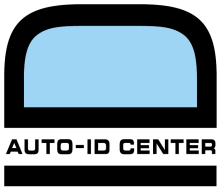
Updated edition published November 1, 2001. Distribution restricted to sponsors until February 1, 2002.



WHITE PAPER

Integrating the Electronic Product Code (EPC) and the Global Trade Item Number (GTIN)

David L. Brock

MIT AUTO-ID CENTER MASSACHUSETTS INSTITUTE OF TECHNOLOGY, 77 MASSACHUSETTS AVENUE, BUILDING 3-449G, CAMBRIDGE, MA 02139-4307

ABSTRACT

The Electronic Product Code (EPC) was conceived as a means to identify all physical objects. These include not only retail products, but also containers, packages and shipments, as well as more general physical systems, assemblies and components. The EPC is a short, simple and extensible code designed primarily for efficient referencing to networked information. The Uniform Product Code (U.P.C.) and its numbering superset, the Global Trade Item Number (GTIN), perform a similar function for product identification and supply chain logistics. There are, however, many important differences in objective, scope, structure and implementation of these coding methods. This paper explores the differences in approach and present a method for integrating the EPC and GTIN.

WHITE PAPER

Integrating the Electronic Product Code (EPC) and the Global Trade Item Number (GTIN)

Biography

by David L. Brock
Co-Director

Dr. David Brock received Bachelors degrees in theoretical mathematics and mechanical engineering from MIT, and his Masters and Ph.D. degrees from the Department of Mechanical Engineering at MIT with an affiliation to the Artificial Intelligence Lab. He is currently a Principal Research Scientist in the Laboratory for Manufacturing and Productivity and Co-Director of the MIT Auto-ID Center. Dr. Brock is also the Founder of Brock Rogers Surgical, a manufacturer of robotic medical devices. Dr. Brock has worked with a number of organizations including the Artificial Intelligence Laboratory, the Massachusetts Eye and Ear Infirmary, DARPA, Lockheed-Martin, Loral, BBN and Draper Laboratories.

WHITE PAPER

Integrating the Electronic Product Code (EPC) and the Global Trade Item Number (GTIN)

Contents

The Electronic Product Code (EPC) was conceived as a means to identify all physical objects. These include not only retail products, but also containers, packages and shipments, as well as more general physical systems, assemblies and components. The EPC is a short, simple and extensible code designed primarily for efficient referencing to networked information. The Uniform Product Code (U.P.C.) and its numbering superset, the Global Trade Item Number (GTIN), perform a similar function for product identification and supply chain logistics. There are, however, many important differences in objective, scope, structure and implementation of these coding methods. This paper explores the differences in approach and present a method for integrating the EPC and GTIN.

1. INTRODUCTION

The Electronic Product Code (EPC) was conceived as a means to uniquely identify all physical objects [3]. This includes not only retail and commercial products, as with the current Uniform Product Code, but also physical assemblies, components and systems. The Code itself contains very little information, but essentially serves as a reference to networked information. Together with the Object Name Service (ONS) and the Physical Markup Language (PML), the EPC connects the physical world and virtual world [1,2,4].

The Global Trade Item Number (GTIN) is a family of item and shipment codes, which has for years identified products and containers. Since the EPC and GTIN perform similar functions, we must consider the necessity, compatibility and application of these standards.

Although similar in function, there are a number of important differences. First, the GTIN is actually a family of coding standards, which differ between North America and the rest of the world. The EPC, on the other hand, has a single format serving to identify the physical entity. Second, some variants of the GTIN codes carry information, such as price and weight. The EPC is contains essentially no information beyond object identity. Third, the GTIN codes were design to work in conjunction with linear bar-codes, providing one or more checksum digits to validate scans. The EPC separates the identification number from the communication media. Finally, although the GTIN standards cover a wide range of product identification, the EPC is intended for an even greater breadth of application – enumerating physical objects, assemblies, components and systems.

Rather than ignore globally accepted standards, we will show how the EPC is compatible with elements of the GTIN. Specifically, we propose to map the currently accepted Uniform Code Council (UCC) and European Article Numbering (EAN) International Company and Product Numbering Systems into the EPC Manager and Product Type identifiers. Thus the current standards for product numbering can be integrate into the broader vision of the 'networked physical world.'

2. ELECTRONIC PRODUCT CODE (EPC)

The Electronic Product Code (EPC) is a numbering scheme that provides unique identification for physical objects, assemblies and systems. Information is not stored in the code, but serves only as a reference to on-line – or Internet-based – information.

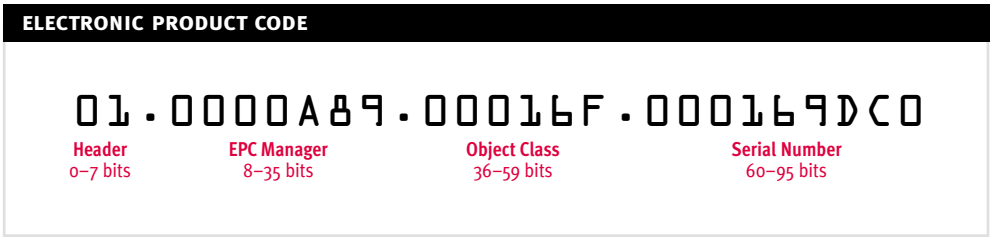
The EPC is a 96-bit code divided into four, fixed length partitions, as shown in Figure 1 [3]. The first partition is an 8-bit **header**, which defines the number, type and length of subsequent data partitions. The purpose of the header is to provide extensibility for future, unanticipated data requirements.

The second partition identifies the **EPC manager**; that is the company, or entity, responsible for maintaining the subsequent codes. In other words, it is the EPC manager's responsibility to maintain both object type codes and serial numbers in their domain. The EPC manager ensures reliable operation of the Object Name Service and maintains and publishes the on-line product documents. The EPC manager partition spans a 28-bit section, encoding a maximum of 2^{28} = 268,435,456 companies.

