The General Ontology Evaluation Framework (GOEF)
& the I-Choose Use Case

1. Challenges of Ontology Evaluation:

Use case driven ontology evaluation is managed through direct inspection by subject matter experts. However, this is a time-consuming effort, which requires individual review of potentially multiple ontologies. Thus, what if we could develop a system which could take in a use case formalism, and give recommendations for ontologies to use?

2. The General Ontology Evaluation Framework (GOEF)

Evolve towards science and engineering discipline for ontology

- Goal: Enable objective evaluation of an **ontology** with respect to a **use case**.
- Both are constructed / deconstructed to extract / expose the evaluation criteria and the ontology-encoded knowledge.
- Facilitates ontology design, modular construction, development management, and evaluation is built into the development process.

Create procedures, processes, methods to help define, adjudicate, and ensure quality of knowledge capture/representation

3. Stages of the GOEF Approach

The GOEF approach consists of two stages:

1) Recast use case into its components:
    - Functional objective
    - Design objective and requirements specification
    - Semantic components required to achieve above

2) Evaluate components using objective metrics
    - Place existing evaluation methods in context by utility

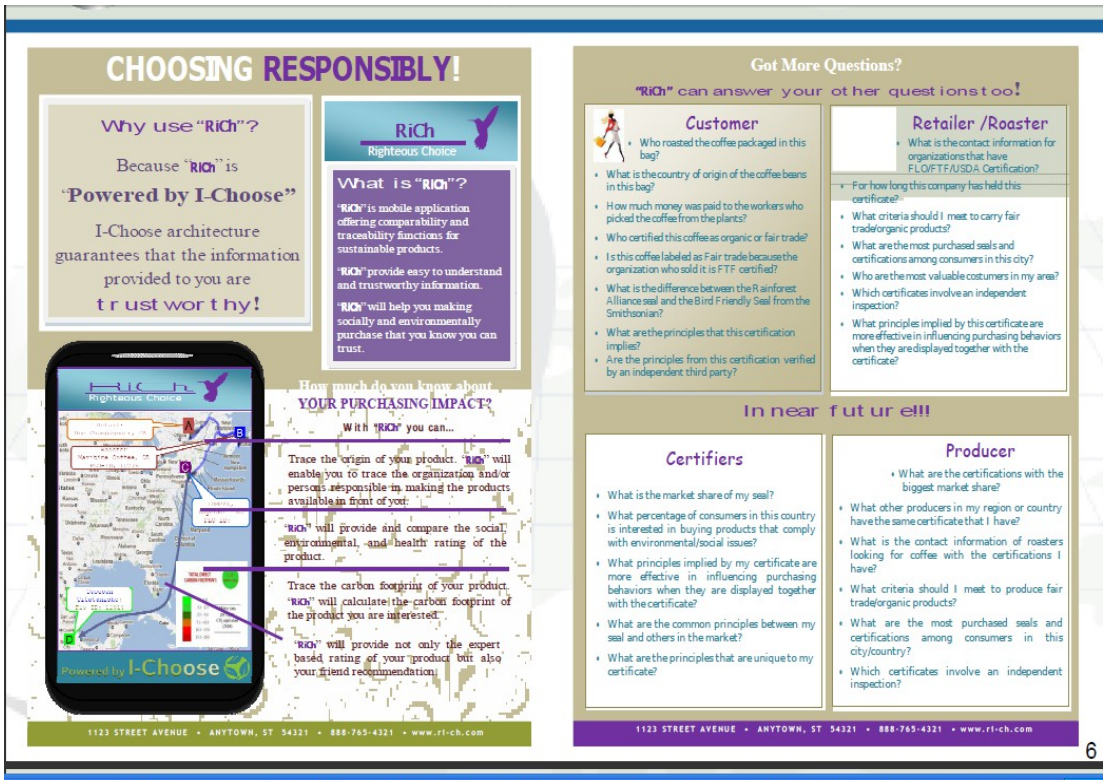
4. Motivating Example - I-Choose:

a. Short description of I-Choose:

