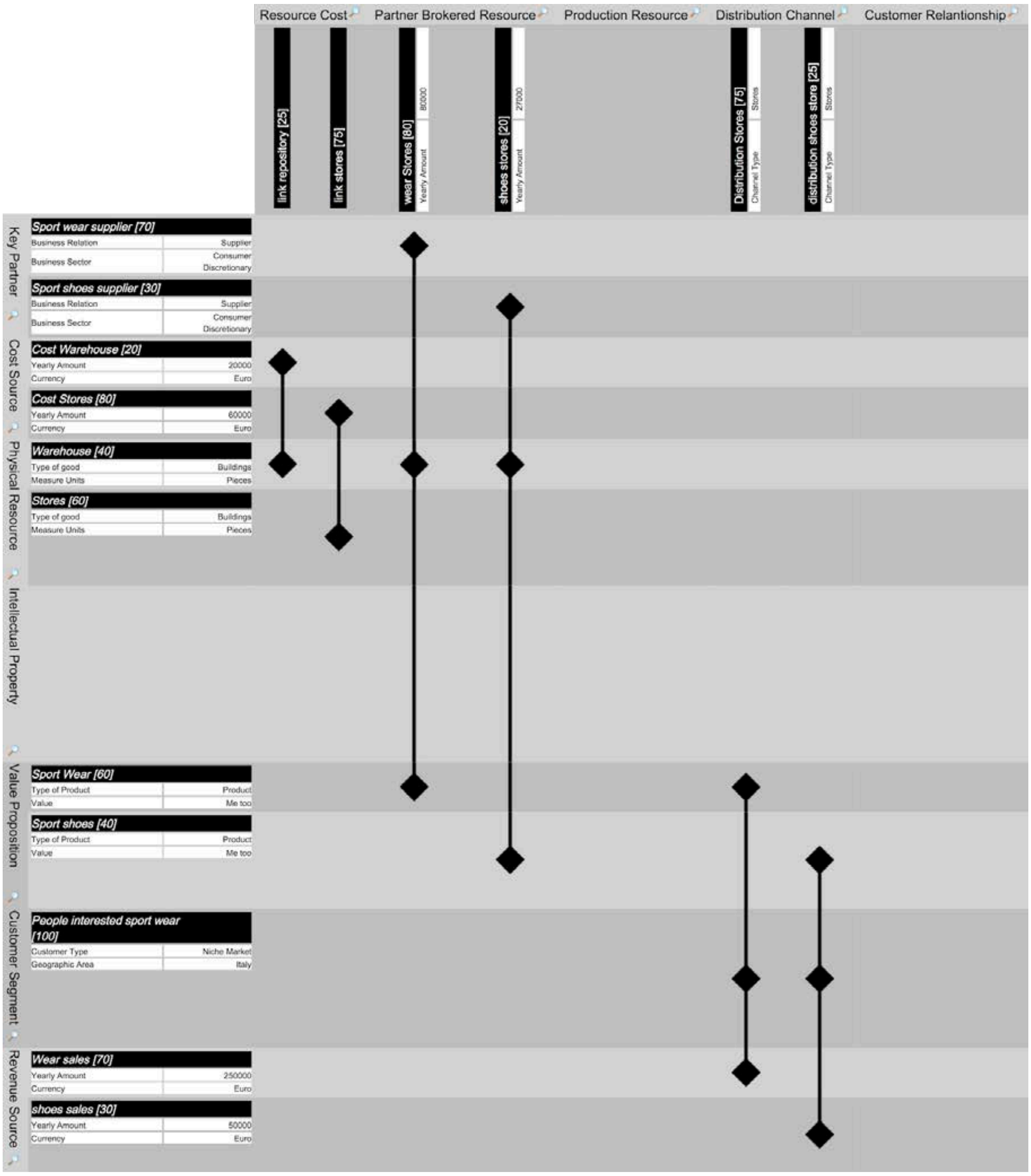

Figure 15 - Business Model Instance of Brema Sport

4.2.1.7. *Case Stradivarius*

Stradivarius is a company that produces and distributes clothing for young women only. Its distribution channel is divided between retail and ecommerce. In our example an external company that deals with both development and maintenance manages the ecommerce channel. Direct costs are those of the company's warehouses and staff that draw new clothes. Revenues are divided into two parts, those resulting from ecommerce and those arising from the shops.

We have instantiated the following building blocks:

Key Partner		
	Platform develop and management	Represents the cost of development and maintenance of the ecommerce site
	Wear supplier	Represents the suppliers of woman clothing
Cost Source		
	Cost warehouse	Is the cost of management of the warehouses owned by the company
	Cost research	Is the cost of staff that creates new models
Physical Resource		
	Factory	Represents the factories used by the company to create its products
	Warehouse	Represents the warehouses used to sort the products by the company
	e-commerce	Is the marketplace used by Stradivarius
Intellectual Resource		
	Brand Stradivarius	Is the brand of the company
	Research new product	Is the team that designs new products
Value Proposition		
	Wear woman	Is the product that produces the company
Customer Segment		
	Young woman	Represents clients who are interested in the product provided by the company
Revenue Source		
	Wear sales on shops	Represents the proceeds from the sale of product in the shops
	Wear sales on ecommerce	Represents the proceeds from the sale of product in the ecommerce

Table 30 - Building Block [Stradivarius]

The enabling relation instantiated to the corresponding building block are the following:

Resource Cost		
	Link warehouse	Connects the costs of the warehouses to the warehouses themselves
	Link research	Connects the costs of staff in charge to improve the motor of suggestion the corresponding intellectual resource
Partner Brokered Resource		
	Platform management	Represents the development and maintenance is the ecommerce site outsourced
	Factory management	Is the production of clothing outsourced
Production Resource		
	Research	Represents the research staff with the product they are designing new models
Distribution Channel		
	Distribution retailer	Is the distribution channel that connects the retail store to the customer
	Distribution ecommerce	Is the distribution channel that connects the ecommerce to the customer
Customer Relationship		
	---	---

Table 31 - Enabling Relation [Stradivarius]

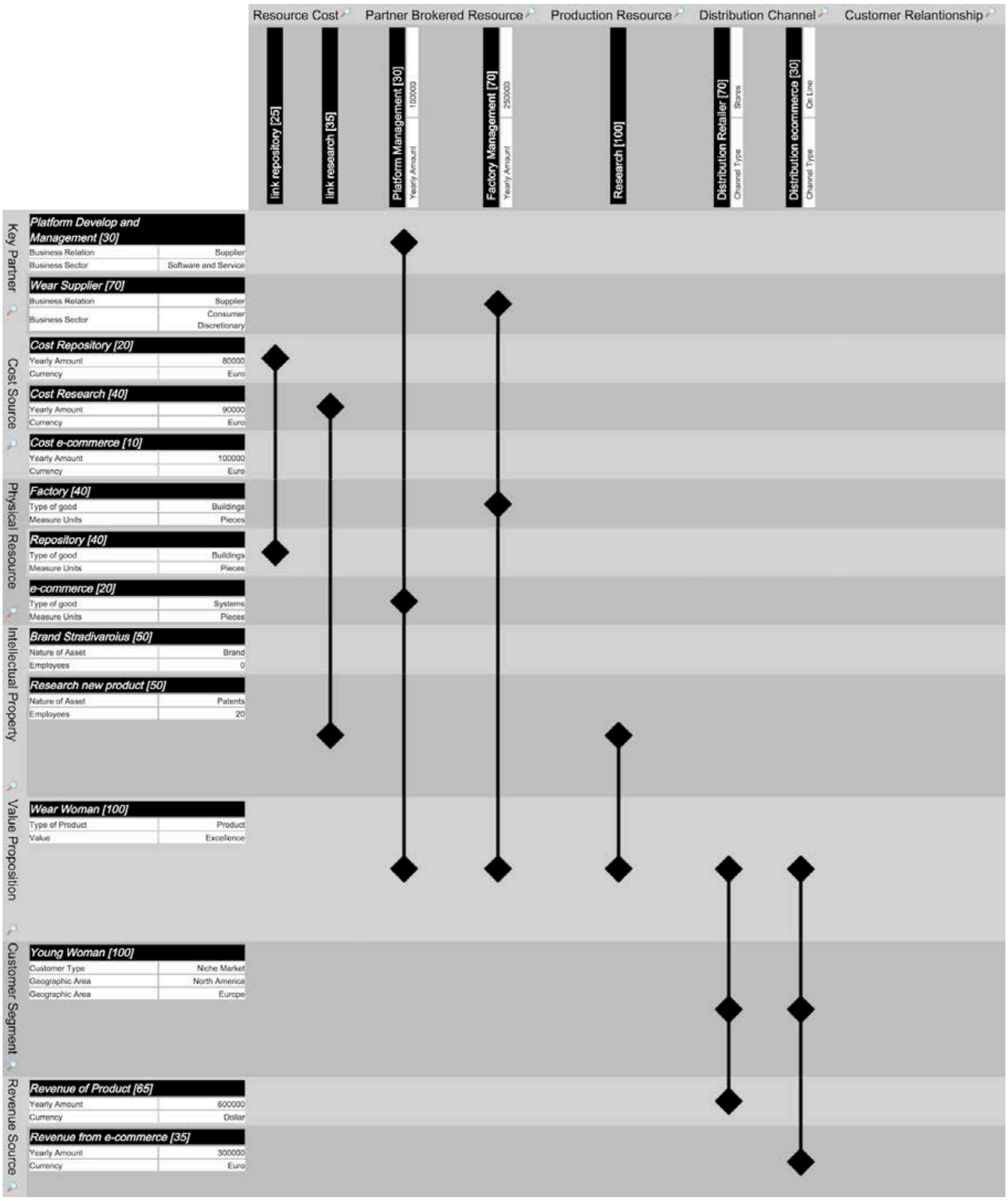

Figure 16 - Business Model Instance of Stradivarius

4.2.1.8. *Case Benetton*

Benetton is an Italian company, which produces clothing of all kinds for all people. The company relies on the manufacturers outside the company and maintains its own staff to the planning and design of new clothes. In our example uses a direct distribution channel formed by its stores located throughout Europe.