The next partition, **object class**, occupies the next 24-bits. When applied to retail products, the object class may be considered the skew or stock keeping unit (SKU). It may also be used for lot number or any other object-grouping scheme developed by the EPC manager. Since each organization is allowed more than 16 million object types, this partition could encode all the current UPC SKUs, as well as many other object classes. This is important as we expand beyond retail applications into general supply chain and manufacturing.

The final partition encodes a unique object identification number. For all objects of a similar type, the **EPC serial number** provides 36-bits, or 2^{36} = 68,719,476,736, unique identifiers. Together with the product code, this provides each company with 1.1×10^{18} unique item numbers – currently beyond the range of all manufactured products.

Figure 1. The Electronic Product Code (EPC) provides unique identification for all physical objects, assemblies and systems.



3. GLOBAL TRADE ITEM NUMBER (GTIN)

The Global Trade Item Number (GTIN) is a system for identifying manufacturers and items through the retail and commercial supply chain. The GTIN is more accurately a family of coding structures, which grew from retail checkout in the United States to general logistics worldwide, as shown in Table 1.

The GTIN began as a means to identify grocery items and facilitate automatic checkout at the supermarket using bar-code laser scanners. The Uniform Product Code (U.P.C.) has for years sped the tracking and identification of merchandise in stores, as well as in industrial and commercial applications.

Since the first scan in 1973, the grocery industry established the U.P.C. as the standard bar code system for product identification. The original U.P.C. grew to include a European variant, the EAN/UCC-13, a Japanese Article Numbering Code (JAN) (identical to the EAN codes with flag characters set to "49"),

Table 1. The Uniform Code Council (UCC) and European Article Numbering (EAN) International Global Trade Item Number (GTIN) is a family of coding structures to identify suppliers and items through the retail and commercial supply chain.

TERMS	LEGACY TERMS	DOMAIN
UCC-12	U.P.C. (Universal Product Code)	North America
EAN/UCC-13	EAN Code	Outside North America
EAN/UCC-14	SCC-14 (Shipping Container Code)	Worldwide
EAN/UCC-8	EAN-8	Outside North America

and systems for identifying shipments, the Shipping Container Code (SCC-14). By increasing the length of the numbers from 8 to 14, the various codes coexist under a proper superset, the Global Trade Item Number system, as shown in Table 2.

The GTIN is organized and administrated by the Uniform Code Council (UCC) [5] and the European Article Numbering (EAN) International [6]. The UCC and EAN International represent an alliance of organizations with numerous member companies worldwide. The UCC/EAN coding standards includes the GTIN and its subsets, as well as many other coding standards. In the following sections, we will outline in somewhat more detail the code formats of the GTIN family.

Table 2. The Global Trade Item Number (GTIN) system allows coding substructures, including the UCC-12, EAN/UCC-13 and EAN/UCC-14 to coexist under the GTIN superset

CODING SUBSTRUCTURES

CODE	1	2	3	4	5	6	7	8	9	10	11	12	13	14
UCC-12	0	0	1	2	3	4	5	6	7	8	9	10	11	12
EAN/UCC-13	0	1	2	3	4	5	6	7	8	9	10	11	12	13
EAN/UCC-14	1	2	3	4	5	6	7	8	9	10	11	12	13	14
EAN/UCC-8	0	0	0	0	0	0	1	2	3	4	5	6	7	8

3.1. UCC-12

The Universal Product Code UCC-12 (U.P.C.) is 12-digit numeric sequence, which encodes numbering system, manufacturer, item number and checksum digit. The standard symbol, which represents the UCC-12, is a series of light and dark lines, together with a human readable numeric equivalent, as shown in Figure 2.

The most common variant, UPC Version A, has 12 digits, as shown the figure. However, there are four other versions, UPC Version B, C, D and E. A zero-suppressed version, UPC Version E, is an 8-digit code used on items with packaging too small to accommodate the full bar code, and is shown in Figure 3. Obviously, to convert from a UPC Version A to UPC Version E, there must be at least four zeros in the data set.

Figure 2. The Universal Product Code UCC-12 (U.P.C.) (UPC version A shown) is a numeric sequence representing numbering system, manufacturer, product and checksum.

UNIVERSAL PRODUCT CODE

The other three versions of the UPC code were in the original specification, but for the most part are abandoned. UPC Version B is a special version originally intended to handle the National Drug Code

and National Health Related Item Codes. It allowed 11 digits and one product type code and did not have a modulo check digit. UPC Version C is a special code designed to promote industry-wide compatibility. Finally, a UPC Version D is variable length version used in limited special applications. The code contained at least 12 digits. The first digit is a product type code, the next 10 digits carried data and the twelfth digit was a check sum. There followed a variable number of digits that encoded data.

Figure 3. The UPC version E is a zero suppression version of the UPC and is intending to be used on packaging too small to accommodate the full version A. These are commonly seen on soda cans and bottles.

The UPC Version A numeric code (which is distinct from its "bar code symbol" representation) consists of four partitions [7]. The first partition consists of a single digit indicating the numbering system used to interpret the remaining characters. A '0', for example, designates a regular UPC code, a '3' a National Drug Code, and a '5' a coupon. A complete list is given in Table 3.

Table 3. The UCC-12 provides a single digit to represent the numbering standard.

NUMBER	USAGE
0	Regular UPC Codes
1	Reserved
2	Random weight items
3	National Drug Code
4	Internal company use
5	Coupon
6	Regular UPC Codes
7	Regular UPC Codes
8	Variable partition codes
9	Reserved

The next five digits designate the manufacture identification number, and the following five, the item number. Item numbering is maintained by the manufacture that must ensure unique numbers for each product type. Finally, a single digit is added as a check character used to validate the correct interpretation of the machine scan. Thus, the UPC code can, provide up to 10 unique numbering systems, 100,000 manufacturer identifiers, and 100,000 product types for each manufacturer.

3.1.1. Random Weight Items

Not all UPC codes, however, follow the regular structure given above [8]. North American retail products sold after being weighed or measured have a different partitioning structure, as shown in Figure 4.

Designated with a '2' in the first digit, the number is divided into five additional sections. The second partition is a 1-digit code called the Packager Code. Unlike fixed content products, the random weight ID Number allows only one position to identify the supplier of the product. Traditionally digits '0,' ..., '3' indicate the retailer and digits '4,' ..., '9' their suppliers. The supplier and retailer must manually agree to the Packager Code assignment.