The objective of I-Choose project is to develop an interoperable data framework to provide consumers with a wide range of information about how, where, and by whom products are manufactured and brought to market. More specifically, I-Choose ontology based architecture should enable consumer advocates to retrieve and verify information acquired from social and environmental certification procedures, such as inspections undertaken to acquire FairTrade and Organic certifications. The I-Choose network and

ontology team have decided to focus first on an ontology based on FairTrade coffee certification. This preliminary ontology would allow for a prototype application to be developed and tested with an actual (or artificial) dataset. Ultimately, the I-Choose project is interested in developing a data standard for this sort of inspection/certification.

b. Envision Mock-up Application of I-Choose



The image above shows a mock up application leveraged on a proposed I-Choose architecture. In this application, a user would be able to retrieve "extra" information about a product not provided in its package. The I-Choose project had thus set out to map the entire domain of coffee production, distribution, and consumption. This would have enabled the consumer or consumer advocate to retrieve a number of data points on characteristics of the particular product he/she would be interested in purchasing. To enable this envisioned application, I-Choose architecture would need to gather data from 3rd party certifiers and supply chain companies.

Preliminary research by I-Choose ontology team has shifted the focus of the I-Choose project to target data obtained through inspection/certification processes for sustainable certified coffee as a starting point.

c. **Overview of Sustainability Certification Schemes**

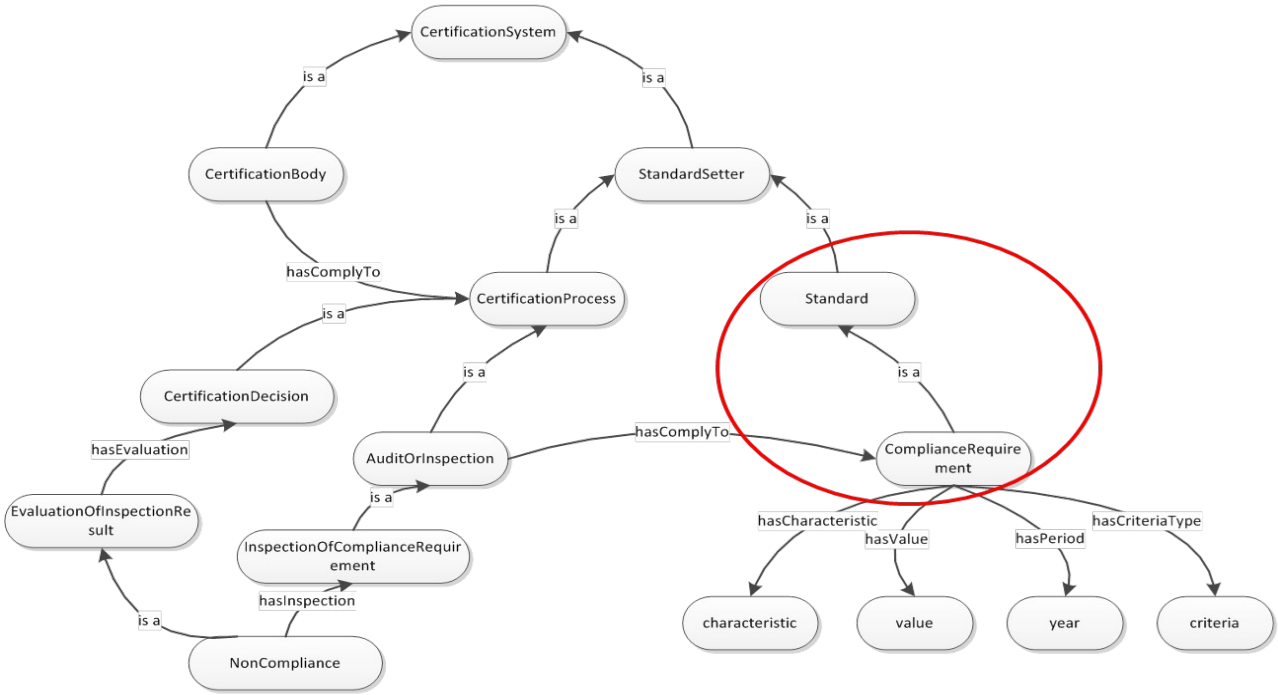
A certification system generally consists of a *standard setter* and a *certification body*. One of the fair trade standard setters is Fairtrade International (FLO). FLO creates standards and manages the labeling initiative.