We have instantiated the following building blocks:

Key Partner		
	Wear factory	Represents the partners managing the production of clothing
Cost Source		
	Cost warehouse	Is the cost of management of the warehouses owned by the company
	Cost research	Is the cost of staff that creates new models
	Cost retail	Represents the cost of the retail stores
Physical Resource		
	Factory	Represents the factories used by the company to create its products
	Warehouse	Represents the warehouses used to sort the products by the company
	Retail	Represents the stores of the company
Intellectual Resource		
	Brand Benetton	Is the brand of the company
	Research new product	Is the team that designs new products
Value Proposition		
	Wear	Is the product that produces the company
Customer Segment		
	People mass market	Represents clients who are interested in the product provided by the company
Revenue Source		
	Wear sales	Represents the proceeds from the sale of product

Table 32 - Building Block [Benetton]

The enabling relation instantiated to the corresponding building block are the following:

Resource Cost		
	Link warehouse	Connects the costs of the warehouses to the warehouses themselves
	Link research	Connects the costs of staff to the corresponding intellectual resource
	Link retail	Connects the costs of the stores with the stores
Partner Brokered Resource		
	Wear production	Is the production of clothing outsourced
Production Resource		
	Make research	Represents the research staff with the product they are designing new models
Distribution Channel		
	Distribution retailer	Is the distribution channel that connects the retail store to the customer
Customer Relationship		
	---	---

Table 33 - Enabling Relation [Benetton]

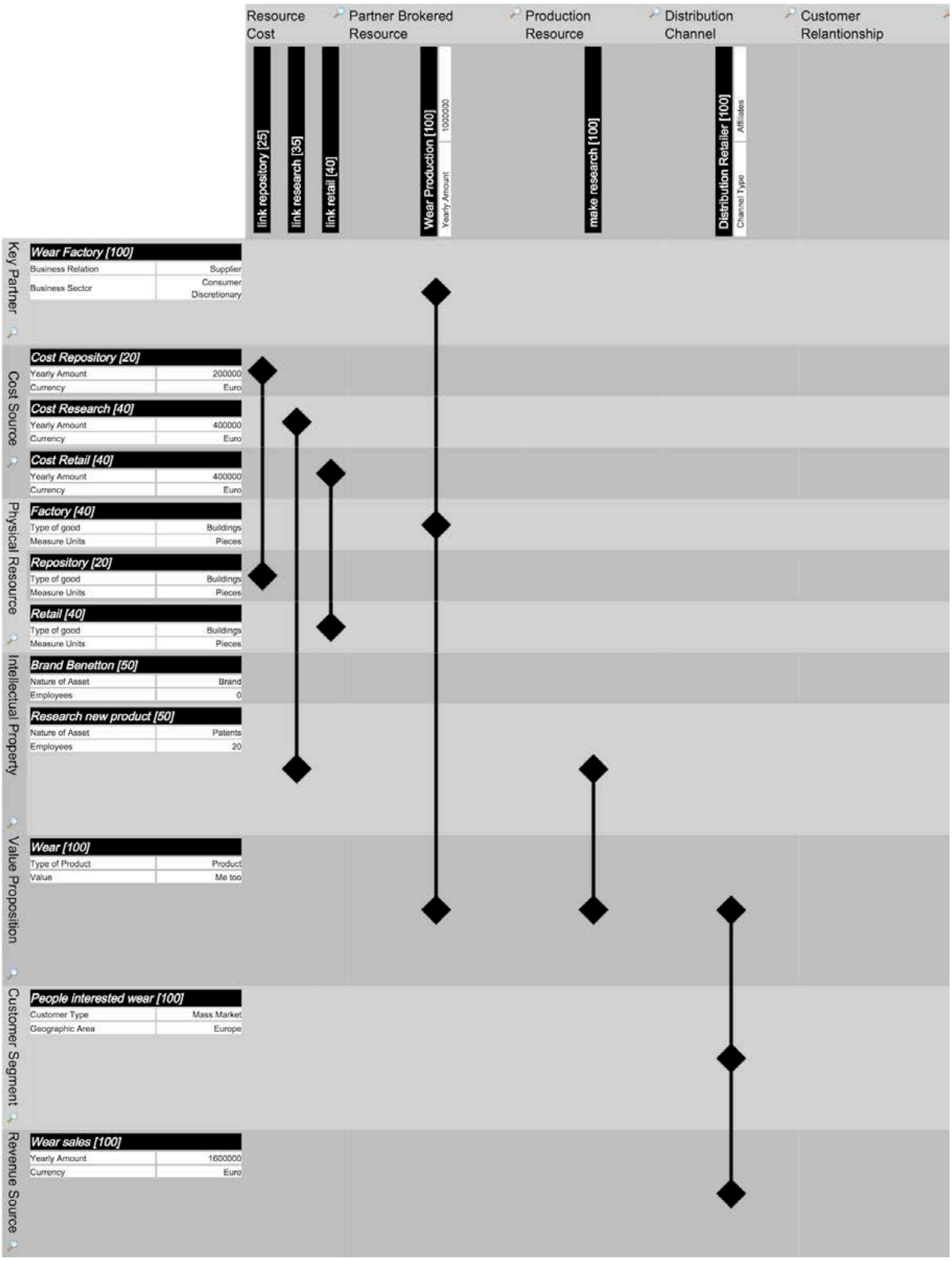

Figure 17 - Business Model Instance of Benetton

4.2.1.9. Case Heels

Heels is the only company that uses the Web distribution channel through its ecommerce site. The products we sell are the only shoes for women. His business shown in this example is very simple. The company supplies through suppliers and fills its stores. Once sold the product from the ecommerce site send your product to the customer.

We have instantiated the following building blocks:

Key Partner		
	Shoes woman supplier	Represents the suppliers of elegant woman shoes
Cost Source		
	Cost warehouse	Is the cost of management of the warehouses owned by the company
	Cost ecommerce	Represents the cost of development and maintenance of the marketplace
Physical Resource		
	Warehouse	Represents the warehouses used to sort the products by the company
	Ecommerce	Is the marketplace used by heels
Intellectual Resource		
	---	---
Value Proposition		
	Elegant woman shoes	Is the product that sell the company
Customer Segment		
	Woman	Represents clients who are interested in the product provided by the company
Revenue Source		
	Shoes sales	Represents the proceeds from the sale of the shoes

Table 34 - Building Block [Heels]

The enabling relation instantiated to the corresponding building block are the following:

Resource Cost		
	Link warehouse	Connects the costs of the warehouses to the warehouses themselves
	Link ecommerce	This relationship linking the ecommerce platform to its cost
Partner Brokered Resource		
	Shoes supplier to warehouse	Represents the providing of elegant shoes from the supplier to the warehouse
Production Resource		
	---	---
Distribution Channel		
	heels.com	Is the distribution channel that utilize the company
Customer Relationship		
	---	---

Table 35 - Enabling Relation [Heels]

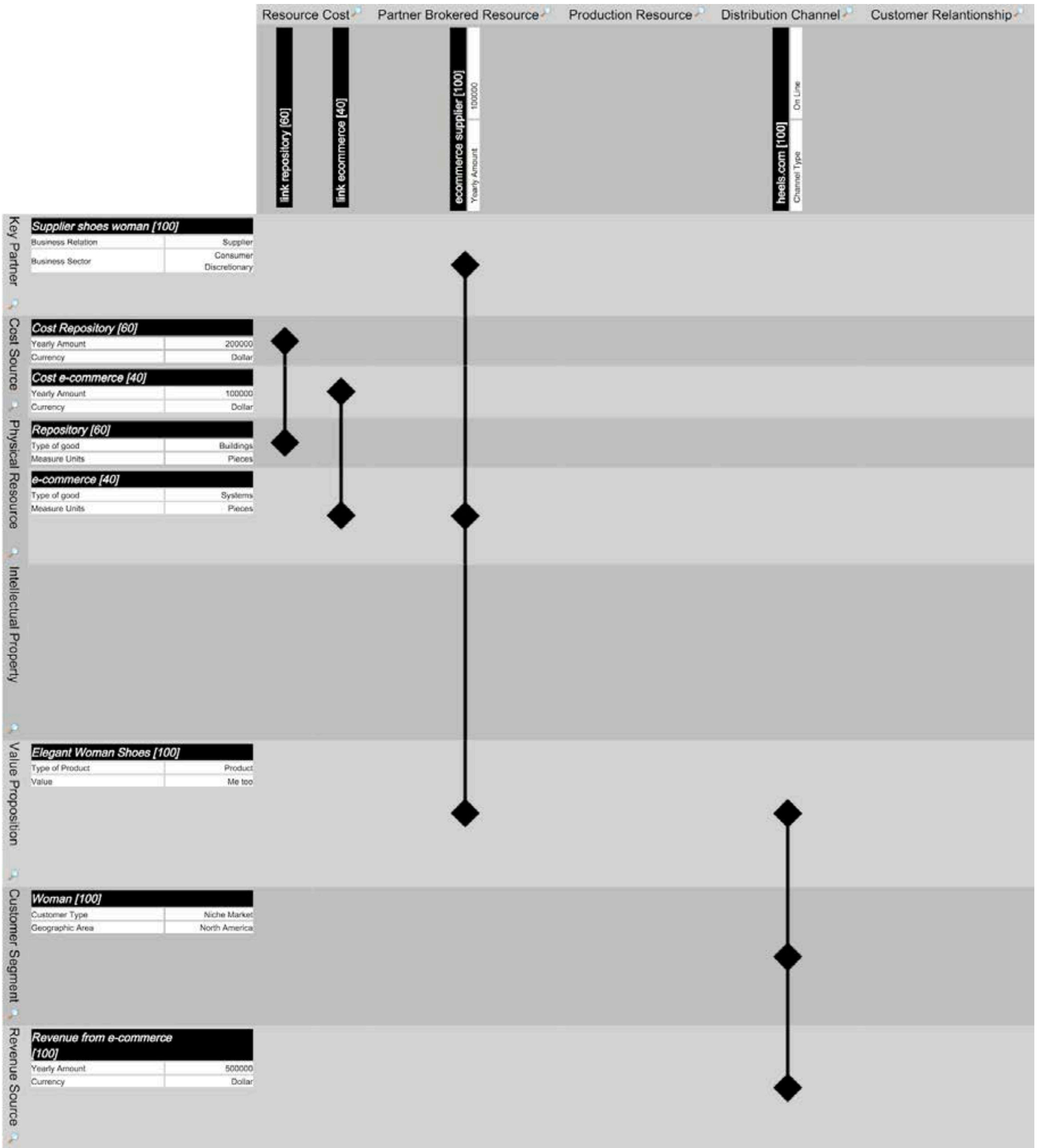

Figure 18 - Business Model Instance of heels

4.2.1.10. Case Decathlon

Decathlon is a company that produces sportswear for all sports and manages the sales through its stores spread throughout Europe. In addition to selling its products also sells products of other manufacturers. In addition to the distribution channel represented by its network of retail uses also its ecommerce website which in our example it maintains itself. Given the two channels of distribution revenues are divided into those arising from the shops and those arising from e-commerce.