The next partition contains 4-digits indicating the commodity identification number. The Commodity ID allows 10,000 product types, which are divided among the trade associations for beef, veal, pork, chicken and others. The trade associate for beef, for example, uses '1314' to identify ground round. The fourth partition is a Price Check Digit, followed by a 4-digit price in dollars and cents. The final digit is the modulo checksum.

Figure 4. Random weight items use the bar code structure to indicate the type of commodity and the price

RANDOM WEIGHT ITEMS

3.1.2. National Drug Codes or National Health Related Item Codes

A '3' in the first digit of the GTIN UPC subset identifies a pharmaceutical product using the National Drug Code (NDC) or a medical/surgical product using the National Health Related Item Code (NHRIC). The Food and Drug Adminstration (FDA) assigns the data portion of the code. Placing a '3' in the first position and a checksum in the last forms the UPC-12 number, as shown in Figure 5.

The FDA NDC or NHRIC number consists of a variable length Labeler Code of either 4 or 5 digits followed by 5 or 6 digits for the Product/Package. The first digit of the Labeler Code indicates the organization of the remaining positions. A '0' in the first data digit indicates the company has a 4-digit NDC Labeler Code, a '1,' ..., '7,' a 5-digit NDC Labeler Code, an '8' a 4-digit NHRIC Code and a '9' a 5-digit NHRIC Code.

Figure 5. National Drug Codes (NDAC) and National Health Related Item Codes (NHRIC) are embedded in the UPC.

NATIONAL DRUG CODES AND NATIONAL HEALTH RELATED ITEM CODES

3.1.3. Internal Use

A '4' in the first digit of the UPC code indicates an internally assigned number. UCC member companies cannot use this type of bar code in the open supply chain, since the number is **not unique**. The numbers used internally within a company are at their complete discretion. No specification is made on the organization and structure of the number, or the format, size and quality of the bar-code symbol.

3.1.4. Coupons

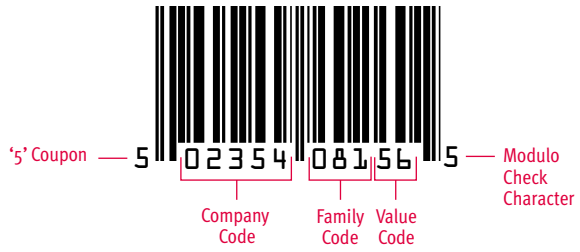
The UPC Code also identifies and describes coupons. A '5' in the first digit indicates a manufactures' coupon for redemption. The identification number is formed by supplementing the UPC Coupon Code with one of five types of Coupon Extended Code Formats. Each of these formats varies in length and structure.

The UPC Coupon Code including the Couple Extended Code Format 1 is shown in Figure 6. The first portion, the UPC Coupon Code, has the same structure independent on coupon type. It consists of five partitions. The first partition is the single digit type, which for coupons is indicated by a '5.' The second is the 5-digit UCC Company Code.

The third partition consists of 3 digits which indicate the Family Code, Summary Codes and Super Summary Codes. The Family Codes relate to individual products, while the other are assigned to larger groups of products. The Uniform Code Council details the Family Codes, as described in the **U.P.C. Coupon Code Guidelines Manual** [9].

Figure 6. The UPC Code allows coupon descriptions using a main UPC Coupon Code and one of five types of Coupon Code extensions. UPC Coupon Code Format 1 is shown here.

UPC COUPON CODE FORMAT 1



The fourth partition is the Value Code. The Value Code is a 2-digit number related to the redemption value in dollars and cents. The mapping from these two digits into the savings is complex and depends on the amount, multiple purchase requirements or free merchandise offers. The final digit of the UPC Coupon Code is the modulo checksum.

The UPC Couple Code includes five types of extensions. The extensions have variable length and structure. The UPC Coupon Code Format 1, shown in the figure, has three partitions. The first is a 4-digit Application Identifier. In the example, the '8100' indicates a 6-digit number follows.

The first position of the number is the first digit of the UCC Company Code from the company issuing the coupon. The remaining five digits make up the Offer Code. This 5-digit code refers to the geographic distribution of coupons or to a distribution medium or some other definition.
The supplemental codes range in size from 6 digits to as many as 24. They include information on the offer, expiration date and household identification.

3.1.5. Supplemental Codes

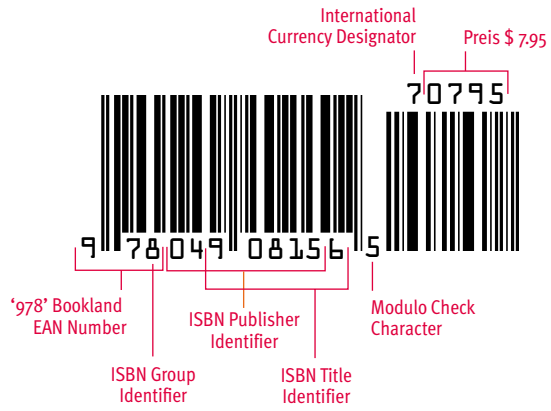
In addition to the main code, both the UPC Version A and E may add a 2 digit or a 5 digit supplemental code. Together with the 12-digit UPC code, these supplemental extensions are used for books, magazines, greeting cards and audio/video products.

A mass-market paperback book sold in North America has an identification number composed of the Bookland EAN Number and a 5-digit price code, as shown in Figure 7. The Bookland EAN number is divided into three main partitions, a '978' indicating a EAN prefix to the International Standard Book Number (ISBN) (see Appendix A), the ISBN Number, and a checksum digit.

The ISBN number contains three partitions. The first partition is a single digit (or digits) representing the Group Identifier, which may be either a language group ('0' or '1' for English-speaking country) or a country designator (for example a '2' for France). The second partition is a variable length publisher identifier, which may be 2 to 7 digits. The third section is a variable length title identifier, which may be 1 to 6 digits, and when combined with the group and publisher identifier are a fixed 9 digits.