The standard setter relies on an independent certification body to verify compliance to the standard. The certification body for FLO is Flo-cert—an internal body but independent to FLO. Flo-cert interprets the standard into verifiable control points called *compliance criteria*. Other sustainable certification schemes may use independent an external certification body (such as: UTZ Good Inside) or use a non-independent internal certification body (such as: Rainforest Action Network/RAN).

An applicant wishing to receive the Fairtrade (FLO) certification will be evaluated against the compliance criteria by an inspector/auditor appointed by Flo-cert (*figure 1 below*). Inspector/Auditor will follow each compliance criteria specified by the certification body—applicable to the entity being evaluated (e.g. small producer organization). The result from the audit/inspection will be used for the ultimate certification decision which, in this case, will grant the particular Fairtrade International certification mark/label.

In this GOEF-I-Choose use case, we focus specifically on the standard and compliance criteria (*area in red circle in figure 1 below*).

Figure 1. General Overview of Certification System



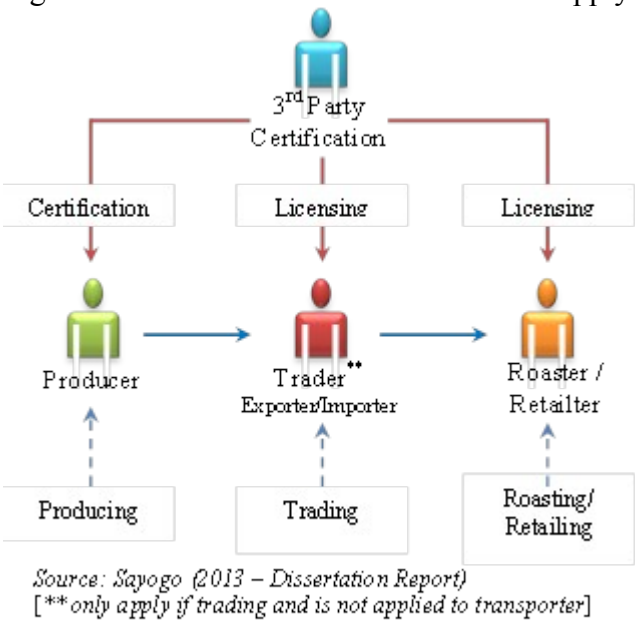
d. **Connecting Consumers to Fairtrade Certification Result**

d.1 Overview of the Chain

Referring to point 4.b, I-Choose aims to support/facilitate the generation of an application that can link consumer advocate (power user) to certification results, as a starting point. Using I-Choose architecture, consumer advocate will be able to retrieve and verify specific criteria evaluations related to the topic of interest.

To provide a brief overview on connecting consumers to fairtrade certification results, figure 2 below illustrates direct Fairtrade chain which is part of Fairtrade policy to eliminate the intermediary organizations. The red arrow indicates the applicability of certification. Each company in the chain will be assigned unique ID which will be available in every documentations exchange between companies and with the certification.

Figure 2. Direct Sustainable Certification Supply Chain.