We have instantiated the following building blocks:

Key Partner		
	Sport wear supplier	Represents the suppliers of sports clothing
Cost Source		
	Cost factory	Is the cost of management of the factory
	Cost warehouse	Is the cost of management of the warehouses owned by the company
	Cost research	Is the cost of staff that creates new models
	Cost stores	Is the cost of management of the shops owned by the company
	Cost platform	Represents the cost of development and maintenance of the ecommerce site
Physical Resource		
	Factory	Represents the factories used by the company to create its products
	Warehouse	Represents the warehouses used to sort the products by the company
	Stores	Represents the entire chain of stores owned by the company
	Platform	Is the marketplace of Decathlon
Intellectual Resource		
	Brand decathlon	Is the brand of the company
	Research new product	Is the team that designs new products
Value Proposition		
	Sport wear	Is the product that produces the company
Customer Segment		
	People interested sport	Represents clients who are interested in the product provided by the company
Revenue Source		
	Wear sales on stores	Represents the proceeds from the sale of product in the shops
	Wear sales on ecommerce	Represents the proceeds from the sale of product in the ecommerce

Table 36 - Building Block [Decathlon]

The enabling relation instantiated to the corresponding building block are the following:

Resource Cost		
	Link factory	Is the link between the cost of the factory and the factory itself
	Link warehouse	Connects the costs of the warehouses to the warehouses themselves
	Link research	Connects the costs of staff in charge to improve the motor of suggestion the corresponding intellectual resource
	Link stores	Connects the costs of the stores to the stores themselves
Partner Brokered Resource		
	Supplier of wear	Is the business of providing clothing linking suppliers with the product supplied
Production Resource		
	Make product	Is the activity that connects the production with the product in question
	Make research	Represents the research staff with the product they are designing new models
Distribution Channel		
	Distribution stores	Is the distribution channel used by the company to reach its customers, which is made from the warehouses that distribute their products
	decathlon.com	Is the distribution channel between the company and customers interested in sport wear
Customer Relationship		
	---	---

Table 37 - Enabling Relation [Decathlon]

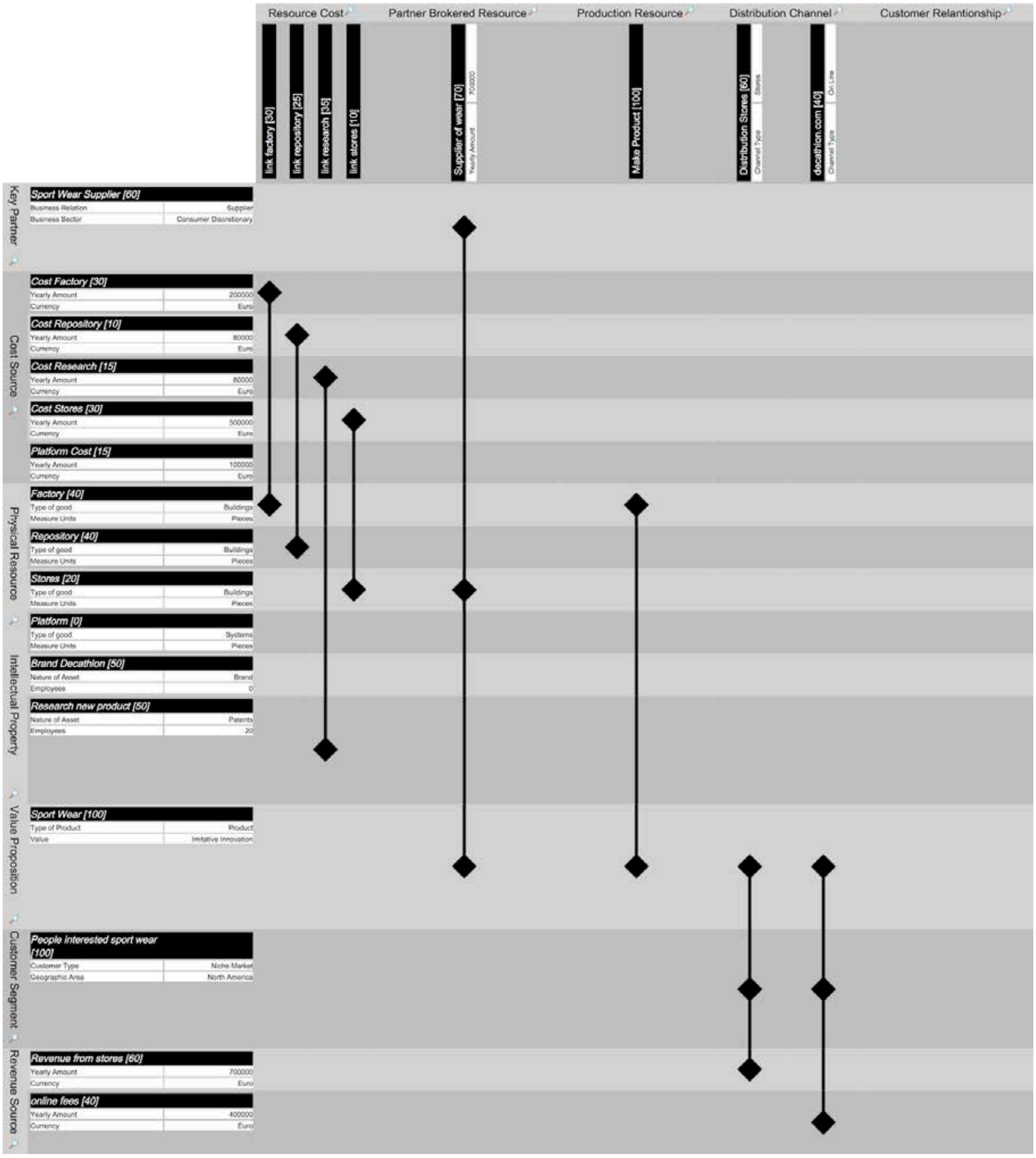

Figure 19 - Business Model Instance of Decathlon

4.2.1.11. *Case New Lions*

New Lions is a company that has as its purpose the sale of footwear of all kinds. The company is only reseller and then gets the products that they sell by the suppliers. Its distribution channel is direct and occurs through the shops spread in the northeast of Italy.

We have instantiated the following building blocks:

Key Partner		
	Shoes supplier	Represents the suppliers of shoes
Cost Source		
	Cost Warehouse	Is the cost of management of the warehouses owned by the company
	Cost Stores	Is the cost of management of the shops owned by the company
Physical Resource		
	Warehouse	Represents the warehouses used to sort the products by the company
	Stores	Represents the entire chain of stores owned by the company
Intellectual Resource		
	---	---
Value Proposition		
	Shoes	Is the product on which the company aims
Customer Segment		
	People interested shoes	Are the customers interested in the company's product
Revenue Source		
	Shoes sales	Represents the proceeds from the sale of shoes

Table 38 - Building Block [New Lions]

The enabling relation instantiated to the corresponding building block are the following:

Resource Cost		
	Link warehouse	Connects the costs of the warehouses to the warehouses themselves
	Link stores	Connects the costs of the stores to the stores themselves
Partner Brokered Resource		
	Supplier stores	It is the business of providing shoes that connects supplier, stores and the company product
Production Resource		
	---	---
Distribution Channel		
	Distribution stores	Company uses its stores as a distribution channel to the end customer
Customer Relationship		
	---	---

Table 39 - Enabling Relation [New Lions]

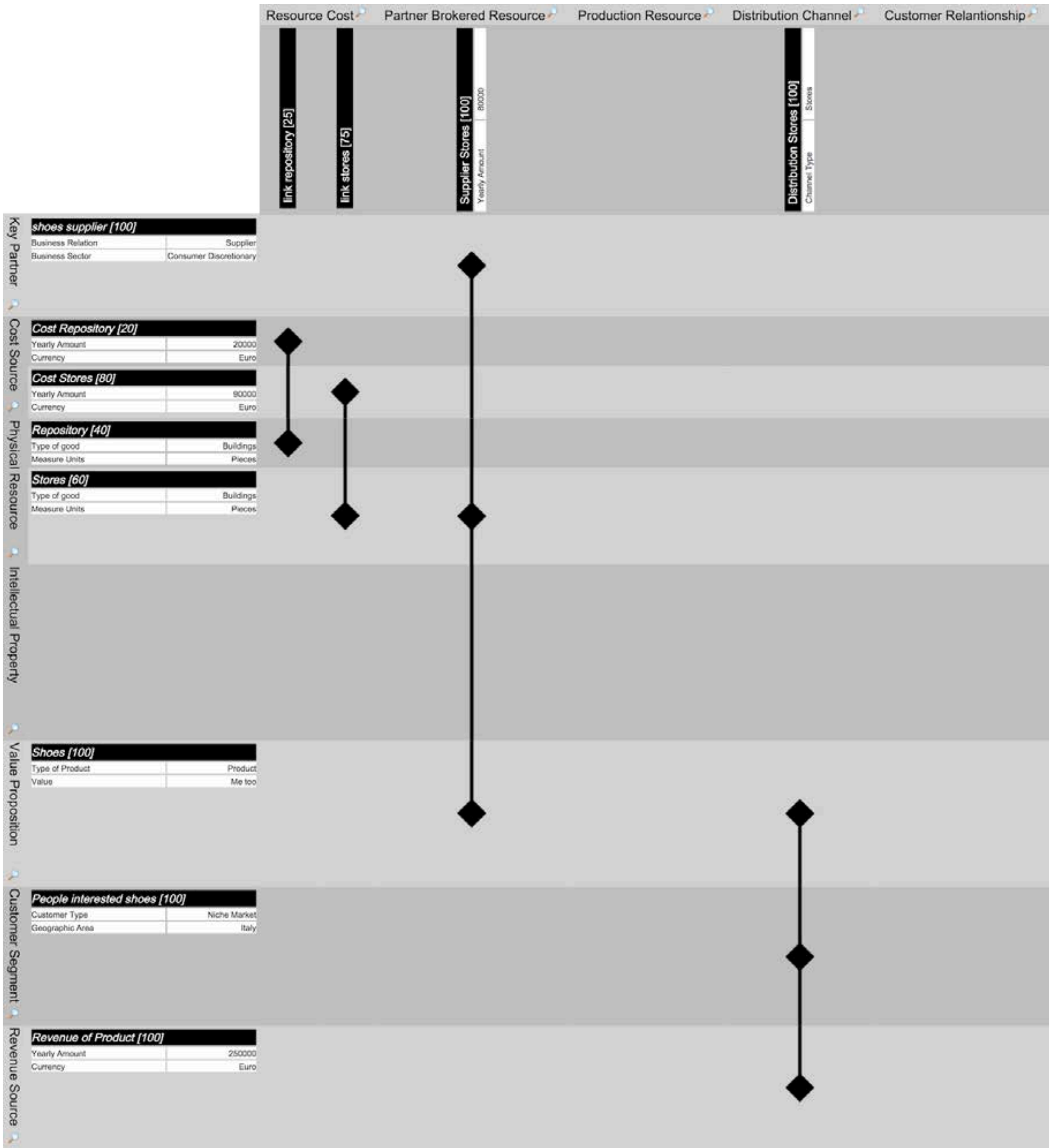

Figure 20 - Business Model Instance of New Lions

4.2.1.12. Case Metropolis Dischi

Metropolis is a business located in Italy that specializes in the sale of vinyl records. This company operates a warehouse and a shop, the latter used as a distribution channel. For the product that sells you can easily understand that focuses on a niche sector.