Figure 7. Mass-market paperback books sold in North America contain a supplemental bar code which records the International Standard Book Number (ISBN), was well as the currency and price.

MASS-MARKET PAPERBACK BOOKS



The Price Code is divided into two sections. The first is a single digit International Currency Designator, indicating the type of currency applicable to the price code. For North America a '5' represents U.S. dollars and a '6' Canadian dollars. The last 4 digits are price, allowing books ranging from $0 to $99.99.

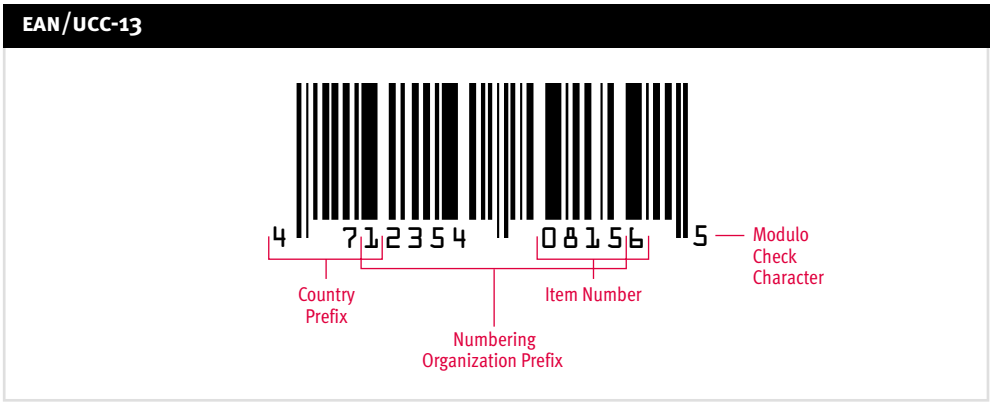
In addition to mass-market paperback books for North American markets, there are EAN/UCC codes with extensions, for hard cover books, magazines, greeting cards and audio/video products. Each product type has a slightly different partitioning scheme. Most schemes include information on price and title.

3.2. EAN/UCC-13

Growth of the North American UPC code lead to the development of a European standard in 1976. Aside from the extra digit and the means for encoding country of origin, the EAN Code is nearly identical to the original UPC standard. The Japanese Article Numbering (JAN) System and the International Article Number System (IAN) are also identical to the EAN/UCC-13.

The EAN/UCC-13 is a 13-digit code, which contains four partitions, as shown in Figure 8. The first partition is the UCC/EAN Numbering Organization (NO) Prefix containing 2 or 3 characters. The UCC/EAN NO Prefix represents the country code of the UCC/EAN Numbering Organization assigning the company number. For compatibility with the UPC, the prefixes 00, 01, 03, 04 and 06 through 13 are assigned to the United States and Canada. A complete list of country codes is given in Appendix A.

Figure 8. The EAN/UCC-13 is a 13-digit code used to identify product in sold in countries outside the United States and Canada.



The second partition is the unique manufacturer number and is usually composed of 4 or 5 characters, but may have as many as 8 characters. For most EAN/UCC-13 numbers, the UCC/EAN NO Prefix together with the unique manufacture identification total 7 digits and is designated the UCC/EAN Company Number. Thus if the country code has 3 digits, for example Taiwan '471,' the company identification number has 4 digits.

With 7 digits, the EAN/UCC-13 could, theoretically, represent as many as 10,000,000 unique companies worldwide. However, the partitioning of Company Number into country and manufacture reduce this theoretic maximum.

The third partition is a 5-digit number representing the item identification – identical to the UPC standard. This allows each manufacture 100,000 unique product types. Finally, the last partition is a single digit encoding the modulo check sum for reader validation.

3.3. EAN/UCC-14

The EAN/UCC-14, previously known as the Shipping Container Symbol (SCC-14), is a 14-digit code used to mark cartons, cases or pallets containing products which have UPC or EAN identification numbers. The EAN/UCC-14 identifies what product is contained within the shipping container. The bar code symbol used to encode the EAN/UCC-14 is the Interleaved 2-of-5 (ITF-14), as shown in Figure 9.

The EAN/UCC-14 contains four partitions. The first partition is the Package Indicator (PI), which represents the level of packaging. Level '0,' for example is used for products at the unit level (i.e. the lowest level of packaging). Numbers '1' through '8' indicate different packing configurations. A '9' indicates that the amount of product varies from container to container.

The second partition is a single digit representing the management organization. A '0' indicates the structure is managed by UCC and a non-zero means the number is managed by the EAN. Essentially, the remaining partitions match the UCC-12 or EAN/UCC-13 product codes indicating the contents of the package.

Figure 9. The EAN/UCC-14, formally the Shipping Container Code (SCC-14), is used to mark cartons, cases or pallets that contain UPC or EAN name products.

EAN/UCC-14

The third partition are the manufacture identification number and the fourth the 5-digit item number. The last partition, as with other EAN/UCC numbers, is the modulo checksum character.

3.4. EAN/UCC-8

The EAN/UCC-8 is an abbreviated version of the EAN/UCC-13 and is used primarily outside the United States and Canada. The 8-digit code includes a 2 or 3 digit country code, 4 or 5 data digits and a single checksum, as shown in Figure 10. The combined data digits specify a particular product within a particular country. A 3-digit country code, such as Taiwan '471,' for example, would only allow 10,000 products.

Unlike the UPC-E symbol, which is a compressed version of the UPC-A, the EAN/UCC-8 directly identifies a particular manufacture and product. Therefore, there is no simple algorithm to map a EAN/UCC-8 into a EAN/UCC-13.

The EAN/UCC-8 may include a 2-digit or 5-digit extension, though the primary purpose of this code is to use as little space as possible.

Figure 10. The EAN/UCC-8 is an 8-digit bar code used primarily outside United States and Canada on products with limited printable space.

EAN/UCC-8

4. COMPARISON OF EPC AND GTIN

The UCC/EAN GTIN family and the EPC code are similar, but have fundamental differences in their history, scope and structure.