Source: Sayogo (2013 – Dissertation Report)
*[** only apply if trading and is not applied to transporter]*

d.2 Data Provenance

Within the certified coffee supply chain, certification information is extracted from each company during the certification process. These are information pertinent to the compliance of certification criteria. This information is all kept in the database of the 3rd party certification. In the case of Fairtrade (FLO), this data is kept in Flo-cert database. In current condition, this data is proprietary and owned by each company certified by the certification. The release of this data depends on the consent of the company. In relation, I-Choose will use a mock-up data of the certification results pertinent to Child Labor as the first iteration in developing the architecture. Subsequently, the architecture will be tested against the actual data derived from the certification. Below is the example of the mock-up data:

AUDIT

FLO-ID	YEAR	AUDIT NO.	AUDIT_DATE	AUDIT_TYPE	SERVICE	AUDITOR	CERT_OFFICER
24525	2010	AO-00158	30/03/2010	Initial Audit	Producer-HL-Single Plantation	Nyagoy-Kenya	Janssen Martina

AUDIT DETAILS

AUDIT NO.	CC NO.	COMPLIANCE CRITER	NC_AUDIT	STATUS_NC	REASON_FOR_CHANGE	NC_CERT	CM_AUDIT
AO-00158	1.2.1.2.1	More than 50% of the n	less than 50%				
	1.2.1.2.2	Small producers are un					
	3.3.3.01	children below 15	indication of child				
	3.3.3.02	non-suitable work cond	indication of non-				

CC No.	:	The number of the checkpoint of the checklist used by the auditor
COMPLIANCE CRITERIA	:	The text of the checkpoint of the checklist used by the auditor
NC_AUDIT	:	the non-conformity as documented by the auditor during the audit
STATUS_NC	:	indicator whether the text of the non-conformity was change after the audit by the certification analyst

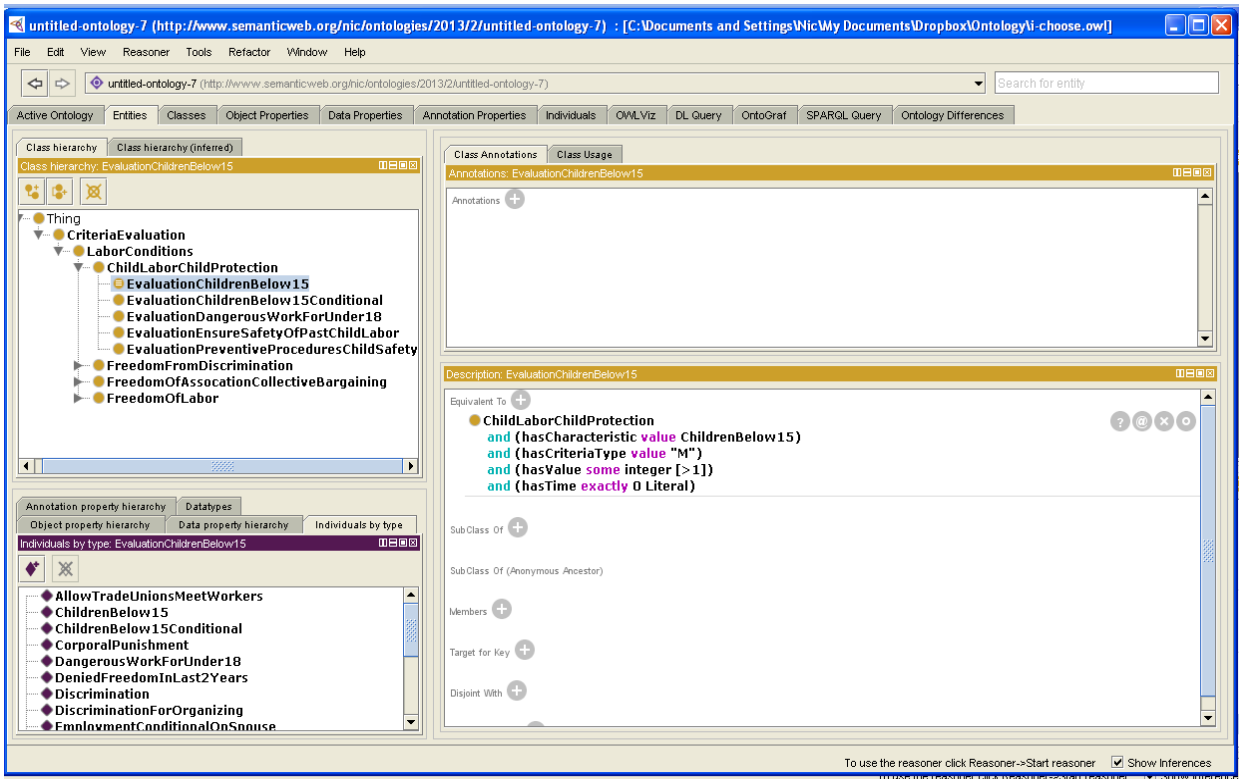
Source: Sayogo (2013) – Dissertation Report – adapted from Flo-cert