We have instantiated the following building blocks:

Key Partner		
	Vynil supplier	Represents the suppliers of vynil
Cost Source		
	Cost warehouse	Is the cost of management of the warehouses owned by the company
	Cost stores	Is the cost of management of the shops owned by the company
Physical Resource		
	Warehouse	Represents the warehouses used to sort the products by the company
	Stores	Represents the entire chain of stores owned by the company
Intellectual Resource		
	---	---
Value Proposition		
	Vynil	Is the product on which the company aims
Customer Segment		
	People interested vynil	Are the customers interested in the company's product
Revenue Source		
	Vynil sales	Represents the proceeds from the sale of vynil

Table 40 - Building Block [Metropolis Dischi]

The enabling relation instantiated to the corresponding building block are the following:

Resource Cost		
	Link warehouse	Connects the costs of the warehouses to the warehouses themselves
	Link stores	Connects the costs of the stores to the stores themselves
Partner Brokered Resource		
	Supplier stores	It is the business of providing vynil that connects supplier, warehouse and the company product
Production Resource		
	---	---
Distribution Channel		
	Distribution stores	Company uses its stores as a distribution channel to the end customer
Customer Relationship		
	---	---

Table 41 - Enabling Relation [Metropolis Dischi]

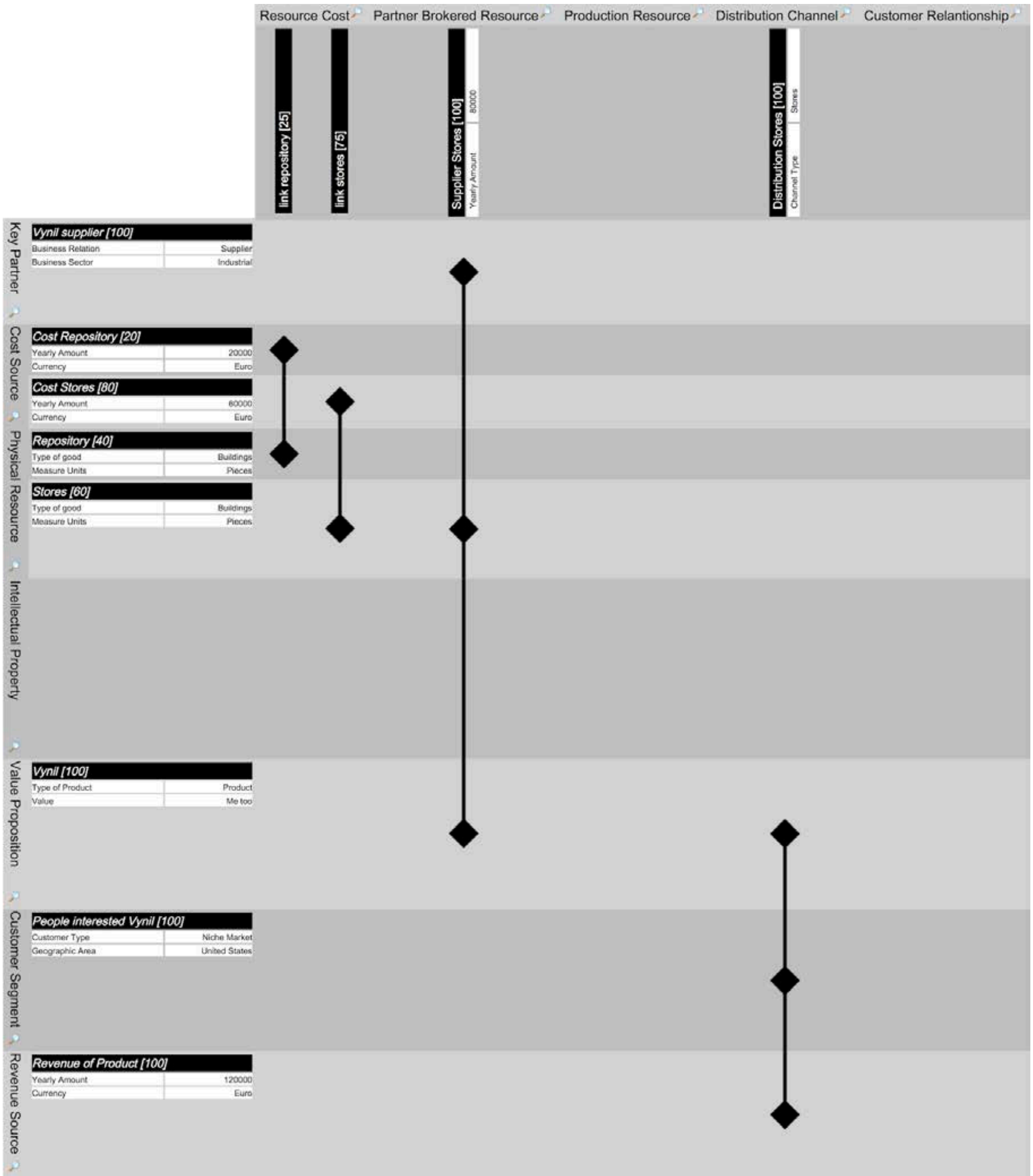

Figure 21 - Business Model Instance of Metropolis Dischi

4.2.1.13. Case iTunes

ITunes is one part of Apple that is responsible for sales of digital music. Was one of the first companies that have revolutionized the music industry, in that it have persuaded the record companies to sell individual songs even at very low prices. In our example is very simple to understand how its business model. iTunes makes available for sale the songs that are made available by its partners, record labels, and uses its ecommerce site as a distribution channel to reach its customers.

We have instantiated the following building blocks:

Key Partner		
	Record company	Represents the suppliers of music
Cost Source		
	Cost ecommerce	Represents the cost of development and maintenance of the ecommerce site
Physical Resource		
	Ecommerce	Is the marketplace of iTunes
Intellectual Resource		
	---	---
Value Proposition		
	Music a very low price	Represents clients who are interested in the product provided by the company
Customer Segment		
	People interested music	Represents clients who are interested in the product provided by the company
Revenue Source		
	Music sales	Represents the proceeds from the sale of product in the shops

Table 42 - Building Block [iTunes]

The enabling relation instantiated to the corresponding building block are the following:

Resource Cost		
	Link ecommerce	Linking the cost of development and maintenance of the marketplace to the resource marketplace
Partner Brokered Resource		
	Ecommerce supplier	Represents the mode of delivery of the music, connecting the suppliers, to the ecommerce until the product supply
Production Resource		
	---	---
Distribution Channel		
	Distribution ecommerce	Is the distribution channel between the company and customers interested in music
Customer Relationship		
	---	---

Table 43 - Enabling Relation [iTunes]

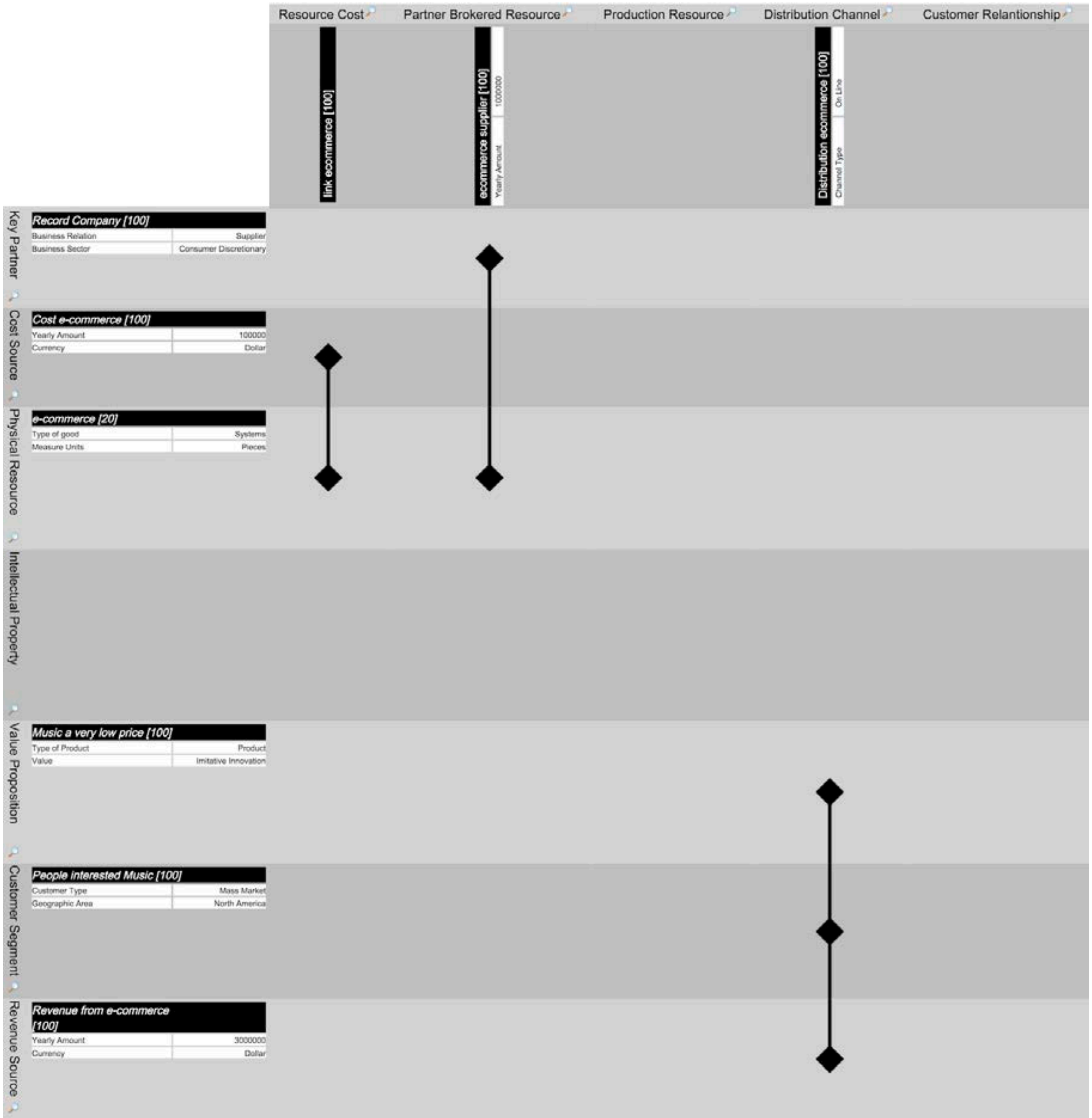

Figure 22 - Business Model Instance of iTunes

4.2.1.14. Case PC City

PC City is primarily a reseller of computer hardware and computer software. In addition offers the repair service of personal computers of customers. For this reason, in our example, there is the intellectual resource "repair team" that represents the staff responsible of this service. As a reseller ordering the necessary components from suppliers, places them in the warehouse and then distribute them at all its stores.