4.1. History

The family of codes, including the UPC, EAN/UCC-13, EAN/UCC-14, EAN-8 and their extensions, known collectively as the Global Trade Item Number (GTIN), grew over the years from the original supermarket checkout application in the United States to a worldwide system of supply chain identification and description. Because of its evolutionary growth, the GTIN system includes many modifications and extensions, which have effectively met the demands of targeted applications.

From its initial development in the early 1970s, the U.P.C. could not easily rely on computer networking and open database standards. Information had to be locally accessible. In the case of variable weight items and coupons, information had to be embedded directly in the code.

The success of the U.P.C. standard encouraged its spread beyond grocery retail, into pharmaceuticals, coupons, books, magazines, videos and shipments. The basic identification code was extended to include country, classification and product information.

The UCC/EAN data structures are also closely linked to their bar code implementations. The length of the codes, checksums and data structures are designed to be compatible with and facilitate scanning with a linear bar code reader.

The Electronic Product Code (EPC), on the other hand, has a much shorter history. The EPC grew explicitly from the desire to link physical objects to the Internet. In this sense, the EPC is much more closely related to the Internet Protocol (IP) Address, used to reference networked computers, than to the GTIN data structures.

The combination of the EPC with the other elements of our proposed ‘intelligent infrastruture,’ including electronic tags (eTags), the Object Name Service (ONS) and the Physical Markup Language (PML), create a system in which the function of the EAN/UCC architecture can be reproduced and enhanced [1-4].

4.2. Scope

The UCC/EAN system of numbering and bar-coding was designed to address specific needs in retail and distribution. Different applications had different requirements and demanded different data structures. The EPC, on the other hand, was intended to enumerate all physical objects, in the same way the IP address was designed to identify all networked computers. The Electronic Product Code attempts to organize and identify physical objects across industries, including manufacturing, healthcare, defense, electronics, services, etc., as well as the traditional retail and distribution. For this reason, the EPC code had to be broadly defined to address these areas.

4.3. Structure

The EPC was designed explicitly not to carry information about the identified object, but to provide an efficient reference to networked information. For this reason, the EPC has an inherently simple data structure. We need only identify the organization and the host computer server, which contains the reference information.

The UCC/EAN system could not rely on networked resources. In many cases, the system had to work effectively without any external connection, and at times, without any computer system at all. Explicit data, such as price and expiration date – particularly for variable items – had to be stored within the code structure.

5. INTEGRATION OF EPC AND GTIN

Over 800,000 member companies in over 90 countries use the UCC/EAN Numbering System. Billions of items use bar codes in the GTIN family of formats, in what is arguably one of the most successful standards ever developed.

With this level of investment, it is of no surprise we wish to provide methods to map elements of the globally accepted UCC/EAN identification systems into the newer Electronic Product Code. Although it may seem complex, the integration of these two approaches may be quite simple. The effective compatibility of the Global Trade Item Numbering family with the Electronic Product Code will allow more rapid adoption of the 'intelligent infrastructure' into the more traditional bar-code industries such as retail and distribution, as well as expand into newer fields of global standardization, including healthcare and manufacturing.

Most of the data elements of the UCC/EAN structures (e.g. shipping data, coupons, expiration dates, price, country, etc.) we propose to move to networked databases. The checksum characters, needed for bar-code scanning, would be removed from the data structure.

The EAN/UCC Company and Product Numbers of the Global Trade Item Number Family would remain elements of the Electronic Product Code in the form of the EPC Manager Number and the EPC Product Type. In the next sections, we will describe the mapping from these GTIN elements to the corresponding components of the EPC.

5.1. Manager Identification

The Regular UPC Code allows five digits for manufacturer, providing 100,000 possible companies. This 5-digit number has no special cases, extensions and categorizations. Therefore the mapping from the UPC Company Number to EPC Manager Identification Number is simple – the numbers are identical.

The EAN-13 UCC/EAN Company Number, on the other hand, includes both a Country Code and a Manufacturer Identifier. The Country Code indicates the country, which manages and issues the UCC/EAN Manufacturer Identification Numbers. Although the UCC/EAN Company Number is a 7-digit number, which provides theoretically as many as 10,000,000 company numbers, the division of country and company codes, reduces this theoretical upper bound.

The 2 or 3-digit country codes theoretically enable as many as 1,000 individual countries (depending on the distribution of numbers and the allocation between the 2 or 3-digit codes). However, the actual distribution of country codes provides approximately 90 unique countries, see Appendix A.

A country with a 2-digit code may identify as many as 100,000 possible companies, and a 3-digit code, 10,000 companies. The segmentation and allocation of codes, reduces the practical upper bound on identifiable companies. There are 42 2-digit codes and 105 3-digit codes providing as many as 6,300,000 identifiable companies. (Note, a number of countries have more than one code. The United States and Canada have 12 2-digit codes and Germany 41 3-digit codes). The number of identifiable companies may be further reduced, since some countries may not fully use their allocation.

The Electronic Product Code attempts to limit the amount of inherent information and categorization within the code structure. The partitioning of companies by country (or company codes managed by country) are avoided. Similar to the Internet Protocol (IP) Number, which makes no distinction of country of origin, the EPC is globally oriented.

If we wish to identify a company independent of country, we must map the country code and company code of the EAN/UCC-13 into a single identification number. Perhaps the easiest method – and one which allows the greatest data storage efficiency – would be to add company numbers sequentially without regard to the country partition.

In other words, companies numbers for Estonia '472,' for example, would be listed immediately after Taiwan '471.' The list of EPC manager numbers would then correspond one-to-one with the GTIN Company Numbers. The approximately 800,000 EAN/UCC member companies would embed easily within the 268,435,456 manager identifiers allocated for the EPC. The result would be a simple, compact enumeration of globally unique manager identifiers.

All future company numbers could be issued via an automated registration system, similar to the domain name system used today. In this way, we could smoothly transition from the current GTIN company numbering system to a more globally defined EPC manager number

5.2. Product Number

Both the UCC-12 and EAN/UCC-13 Codes allow 5 digits for the product number. The mapping of these numbers into an EPC product type is straightt-forward. The GTIN product number and EPC product type would be identical.

5.3. Serial Number

Although the GTIN allows serialization as an option, the EPC intrinsically includes a serial number. Since serial numbers are managed by the individual company for each individual product type, the specification and allocation numbers are as they would be for GTIN serialization.

5.4. UCC/EAN Data Carrying Formats

The UCC/EAN GTIN specification, as outline in earlier sections, includes many formats which intrinsically carry data. These include formats for variable weight, coupons, publications, videos, shipments, etc. For example, the variable weight number system, identified by a '2' in the first digit of the UCC-12 code, stores

the packer, commodity type and price of the item. Since some items may have the same weight, this number is not unique and, therefore, cannot be mapped into a unique identification system, such as the EPC.

Because of the inherent architecture of the EPC system, all of product data – including weight, rebates, product names and shipments – would exist on the network. These data structures would therefore not carry over into the EPC format.

6. EXAMPLES

To illustrate the integration of the GTIN family with the EPC format, we present the following examples. The regular U.P.C. code, as shown in Figure 11, maps directly into an EPC number. The manufacturer and item number of the UCC-12 correspond to manager and object class numbers of the EPC. Note the decimal number of the U.P.C. is written in the hexadecimal notation of the EPC.

Only regular U.P.C. numbers can be mapped in this manner. The National Drug Codes and National Health Related Item Codes, identified by a '3' may also be mapped into an EPC number, if the FDA labeler and product/packaging codes are unique.

As previously described, data carrying UCC-12 formats, including random weights '2,' internal numbers '4' and coupons '5,' cannot correspond to a unique identifier. This information, however, is included within the EPC architecture, in that all product data is stored and retrieved via the Internet [1,2].

Figure 11. The manufacturer and item numbers of the U.P.C. code map directly into the manager and object class numbers of the EPC.

The EAN/UCC-13 code can also be mapped into a unique EPC number. In the example, shown in Figure 12, the EPC manager number is formed by adding an offset to the manufacturer identification number. The exact offset number system has not been determined, but will correspond to the EAN/UCC Country Code. Each EAN/UCC Country Code will have a unique offset, which is added to the manufacture identification number to yield a globally unique EPC manager number.

In the example, the offset for Taiwan '471' is arbitrary set to 900,000. The manufacturer identification number '2354' is thus added to the offset '900,000,' to yield the EPC manager number '902,354.'

The EAN/UCC-13 item number corresponds directly to the EPC product class. In this case, the item number '8156' corresponds to the EPC product class '8156,' or '1FDC' when written in hexadecimal notation.

Figure 12. Adding an EAN/UCC Country Code specific offset to the EAN/UCC Manufacturer Identification Number forms the EPC Manager Number. The EPC Product Class and EAN/UCC item number are identical.

The GTIN EAN/UCC-14 shipment numbers do not correspond to EPC codes, since these numbers are not unique. In addition, all shipment information, as well as all product data is stored on the computer network in the new EPC infrastructure.

The GTIN EAN/UCC-8 format would correspond to EPC codes in much the same way as the EAN/UCC-13. However, since we envision the use of electronic tags to store and transmit identification information, the reduced spaced symbology is not as necessary.

In any case, the first step would be to convert the EAN/UCC-8 to a unique EAN/UCC-13. Since there is no analytic correspondence, the product code number would have to be generated by the associated manufacturer.

In the example shown in Figure 13, the '4081' product code corresponds to the manufacturer identification number '2354' and product code '8156.' Thus, adding the offset for Taiwan '471' (again arbitrarily given as '900,000') yields an EPC manager number '902,354. The EPC product class is the same as the product code '8156.'

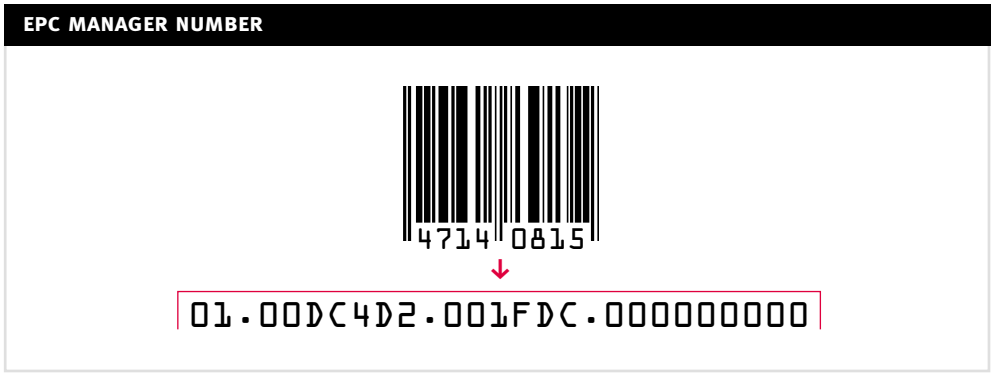


Figure 13. Adding an EAN/UCC Country Code specific offset to the EAN/UCC Manufacturer Identification Number forms the EPC Manager Number. The EPC Product Class and EAN/UCC item number are identical

7. GTIN AS EPC

The UCC/EAN GTIN family of data formats has been globally accepted as a means to identify and track items in retail and commercial distribution. It seems reasonable that the GTIN should be used as the reference to networked information in the same way as the Electronic Product Code.

There were, however, a number of issues, which make the direct use of the GTIN as the EPC less desirable. First, the GTIN represents not a single numbering scheme, but a family of codes, including U.P.C., EAN-13, SCC-14 and EAN-8.

Second, GTIN carries a number of extensions and exceptions. These include designations for weighed items, drugs, unregulated sequences, coupons and publications.

Third, the shipping variant of the GTIN, that is the Shipping Container Code – SCC-14, describes container contents. Although useful and necessary for labeled packaging, this differs from the intent of the EPC. The latter will use networked information to track shipments. Mixed loads and variable packing can be dynamically recorded and updated as the shipment moves through the supply chain.

Fourth, the GTIN is somewhat limited in its total address space. The EAN/UCC-13 component of GTIN allows theoretically as many as 10,000,000 companies and as many as 100,000 items per company. While this is significant, it is less than the envisioned application of the Electronic Product Code.