We have instantiated the following building blocks:

Key Partner		
	Software supplier	Represents the suppliers of the components of software
	Hardware supplier	Represents the suppliers of the components of hardware
Cost Source		
	Cost stores	Is the cost of management of the shops owned by the company
	Cost warehouse	Is the cost of management of the warehouses owned by the company
	Cost team	Is the cost of staff that repair pc
Physical Resource		
	Stores	Represents the entire chain of stores owned by the company
	Warehouse	Represents the warehouses used to sort the products by the company
Intellectual Resource		
	Repair team	Is the team that repair pc
Value Proposition		
	Hardware component	Is the product that produces the company
	Software	Is the product that produces the company
	Repair PC	Is the service that offers the company
Customer Segment		
	People interested hw and sw	Represents clients who are interested in the product provided by the company
	People interested pc repair	Represents clients who are interested in the service offered by the company
Revenue Source		
	Software sales	Represents the proceeds from the sale of software in the shops
	Hardware sales	Represents the proceeds from the sale of hardware in the shops
	Service fees	Represents revenues from repairing pc

Table 44 - Building Block [PC City]

The enabling relation instantiated to the corresponding building block are the following:

Resource Cost		
	Link stores	Connects the costs of the stores to the stores themselves
	Link warehouse	Connects the costs of the warehouses to the warehouses themselves
	Link team	Connects the costs of staff in charge to improve the motor of suggestion the corresponding intellectual resource
Partner Brokered Resource		
	Software to warehouse	It is the business of providing software that connects suppliers, warehouse and the company's product
	Hardware to warehouse	It is the business of providing hardware that connects suppliers, warehouse and the company's product
Production Resource		
	---	---
Distribution Channel		
	Distribution software	The company uses its stores as a distribution channel to the end customer
	Distribution hardware	The company uses its stores as a distribution channel to the end customer
	Distribution service	The company uses its stores as a service to the end customer
Customer Relationship		
	---	---

Table 45 - Enabling Relation [PC City]

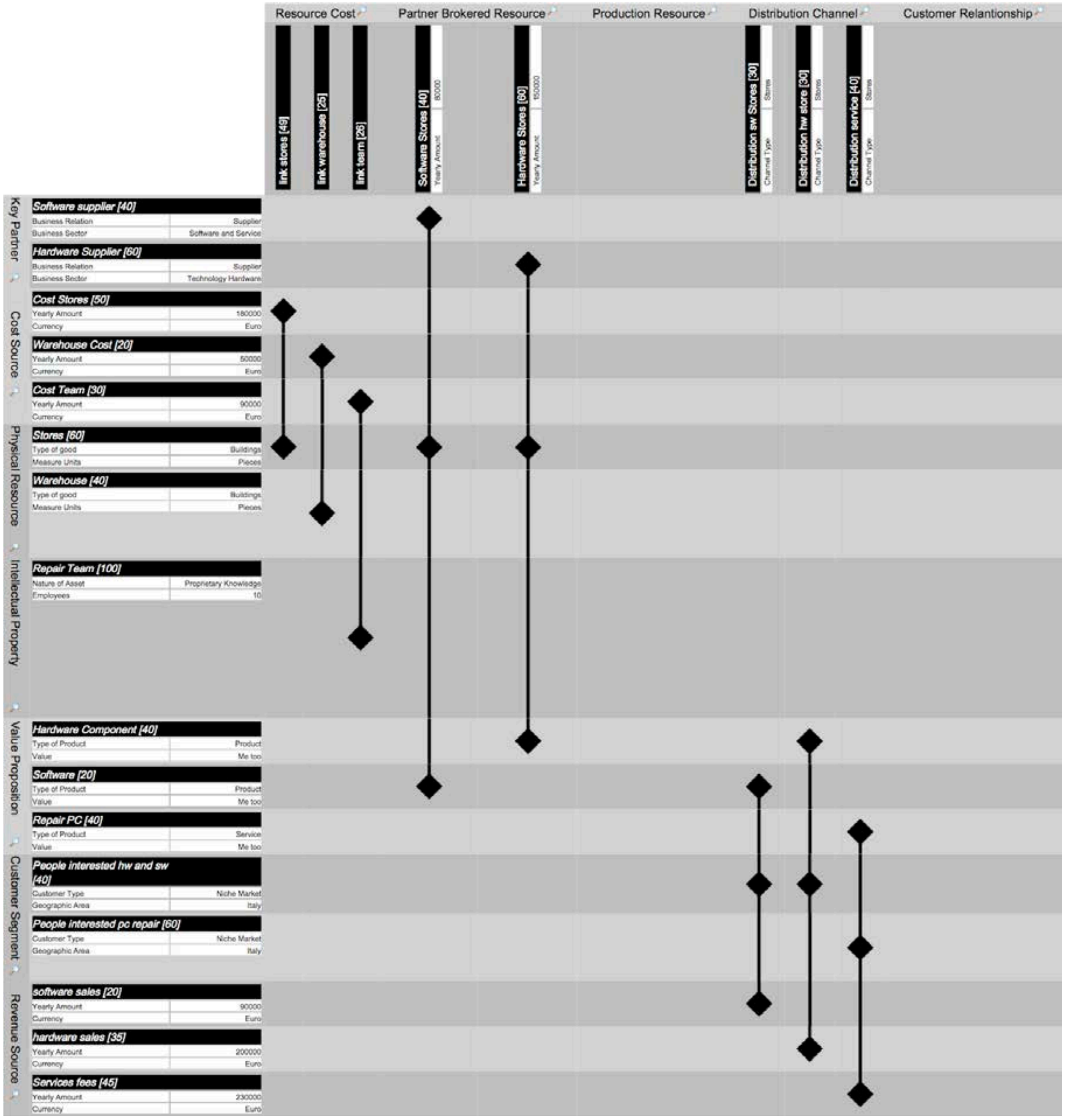

Figure 23 - Business Model Instance of PC City

4.2.1.15. Case MediaWorld

Mediaworld is a dealer of electronic components, starting from small household appliances to large. Its business model is the order of the components from specialized suppliers and uses two distribution channels to reach their customers. The classic stores represent the first while the second is via the Web through its ecommerce site.

We have instantiated the following building blocks:

Key Partner		
	Supplier of electronics	Represents the suppliers of the components of electronic
Cost Source		
	Cost Warehouse	Is the cost of management of the warehouses owned by the company
	Cost ecommerce	Represents the cost of development and maintenance of the marketplace
	Cost Stores	Is the cost of management of the shops owned by the company
Physical Resource		
	Warehouse	Represents the warehouses used to sort the products by the company
	Ecommerce	Is the marketplace of Mediaworld
	Stores	Represents the entire chain of stores owned by the company
Intellectual Resource		
	---	---
Value Proposition		
	Electronics	Is the product on which the company aims
Customer Segment		
	People interested electronics	Are the customers interested in the company's product
Revenue Source		
	Electronics sales on stores	Represents the proceeds from the sale of stores
	Electronics sales on ecommerce	Represents the proceeds from the sale of ecommerce

Table 46 - Building Block [MediaWorld]

The enabling relation instantiated to the corresponding building block are the following:

Resource Cost		
	Link stores	Connects the costs of the stores to the stores themselves
	Link warehouse	Connects the costs of the warehouses to the warehouses themselves
	Link ecommerce	Linking the cost of development and maintenance of the marketplace to the resource marketplace
Partner Brokered Resource		
	Supplies electronics	Represents the mode of delivery of the electronic components, connecting the suppliers, to the warehouses until the product supply
Production Resource		
	---	---
Distribution Channel		
	Distribution stores	Is the distribution channel used by the company to reach its customers, which is made from the warehouses that distribute their products
	Distribution ecommerce	Is the distribution channel between the company and customers interested in sport wear
Customer Relationship		
	---	---

Table 47 - Enabling Relation [MediaWorld]

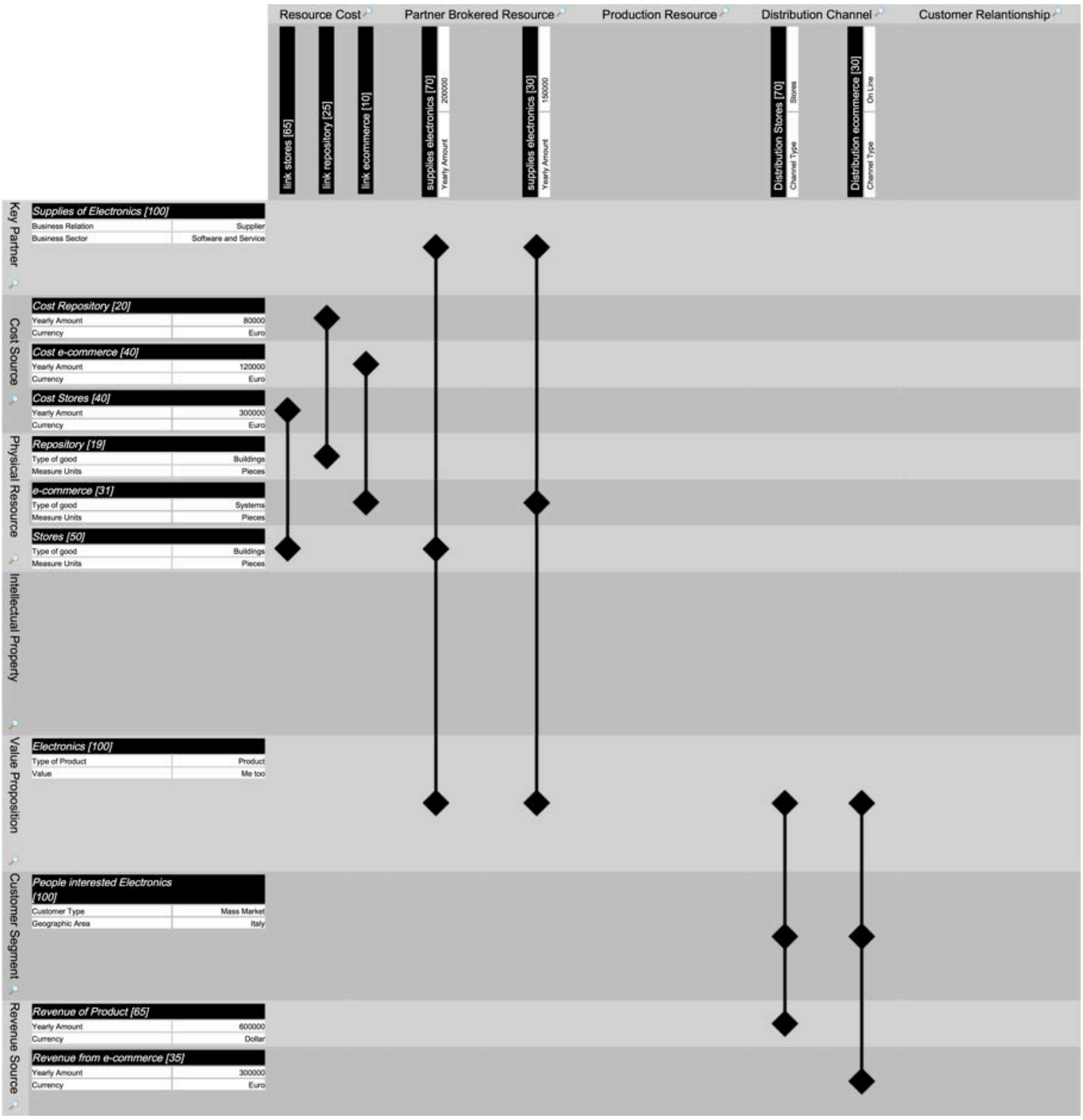

Figure 24 - Business Model Instance of PC City

5. Mining Structured Business Model Database

In this chapter we will discuss how you can make mining operations on our database of business model instance. The main purpose of this analysis is to help the human operator to make decisions thanks to computer support. We have identified three possible types of analysis that can help experts in business administration: detection of such behavior among the various companies in our database, find synergies between companies and finally by creating custom algorithms, possibly with the help of an expert, finding useful patterns to specific analysis. The identification of such behavior can help, for example, whether a company is using the right suppliers, or if there is other more efficient or cheaper used by other companies. Or you can tell if a particular company can expand its market through other channels of distribution or gain new customers whereas other similar companies already do. With regard to the synergies between companies, for example, if more companies have customers in the same geographic area is possible to evaluate whether they can use a warehouse in common, thereby reducing overall costs. Finally, with the creation of custom algorithms it is possible for the operator look for certain useful behaviors that can help you make decisions through analysis of all the companies in the database, not just those known to him. After discussing what can be achieved thanks to the mining of our database, let's see how we can achieve this analysis.

5.1. State of the art

In the literature, there are two distinct formulations for Frequent Subgraph Mining that can be identified by:

- transaction graph based;
- single graph based.

The main distinction between the two formulations consists of the input data. In the first formulation, transaction based graph, the input data is a collection of relatively simple small graphs [called transaction by Agrawal and Srikant, 1994]. For this purpose several algorithms have been developed, [Holder] has proposed an algorithm to find the best structures compressed, [Inokuchi] proposed an algorithm based on Apriori for discovering all frequent substructures, [Coatney] has proposed an algorithm to find common substructures in biochemical molecules and [Cong] applied this model in hierarchical semi structured data. While in the second, simple graph based, the input data is a very large single graph. The typical example of this formulation is to discovering structural regularities or anomalies in social network or web structures, which are single graphs and we not want to split them in parts [Ruoyu Zou].

5.2. Overview

First we need to define how the business models are represented in our database. The instances of business models have been defined as the graph with the vertex set corresponding to the building block and the set of edge corresponding to the enabling relation. In addition, in our case, all vertexes and all edges are labeled by type of building block instantiated then we can talk about labeled graph.

Labeled graph*: A labeled graph can be represented as $G(V,E,L_V,L_E, \varphi_v, \varphi_e)$, where V is a set of vertexes, $E \subseteq V \times V$ is a set of edges; LV and LE are vertex and edge labels respectively; and φ_v and φ_e are the corresponding functions that define the mapping $V \rightarrow L_V$ and $E \rightarrow L_E$.*

As defined by our structure appears to be a weight applied to each vertex and each edge. In fact at each instance of building block or enabling relation is given a weight, the sum of all the weights of the instances of a single type is a hundred. The aim of our analysis is to find common subgraphs, so for completeness of information we give the definition of subgraph.

Subgraph*: A graph $G' = (V', E')$ is a subgraph of another graph $G = (V, E)$ if its vertex set V' is a subset of V and its edge set E' is a subset of E. The subgraph relationship is denoted as $G' \subseteq_S G$.*

Now we will give the definition of our problem, Frequent Subgraph Mining, but first have to define the support of a subgraph.

***Support**: Given a collection of graphs $\mathcal{G}$, the support for a subgraph g is defined as the fraction of all graphs that contain g as its subgraph, i.e.:*

$$s(g) = \frac{|\{G_i | g \subseteq_S G_i, G_i \in \mathcal{G}\}|}{\mathcal{G}}$$

***Frequent Subgraph Mining**: Given a set of graph $\mathcal{G}$ and a support threshold, minsup, the goal of frequent subgraph mining is to find all subgraphs g such that $s(g) \geq minsup$.*

While this formulation is applicable to any type of graph we will focus only on graphs and associated indirect because this is the representation in our case. The definition of these graphs are given below:

*A graph is **connected** if there exists a path between every pair of vertices in the graph, in which a path is a sequence of vertices <$v_1 v_2 \ldots v_k$> such that there is an edge connecting between every pair of adjacent vertices (v_i, v_{i+1}) in the sequence.*

*A graph is **undirected** if it contains only undirected edges. An edge (v_i, v_j) is undirected if it is indistinguishable from (v_j, v_i).*

Mining of frequent subgraphs is a computationally expensive task because of the search space exponentially. To understand the complexity of this activity considers a set that contains k entities. In frequent itemset mining each entity is seen as an item and the size of the search space to be explored is 2^d, which is the number of candidate itemsets that can be generated. In frequent subgraph mining, each entity is a vertex and can have up to d-1edges to other vertices. Considering that labels associated with the vertices are unique then it follows that the total number of subgraphs is:

$$\sum_{i=1}^{k} \binom{k}{i} \times 2^{i\left(\frac{i-1}{2}\right)}$$

Where $\binom{k}{i}$ is the number of ways to choose *i* vertices to form a subgraph and $2^{i\left(\frac{i-1}{2}\right)}$ is the maximum number of edges between vertices. The number of candidate subgraphs is much lower because we consider only those connected while the previous formulation also included those disconnected. A brute force method to do this is to generate all connected subgraphs as candidates and calculates their support. Given the large size of the set of candidates this method may break down even for moderately sized graphs.

5.3. Apriori algorithm

Now we see how an Apriori algorithm can be developed to find frequent subgraphs as proposed by [Inokuchi]. First we need to transform each graph in a transaction-like format so that the algorithm Apriori can be applied. In this way any combination of the label of an edge and the labels of the corresponding vertices is mapped to an item. The width of the transaction is given by the number of edges in the graph. The steps of an Apriori algorithm for mining frequent subgraphs consist of:

- **Candidate generation**, which is the process of merging pairs of frequent (k-1)-subgraphs to obtain a candidate k-subgraph.
- **Candidate pruning**, which is the process of discarding all candidate k-subgraphs that contain infrequent (k-1)-subgraphs.
- **Support counting**, which is the process of counting the number of graphs in $\mathcal{G}$ that contain each candidate.
- **Candidate elimination**, which discards all candidate subgraphs whose support counts are less than *minsup*.

5.3.1. Candidate generation

There are two possible strategies for solving the first step, one is vertex growing while the other is the edge growing. During this phase, we form the candidate k-subgraphs by merging pairs of frequent (k-1)-subgraphs. The value k can be the number of vertex or the number of edge. To avoid generating duplicate candidates, we may impose an additional condition for merging, that the two (k-1)-subgraphs must share a common (k-2)-subgraph. The common (k-2)-subgraph is known as their core. The first strategy, vertex growing, is the process of generating new candidates by adding a new vertex into an existing frequent subgraph. This procedure is done by creating an $k \times k$ adjacency matrix by the merging of a pair of $(k-1) \times (k-1)$ adjacency matrices. The result of this operation is a graph that contains one or two edges more than the original graph. The operation of subgraph merging via vertex growing is done in this way: an adjacency matrix is merged with another matrix if the submatrices obtained by removing the last row and the last column are identical to each other. The resulting matrix is the first matrix, appended with the last row and the last column of the second. The remaining entries of the new matrix are either zero or replaced by all valid edge labels connecting the pair of vertices. Instead the second strategy is to insert an edge to an existing subgraph frequently during the step of generating candidates. Unlike the previous strategy, the subgraph resulting not necessarily increases the number of vertices in the original graphs. The procedure of subgraph merging via edge growing is done in this way: a frequent subgraph is merged to another frequent subgraph only if the subgraph obtained by removing an edge from the first is topologically equivalent to the subgraph obtained by removing an edge from the second. After merging, the resulting candidate subgraph is the first with the addition of an edge taken from the second.

5.3.2. Candidate pruning

So the result of this first step is the generation of candidate k-subgraphs during the second step, candidate pruning, we must eliminate those candidates who do not frequent (k-1)-subgraphs. To do this, we proceed by removing an edge from the candidate k-subgraph and checking if the corresponding (k-1)-subgraph is connected and frequent, if not the candidate is eliminated. To check if a (k-1)-subgraph is frequent should be matched with another frequent (k-1)-subgraph. Determine whether two graphs are topologically equivalent, or isomorphic, is called *graph isomorphism problem*. The standard way to handling the problem of isomorphism between graphs is mapping every graph in a single string known as the canonical label as proposed by [Babai, 1983]. This has the property that its code is the same whether two graphs are isomorphic. With this property, we can test the isomorphism between graphs comparing the canonical labels of graphs. To construct the canonical label of a graph we proceed in two steps. The first is to find a representation of the graph with an adjacency matrix. First of all we must say that a graph has multiple representations can be made using adjacency matrices because there are several ways to sort the vertices. To find all possible representations of adjacency matrices of a graph we have to consider all possible permutations of rows in the matrix. Mathematically, each permutation corresponds to a multiplication of the adjacency matrix initial with the corresponding permutation matrix. The second step is to determine the string representation of each adjacency matrix. Since the adjacency matrix is symmetric it is sufficient to construct the string representation based on the upper triangular of the matrix. Finally, we must compare all the string representations of the graph and choose the one that has the greater or lesser lexicographic value. This method is very expensive because it requires the computation of all possible adjacency matrices of a graph and computes for each of them the string representation in order to find the canonical label. Has been developed several methods to reduce the complexity of this algorithm, some include a cache of labels canonical already calculated others by reducing the number of permutations necessary to determine the label canonical adding information to the vertices.

5.3.3. Support Counting

Also this operation is potentially very expensive because they must be certain all subgraphs contained in each graph belonging to $\mathcal{G}$. One possible strategy, as proposed by [Kuramochi], to decrease the computational cost is to maintain a list of IDs of graphs associated with each frequent subgraph. Each time a new candidate k-subgraph is generated, from the union of a pair of frequent (k-1)-subgraphs, its corresponding list of IDs of the graphs is intersected. Finally, the test of isomorphism of subgraphs is performed in the graphs on the lists of intersection to determine if they contain a particular candidate subgraph.