Fifth, GTIN includes checksum digits (one for most codes and two for some codes with extensions), which were designed for linear bar codes. The Electronic Product Code is simply an identification and enumeration scheme, independent of communication method.

Sixth, the GTIN, in its current state, is not easily extensible; that is, GTIN would require a header and version system as the proposed EPC. While the Application Identifiers (AI) may serve this function, this could be confused with their current use.

Seventh, the GTIN EAN/UCC-13 includes a country designation, which indicates the UCC/EAN member organization that manages the company numbers. The EPC, like the Internet Protocol (IP) address, does not include country identification. Information about the company including country of origin and all relevant business information would be available from on-line resources.

Finally, GTIN includes a number of information carrying elements in its specification. These include information on pricing, expiration date and packing. Information carrying elements were explicitly removed from the definition of the EPC. Information carrying elements were expressly removed from the EPC, since they prohibit unique identification of objects and prevent hierarchical resolution systems like the Object Name Service.

8. CONCLUSION

The proposed 'intelligent infrastructure,' including electronic tagging, the Electronic Product Code, the Object Name Service and the Physical Markup Language, offers the potential to revolutionize supply chain management and control [1,2]. As we move toward this new architecture, we must consider how to incorporate current methods and standards.

The Global Trade Item Number (GTIN) system is a globally accepted means of identifying and tracking products and shipments. The GTIN family of formats, however, includes many elements which address specific needs in the commercial and retail supply chain. In its current form, the GTIN does not provide a simple, unique and globally scalable method for object identification.

However, as we move toward a new architecture, we should not ignore current standards and methods. Therefore, we propose to use elements of the Global Trade Item Number within elements of the Electronic Product Code. Specifically, the UCC/EAN identification for company, product and serial number can map almost identically into the EPC manager, type and instance components. The Electronic Product Code – like the Internet Protocol Address – thus becomes a globally scalable method to identify and track objects. In this way current standards and methods are used efficiently to enable our proposed 'Intelligent World.'

9. REFERENCES

1. **Brock, D. L, "Intelligent Infrastructure – A Method for Networking Physical Objects,"**
   Presentation, MIT Smart World Conference, Apr 2000.
   http://auto-id.mit.edu/whatsnew/download/DB_Smart_World.pdf

2. **"The Networked Physical World – Proposal for Engineering the Next Generation of Computing, Commerce and Automatic-Identification,"**
   Auto-ID White Paper, WH-001, Dec 2000.
   http://auto-id.mit.edu/pdf/MIT-.AUTOID-WH-001.pdf

3. **Brock, D. L, "The Electronic Product Code – A Naming Scheme for Physical Objects,"**
   Auto-ID White Paper, WH-002, Jan 2001.
   http://auto-id.mit.edu/pdf/MIT-AUTOID-WH-002.pdf.

4. **Brock, D. L, "The Physical Markup Language – A Universal Language for Physical Objects,"**
   Auto-ID White Paper, WH-003, Feb 2001.
   http://auto-id.mit.edu/pdf/MIT-AUTOID-WH-003.pdf.

5. **The Uniform Code Council (UCC)**
   http://www.uc-council.org

6. **The European Article Numbering (EAN) International**
   http://www.ean.org

7. **Uniform Code Council (UCC), Uniform Product Code (U.P.C.) Symbol Specification Manual,**
   January 1986.
   http://www.uc-council.org/reflib/01302/d36-t.htm.

8. **Uniform Code Council (UCC), Reduced Space Symbology.**
   http://www.uc-council.org/rss14/.

9. **Guidelines for Supply Chain Identification, Uniform Code**
   Council Publication, August 1999.

APPENDIX A

EAN/UCC Member Organization Prefixes

00–13	USA & Canada
20–29	Reserved for local use (store/warehouse)
30–37	France
400–440	Germany
45	Japan
46	Russian Federation
471	Taiwan
474	Estonia
475	Latvia
477	Lithuania
479	Sri Lanka
480	Philippines
482	Ukraine
484	Moldova
485	Armenia
486	Georgia
487	Kazakhstan
489	Hong Kong
49	Japan
50	UK
520	Greece
528	Lebanon
529	Cyprus
531	Macedonia
535	Malta
539	Ireland
54	Belgium & Luxembourg
560	Portugal
569	Iceland
57	Danmark
590	Poland
59	Romania
599	Hungary
600–601	South Africa
609	Mauritius
611	Morocco
613	Algeria
619	Tunisia
622	Egypt
625	Jordan
626	Iran
64	Finland
690–692	China
70	Norway
729	Israel
73	Sweden
740–745	Guatemala, El Salvador, Honduras, Nicaragua, Costa Rica & Panama
746	Republica Dominicana
750	Mexico
759	Venezuela
76	Switzerland
770	Colombia
773	Uruguay
775	Peru
777	Bolivia
779	Argentina
780	Chile
784	Paraguay
785	Peru
786	Ecuador
789	Brazil
80–83	Italy
84	Spain
850	Cuba
858	Slovakia
859	Czech
860	Yugoslavia
869	Turkey
87	Netherlands
880	South Korea
885	Thailand
888	Singapore
890	India
893	Vietnam
899	Indonesia
90–91	Austria
93	Australia
94	New Zealand
955	Malaysia
977	ISSN (International Standard Serial Number for periodicals)
978	ISBN (International Standard Book Number)
979	ISMN (International Standard Music Number)
980	Refund receipts
99	Coupons

APPENDIX B – CONVERTING GTIN TO EPC

B.1. Introduction

The Electronic Product Code (EPC) is a simple product identifier that references networked data. Very little information is actually stored in the EPC – manufacturer (EPC manager), product type and serial number. As we have shown, the Global Trade Item Numbering (GTIN) is family of codes that identify the product, but may also encode other information, such as shipment data, price, currency, coupon, weight and product type.

Since the Global Trade Item Number family includes information **not** part of the Electronic Product Code, it is not possible to translate the entire contents of the GTIN into the EPC. The product identification portion of the GTIN, however, **does** map into the EPC, but any additional data must be stored remotely in a Physical Markup Language (PML) data file. In this appendix, we will show how to convert a GTIN number directly into an EPC and PML data.