5.4. Subgraph weighting schemas

To improve the significance of subgraphs found you can use different weighting schemes associated with a weighting function. In the literature there are different weighting functions and are divided into two categories: according to the structure or according to content. Considering the structure of our graphs we can avoid this kind of functions because we already have the relative weights of both vertices of the edges. For example, given a subgraph g with vertexes $\{v_1, v_2, \dots, v_k\}$, if the corresponding weights for vertexes are $\{w_1, w_2, \dots, w_k\}$, then one common way of computing a weight for g is given by [Chartrand and Zhang, 2004] as $\sum_{v_i \in g} w_i$. From this obligatory choice given the structure of our graphs we can apply two different weighting schemes: Average Total Weighting (ATW) and Affinity Weighting (AW).

5.4.1. Average Total Weighting (ATW) scheme

This scheme has been proposed by [Jiang and Coenen 2008] and can be used both for the weights on the vertices that for those on the edges. Given a graph data set $\mathcal{G} = \{G_1, G_2, \dots, G_n\}$, the weight for a subgraph g is calculated by dividing the sum of the average weights in the graphs that contain g with the sum of the average weights across the entire graph data set $\mathcal{G}$. Let us now see how it can be applied considering the weight on the edges, in the same manner can be used for the weight on the vertices.

Given a graph data set $\mathcal{G} = \{G_1, G_2, \dots, G_n\}$*, if* G_i *is edge weighted by* $\{w_1, w_2, \dots, w_k\}$*, then the average weight associated with* G_i *is defined as:* $W_{avg}(G_i) = \frac{\sum_{j=1}^{k} w_j}{k}$ *where* w_j *is the weight of the edge considered.*

Following we define the total weight of the whole data set: $W_{sum}(\mathcal{G}) = \sum_{i=1}^{n} W_{avg}(G_i)$. Finally, using the last two equations we can find the weight of the subgraph.

Given a graph data set $\mathcal{G} = \{G_1, G_2, \dots, G_n\}$ *and an arbitrary subgraph g, let the set of the graphs where g occurs equal* $\delta_{\mathcal{G}}(g)$*. Then, the weight of g with respect to* $\mathcal{G}$ *is:* $W_{\mathcal{G}}(g) = \frac{\sum_{G_i \in \delta_{\mathcal{G}}(g)} W_{avg}(G_i)}{W_{sum}(\mathcal{G})}$*.*

This equation quantifies the importance of the subgraph considered in the entire data set. The weighed support of a subgraph is then defined as the product of the support of g and the importance factor of g: $wsup_{\mathcal{G}}(g) = W_{\mathcal{G}}(g) \cdot sup_{\mathcal{G}}(g) = \frac{W_{\mathcal{G}}(g) \cdot |\delta_{\mathcal{G}}(g)|}{|\mathcal{G}|}$.

5.4.2. Affinity Weighting (AW) scheme

The second scheme proposed by [Jiang 2010], Affinity Weighting, is based on two elements to prevent the growth of the search space: the graph distance measure and the weighting ratio measure. The first element is defined as follows:

Given a graph data set $\mathcal{G} = \{G_1, G_2, \ldots, G_n\}$*, and a subgraph g, let the set of graphs where g occurs equal* $\delta_{\mathcal{G}}(g)$*. Then, the weight of g is formulated as a distance metric,* $dist(g, \delta_{\mathcal{G}}(g))$ *:* $W_{\mathcal{G}}(g) = \frac{\sum_{G_i \in \delta_{\mathcal{G}}(g)} 1 - \frac{|V(g)|}{|V(G_i)|}}{C}$*, where* $C = |V(g)|$*.*

This measure is based on the number of vertices in a graph, the ratio weighting also relies on the weight of the edges. The weighting ratio of a subgraph g is a function *r(g)* that returns a value between zero and one and is defined: $r(g) = \frac{MIN_{w_i \in S}\{w_i\}}{MAX_{w_j \in S}\{w_j\}}$.

5.5. Handling Categorical and Continuous Attribute

The structure that we have proposed provides the possibility to have for each vertex or edge one or more categorical attributes we called ontologies. With these can be made a more accurate search for common subgraphs between the business models. Ontologies in our structure are represented as a hierarchical tree. By means of this it is possible to choose only those frequent subgraphs that satisfy the selected attribute or that are a child of it. For example, you can choose only the business model that has a client in the geographic "North America" that includes not only herself also the “United States” and “Canada”. It 'easy to see that thanks to these attributes, you can reduce the search space and enable you to search in a more specific way within the database. Another important element that is present in our structure and that can be useful for the search of subgraphs common is the ability to have continuous attributes. In our structure, the attributes are declared as numeric, percentage or vote. With these attributes is possible to use a method of discretization, in order to choose a range for which that particular attribute is interesting and then discard all those subgraphs that have a vertex or edge with that attribute is not inherent for research purposes. For example, in our case studies, you can enter a discriminating on the summit "Revenue Source" to select only those that have the attribute "Yearly Amount" between one hundred thousand and two hundred thousand.

5.6. Future Scenarios

The algorithms developed so far for the problem of finding frequent common subgraphs in a data set of graphs take into account only the associated weight or the vertices or edges. In our case we have instead the weight of both. One solution is to transform the edges in vertices, in fact for our structure is defined as all the enabling relation can have the same characteristics as the building block. In this way it is possible to rely only on the vertices weight and also it is possible to check a reduction in the search time since now we consider two different sets of vertices, which can not be connected to their internal but only towards the other set.

Another feature that should be developed is the inclusion directly in the algorithm of search of the ability to manipulate categorical and continuous attributes in the various types of building blocks and enabling relations.

5.7. Example in our set of Business Model Instance

As described in the introduction of this chapter we can perform different types of analysis in our database. From the research of subgraphs common it is possible to perform two types of analysis. The first is that relating to the interventions about economies of scale or logistics. Is possible to suggest economies of scale for companies not in competition that have a customer segment in the same geographical area and use the same category of suppliers. In this type of situation it is possible to suggest to place orders in common, application of economy of scale, for lowering the price of the goods. From this type of cooperation is possible to reduce the costs and have no problems with competition because the companies interested do not have clients in common. From the logistical point of view it is possible to find situations with companies that have the same type of customer segment but offering a different value proposition with different distribution channels. It is therefore possible to suggest the companies in question, which are not in competition because they have a different value proposition, to use the distribution channels used by others. In this way, companies are using new distribution channels that had not yet thought and save on running costs. Can be made research from knowledge already in possession of management experts. In fact these can perform two types of searches with two different purposes: the first possibility is to search for virtuous patterns while the second is to search for bad patterns. A virtuous pattern is a subgraph that experience management expert brings benefit to the company that employs him. Such research is very useful in order to achieve statistics and for the testing of hypotheses a posteriori. For example, an expert can check if the application of a given mechanism recommended by him has had positive outcomes or not. But a bad pattern is a subgraph that is considered by the expert management as harmful to businesses that use it. In this regard, such research is useful in order to suggest to companies that are still using that particular mechanism to change it to a more efficient one. After this brief explanation we see practical examples applied to our database. It was not possible to use the mining algorithms, because of the small size of our data set. In fact, in order to make a good analysis, we should have a few hundred business model entered by the experts of management. To overcome this problem we have included some forty business model in our database, as seen in chapter four, and we performed a manual of common patterns among them. As a first example, consider the common subgraph consists of the following building blocks and enabling relation [Figure 25].

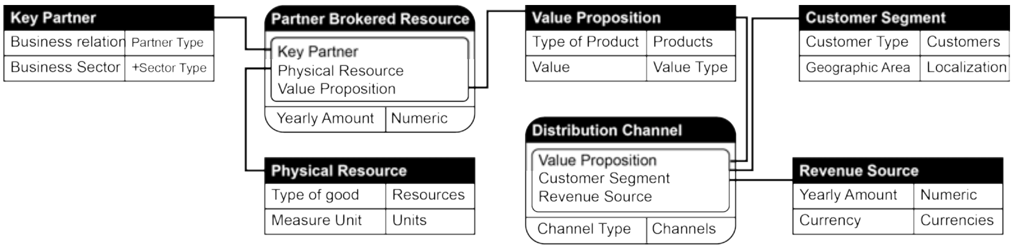

Figure 25 - First Pattern Example

The pattern shown in Figure 25 operates in the following business model of our database:

- *Diamond Comic Distributors (book);*
- *Amazon (book and electronics);*
- *Barnes and Nobles (book);*
- *Body Glove (fashion);*
- *Brema Sport (fashion);*
- *Stradivarius (fashion);*
- *Benetton (fashion);*
- *Heels (footwear);*
- *Decathlon (fashion and footwear);*
- *New Lion (footwear);*
- *Metropolis Dischi (music);*
- *iTunes (music);*
- *PC City (electronics);*
- *Mediaworld (electronics).*

Now insert the specialization that the customer segments is localized in *Italy*, and we have the same type of supplier, we choose *shoes supplier*. At this point, the search provides the following two results:

- *Brema Sport (footwear);*
- *New Lion (footwear).*

A management expert at this point would suggest the two companies to place orders in common to have a more competitive price from the supplier. This is because the value proposition between the two companies is different and therefore the companies are not in competition. We now present a pattern common to most other business model in our data set [Figure 26] that considers only the enabling relations that contain a *channel type* of type "online".

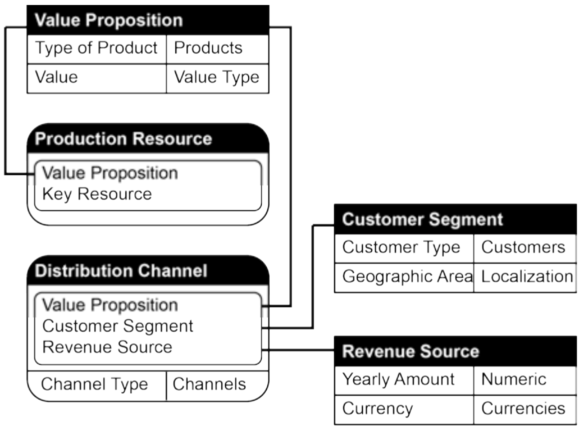

Figure 26 - Second Pattern Example

This pattern is common for the following companies in our database:

- *Lulu (book);*
- *Amazon (book and electronics);*
- *Stradivarius (fashion);*
- *Heels (footwear);*
- *Decathlon (fashion and footwear);*
- *iTunes (music);*
- *Mediaworld (electronics).*

Now a management expert may be noted that among the companies, that use a online distribution channel, there are two, lulu and amazon, using a customer relationship to provide a service loyalty with customers. From this type of analysis could be suggested to companies that do not apply a service customer relationship to design one to allow greater loyalty with customers.