B.2. Converting Regular U.P.C. to EPC

The regular U.P.C. includes a numbering system character, manufacturer identification, item number and check character. To convert a regular U.P.C. into an EPC, the U.P.C. manufacturer identification **is** the EPC manager number and the item number **is** the object class, as shown in Figure B1. Notice, in the figure the manufacture and item number are shown in decimal notional, '02354' and '08156', for the U.P.C. and hexadecimal, '932' and '1FDC', for the EPC.

Finally, the U.P.C. numbering system character and check character are discarded. Since the regular U.P.C. does not include a unique item identification serial number, this is defined in the EPC.

Figure B1. The manufacturer and item numbers of the U.P.C. code map directly into the manager and object class numbers of the EPC.

B.3 Converting Other U.P.C. codes to EPC

The U.P.C. can also store random weight, drug codes, internal company codes, coupons and supplemental codes. Except for the drug code, which converts an EPC in the same manner as a regular U.P.C., the other data types translate into Physical Markup Language (PML) files. In other words, all data content – except for product identity – is stored a PML file.

The random weight item, for example, translates into PML data as shown in Figure B2. The price, $7.56, is encoded as a **price** element in the PML file[1]. Similarly data stored on coupons and supplemental codes will also be translated and stored in PML files. In fact, more detailed information – or even algorithms – may eventually be encoded in PML. This would allow stores, for example, to tailor rebates and discounts based on buying history.

[1] We must note, however, the final PML specification has not been completed as of the date of this writing. Thus the semantics of the elements and attributes may change, but basic approach and the ability to store GTIN data will remain the same.

GTIN DATA STORED IN PML

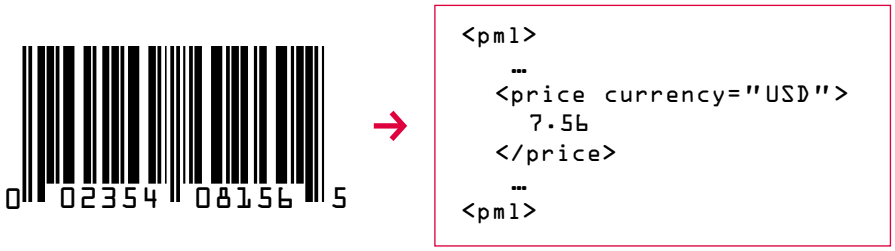


Figure B2. Data in the GTIN will be stored in Physical Markup Language (PML) data files. Here the $7.56 price for a random weight item is recorded as a value in the PML price element.

B.4. Converting EAN-13 to EPC

Converting an EAN-13 number directly into an EPC is somewhat more involved than the regular U.P.C.. The EAN-13 number includes a 2 to 3 digit Numbering Organization Prefix (Country Code) and a 4 to 8 digit manufacturer code. Together the Numbering Organization Prefix and Manufacturer Code form a 6 to 11 digit UCC.EAN Company Number. Allowing the maximum partition size, the EAN-13 **theoretically** provides up to 100 **billion** companies. Given that the current membership in the UCC.EAN is less than 1 million, an allocation this large is unnecessary.

One approach to converting the EAN-13 directly in EPC is to add a multiple of the Numbering Organization Prefix to the Manufacturer Code to yield a **unique** EPC Manager Number, as illustrated in Figure B3. Alternatively, the UCC.EAN Company Numbers could be reassigned **once** to an EPC Manager Number. Thus we provide a one-to-one mapping between UCC.EAN Company Numbers and EPC Manager codes.

As with the regular U.P.C., the EAN-13 Item Number is the EPC Object Class. Finally, the EAN-13 checksum digit is discarded in the conversion to EPC.

CREATING A UNIQUE EPC MANAGER NUMBER

Figure B3. Adding an EAN/UCC Country Code specific offset to the EAN/UCC Manufacturer Identification Number forms the EPC Manager Number. The EPC Product Class and EAN/UCC item number are identical.

B.5. Converting SCC-14 to EPC

The Shipping Container Code (SCC-14), or the EAN/UCC-14, is a case, pallet or package label code that provides some information on the container contents. When using EPC, information on contents, configuration and arrangement are represented in a PML file.

The EAN/UCC-14 is not "converted" to EPC or PML per-se. Rather each container is assigned a unique EPC code, and this EPC references a PML file, which contains the information that would have been stored on the EAN/UCC-14 code.

Figure B4, for example, shows an EAN/UCC-14 container code signifying uniform contents with ‘0 14141 54321 2’ U.P.C. codes. The corresponding PML file contains a node element representing the container’s EPC code (not represented in the EAN/UCC-14) and a list of other node elements signifying each of the individual items. Here we can see mixed loads and individual product tracking are made simple using the EPC and PML system.

Figure B4. The EAN/UCC-14, formally the Shipping Container Code (SCC-14), is used to mark cartons, cases or pallets that contain UPC or EAN name products. Information on shipment contents and configuration are no longer contained in the container identification system, but in the associated Physical Markup Language (PML) data file. Here the contents of the pallet and containers are stored in a “node” structure, which is described more fully in the PML specification [4].

EAN/UCC-14 INCLUDED WITHIN THE PML FILE PLUS OTHER ELEMENTS


```
pml>
  …
  <node label="Container" ePC="01.xxxxxxx.xxxxxx.xxxxxxxxx">
     …
     <node ePC="01.000373D.00D431.xxxxxxxxx">
     </node>
     …
  </node>
</pml>
```

B.6. Converting EAN-8 to EPC

The EAN/UCC-8, formerly the EAN-8, is an abbreviated version of the EAN/UCC-13. Therefore to convert an EAN/UCC-8 into an EPC, first convert the 8-digit code to the corresponding EAN/UCC-13, as shown in Figure B5, and then convert the 13-digit code to an EPC as described in section B3.

Figure B5. To convert an EAN/UCC-8 into an EPC, first convert the 8-digit code to the corresponding EAN/UCC-13 and then convert the 13-digit code to an EPC.

CONVERTING AN EAN/UCC-8 INTO EPS