6. Conclusions

With this thesis we were able to give a complete representation, from a processing point of view, of a business model. We have laid a very flexible structure to represent any type of business model canvas, this will also represent new canvas and is not limited only to existing ones. Not only can you describe any canvas but you can also compare different structures through the use of ontologies. A set of common ontologies used on a canvas with the ability to create mappings between the various structures and then make them comparable although structured differently. We left flexibility for representation of any canvas but we have defined a structure, the language xml, rigid and well structured. Thanks to these features will not be possible to exit from the parameters with which the structure is to be built. In chapter four we described the business model of many companies and all values, channels and different partners. This shows that our model is suitable to represent any type of business. In chapter five we defined as the road is to be taken for the implementation of a data mining algorithm for our data set of graphs. It was not possible to apply the algorithm on our data set because we have not a sufficient number of business models for applying a mining algorithm and check the results. To show that it is possible to obtain results we have described, again in chapter five, the algorithms carried out by hand on our data set of forty companies. Were also defined possible custom searches based on experience of management experts and describes how they can be useful. Finally, we describe possible future scenarios for better results and reduce the search time.

Bibliography

- J.-R. Abrial. Data semantics. In IFIP Working Conference Data Base Man- agement, pages 1–60, 1974.
- R. Agrawal and R. Srikant. Fast Algorithms for Mining Association Rules. 1994.
- H. Akkermans, Z. Baida, J. Gordijn, N. Pena, A. Altuna, and I. n. Laresgoiti. Value webs: Using ontologies to bundle real-world services. IEEE Intelligent Systems, 19(4):57–66, July 2004. ISSN 1541-1672. doi: 10.1109/MIS.2004.35. URL http://dx.doi.org/10.1109/MIS.2004.35.
- L. Babai and E. Luks. Canonical labeling of graphs. Proc. 15th ACM Symposium on Theory of Computing. 1983. pp. 171–183.
- H. A. Babaie and A. Babaei. Developing the earthquake markup language and database with uml and xml schema. Comput. Geosci., 31(9):1175– 1200, Nov. 2005. ISSN 0098-3004. doi: 10.1016/j.cageo.2004.12.010. URL http://dx.doi.org/10.1016/j.cageo.2004.12.010.
- P. Bambury. A taxonomy of internet commerce. First Monday, 3(10), 1998. A. P. G. Brown. Modelling a real world system and designing a schema to represent it. In IFIP TC-2 Special Working Conference on Data Base Description, pages 339–348, 1975.
- T. Burkhart, J. Krumeich, D. Werth, and P. Loos. Analyzing the business model concept - a comprehensive classification of literature. In ICIS, 2011.
- M. Chabrol and D. Sarramia. Object oriented methodology based on uml for urban traffic system modeling. In Proceedings of the 3rd international conference on The unified modeling language: advancing the standard, UML'00, pages 425–439, Berlin, Heidelberg, 2000. Springer-Verlag. ISBN 3-540- 41133-X. URL http://dl.acm.org/citation.cfm?id=1765175.1765219
- G. Chartrand and P. Zhang. Distance in Graphs – Taking the Long View.

- P. P. Chen. A preliminary framework for entity-relationship models. In Proceedings of the Second International Conference on the Entity-Relationship Approach to Information Modeling and Analysis, ER ’81, pages 19–28, Amsterdam, The Netherlands, The Netherlands, 1983. North-Holland Publishing Co. ISBN 0-444-86747-3.
- P. P.-S. Chen. The entity-relationship model toward a unified view of data. ACM Trans. Database Syst., 1(1):9–36, Mar. 1976. ISSN 0362-5915. doi: 10.1145/320434.320440.
- M. Coatney, S. Parthasarathy. MotifMiner: Efficient discovery of common substructures in biochemical molecules.
- G. Cong, L. Yi, B. Liu and K. Wang. Discovering Frequent Substructures from Hierarchical Semi-structured Data.
- Q. Dai and R. J. Kauffman. Business models for internet-based b2b electronic markets. Int. J. Electron. Commerce, 6(4):41–72, July 2002. ISSN 1086-4415.
- De Lucia, C. Gravino, R. Oliveto, and G. Tortora. An experimental comparison of er and uml class diagrams for data modelling. Empirical Softw. Engg., 15(5):455–492, Oct. 2010. ISSN 1382-3256. doi: 10.1007/s10664-009- 9127-7. URL http://dx.doi.org/10.1007/s10664-009-9127-7.
- M. Dean and G. Schreiber. OWL web ontology language reference. W3C recommendation, W3C, Feb. 2004.
- R. Elmasri and G. T. J. Wuu. A temporal model and query language for er databases. In Proceedings of the Sixth International Conference on Data Engineering, pages 76–83, Washington, DC, USA, 1990. IEEE Computer Society. ISBN 0-8186-2025-0. URL http://dl.acm.org/citation.cfm?id=645475.654165.
- S. Ferg Modelling the time dimension in an entity-relationship diagram. In Proceedings of the Fourth International Conference on Entity-Relationship Approach, pages 280–286, Washington, DC, USA, 1985. IEEE Computer Society. ISBN 0-444-87951-X. URL http://dl.acm.org/citation.cfm?id=647510.726517.

- P. Fettke and P. Loos. Reference Modeling for Business Systems Analysis. Gale virtual reference library. Idea Group Pub., 2007. ISBN 9781599040554. URL http://books.google.it/books?id= IRKPgAACAAJ.
- D. J. Flynn and M. D. Jazi. Organisational and information system modelling for information systems requirement determination. In P. Loucopoulos, editor, Entity-Relationship Approach - ER 94, Business Modelling and Re-Engineering, 13th International Conference on the Entity-Relationship Approach, Manchester, U.K., December 13-16, 1994, Proceedings, volume 881 of Lecture Notes in Computer Science, pages 79–93. Springer, 1994. ISBN 3-540-58786-1. doi: db/conf/er/FlynnJ94.html.
- G. L. Geerts and W. E. McCarthy. An accounting object infrastructure for knowledge-based enterprise models. IEEE Intelligent Systems, 14(4): 89–94, July 1999. ISSN 1541-1672. doi: 10.1109/5254.784089. URL http://dx.doi.org/10.1109/5254.784089.
- J. Gordijn, A. Osterwalder, and Y. Pigneur. Comparing two Business Model Ontologies for Designing e- Business Models and Value Constellations. Evo- lution, (Osterwalder 2004), 2005.
- H. Gregersen. Timeerplus: a temporal eer model supporting schema changes. In Proceedings of the 22nd British National conference on Databases: enterprise, Skills and Innovation, BNCOD05, pages 41–59, Berlin, Heidelberg, 2005. Springer-Verlag. ISBN 3-540-26973-8, 978-3-540-26973-1.
- M. Hardwick. Extending the relational database data model for design applications. In Proceedings of the 21st Design Automation Conference, DAC84, pages 110–116, Piscataway, NJ, USA, 1984. IEEE Press. ISBN 0-8186-0542- 1.
- S. Hartmann and S. Link. English sentence structures and eer modeling. In Proceedings of the fourth Asia-Pacific conference on Comceptual modelling - Volume 67, APCCM ’07, pages 27–35, Darlinghurst, Australia, Australia, 2007. Australian Computer

Society, Inc. ISBN 1-920-68285-X. URL http://dl.acm.org/citation.cfm?id=1274453.1274460.

- L. B. Holder, D. J. Cook and S. Djoko. Substructures Discovery in the SUBDUE System.
- A. Inokuchi, T. Washio and H. Motoda. An Apriori-based Algorithm for Mining Frequent Substructures from Graph Data.
- K. E. Iverson. A programming language. In Proceedings of the May 1-3, 1962, spring joint computer conference, AIEE-IRE '62 (Spring), pages 345–351, New York, NY, USA, 1962. ACM. doi: 10.1145/1460833.1460872.
- Jacobson, G. Booch, and J. Rumbaugh. The unified software development process. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 1999. ISBN 0-201-57169-2.
- C. Jiang and F. Coenen. Graph-based image classification by weighting scheme. In Proc. AI'2008, Springer, pages 63–76, 2008.
- C. Jiang, F. Coenen and M. Zito. Frequent Sub-graph Mining on Edge Weighted Graphs. 2010.
- K. M. Khan, M. Kapurubandara, and U. Chadha. Incorporating business requirements and constraints in database conceptual mod- els. In Proceedings of the first Asian-Pacific conference on Concep- tual modelling - Volume 31, APCCM '04, pages 59–64, Darlinghurst, Australia, Australia, 2004. Australian Computer Society, Inc. URL http://dl.acm.org/citation.cfm?id=976297.976304.
- M. Kuramochi and G. Karypis. Frequent Subgraph Discovery. 2001.
- T. Lin, C. Li, M.-C. Shan, and S. Babu. A framework for business operations management systems. In Proceedings of the 8th international conference on Web information systems engineering, WISE'07, pages 461–471, Berlin, Heidelberg, 2007. Springer-Verlag. ISBN 3-540-76992-7, 978-3-540-76992-7.
- URL http://dl.acm.org/citation.cfm?id=1781374.1781425.
- X. Liu, Z. Han, J. Liu, and C.-x. Shen. Develop secure database system with security extended er model. In Proceedings of the 9th

international conference on Knowledge-Based Intelligent Information and Engineering Systems - Volume Part III, KES05, pages 1005–1010, Berlin, Heidelberg, 2005. Springer-Verlag. ISBN 3-540-28896-1, 978-3-540-28896-1.

- E. T. Neal. A new uml-compatible object relationship notation (orn). In Proceedings of the 38th annual on Southeast regional conference, ACM-SE 38, pages 179–183, New York, NY, USA, 2000. ACM. ISBN 1-58113-250-6. doi: 10.1145/1127716.1127756. URL http://doi.acm.org/10.1145/1127716.1127756.
- Osterwalder and Y. Pigneur. Business Model Generation: A Handbook for Visionaries, Game Changers, and Challengers. Wiley, 1 edition, July 2010. ISBN 0470876417.
- Osterwalder, C. Parent, and Y. Pigneur. Setting up an ontology of business models. In J. Grundspenkis and M. Kirikova, editors, CAiSE Workshops (3), pages 319–324. Faculty of Computer Science and Information Technology, Riga Technical University, Riga, Latvia, 2004. ISBN 9984-9767-3-4.
- V. Stefanov, B. List, and B. Korherr. Extending uml 2 activity diagrams with business intelligence objects. In Proceedings of the 7th international conference on Data Warehousing and Knowledge Discovery, DaWaK'05, pages 53–63, Berlin, Heidelberg, 2005. Springer-Verlag. ISBN 3-540-28558-X, 978- 3-540-28558-8.
- P. Timmers. Business models for electronic markets. Electronic Markets, 8: 3–8, 1998.
- T. Weilkiens. Systems Engineering with SysML/UML: Modeling, Analysis, Design. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2008. ISBN 0123742749, 9780123742742.
- Zott, R. H. Amit, and L. Massa. The Business Model: Theoretical Roots, Recent Developments, and Future Research. Sept. 2010. URL http://ssrn.com/abstract=1674384.
- R. Zou and L. B. Holder. Frequent Subgraph Mining on a Single Large Graph Using Sampling Techniques.