In the context of I-Choose, data provenance can expose relations between consumer products and their supply chains to consumer advocates. From an encoding perspective, there are two requirements for provenance records: (i) the presence of lineage for the consumer product, which can be thought of as encodings for the supply chain itself, and (ii) semantics for individual members of the lineage.

Based on the above supply chain, a provenance record in I-Choose could be expected to mention retailers, roasters, traders and producers in the lineage. Furthermore, it would be reasonable to expect details of which standards these individual members are in compliance with to be expressed in the corresponding semantics.

At a more detailed level, determination of what should go into a provenance record is an iterative process, driven by the refinement of both system use cases and user needs analysis. A number of encoding standards for provenance are presently in development, such as the W3C PROV vocabulary (http://www.w3.org/2011/prov/wiki/Main_Page).

e. **Preliminary Lower Level Ontology of Criteria Evaluation**



This partial, preliminary lower level ontology defines the class of "CriteriaEvaluation". Child classes of the CriteriaEvaluation class are the specific evaluations of each compliance criterion. In this ontology each criterion evaluation is modeled as a class. We believe the domain of certification/inspection is similar to pollution evaluation, and thus began to model our ontology after the Tetherless World Semanteco Ontology (http://tw.rpi.edu/web/project/SemantEco).

As noted above (4.d.2) each compliance criterion evaluation has a number of characteristics (data elements) associated with it.

For further details on the use case, and the elements associated with granting a particular certification see section f. below.

f. The I-Choose Use Case

Use Case Name:

Customer Advocate verifying compliance to Fairtrade certification criteria

Scenario *- Retrieving Workers Welfare and Child Labor Evaluation*

Goal:

A customer might want to know about a firm's compliance to specific Fairtrade certification criteria. In this scenario the criteria of interest are workers welfare and/or child labor. As result, a consumer advocate needs to be able to verify the certification criteria evaluations by retrieving specific criteria evaluations pertinent to the topic of interest.

Overview:

Consumer advocate (CA) needs to verify the compliance of producer organization (PO) "XX" to Fairtrade criteria pertinent to workers welfare and no child labor practices. The reason being, there are several queries from consumers who are asking for verification of the workers welfare and child labor practice by PO "XX". In relation, CA needs to retrieve the Fairtrade certification results of PO "XX" including the score on criteria for workers' welfare and no child labor practices. To do so, an understanding of how Fairtrade compliance criteria work is critical. Including the role of each instance in the compliance criteria contribute to the certification decision as well as the actors involved in the certification process.

Actors involved in certification decision:

Standard Setter Organization	-	Fair Trade International.
Certification Body	-	FLO-Cert
Inspector/Auditor	-	named_auditor
Small Farmer Organizations	-	named_coop

Mechanism in compliance criteria:

There are 6 columns representing different information in Fairtrade compliance criteria (see figure 3 below). The description of each column is as follows:

a. ***FLO Standard*** – this column indicate the connection or non-connection of the compliance criteria to the source FLO standard. Certification body (Flo-cert) translates the standard into compliance criteria. Thus, compliance criteria correspond to specific standard. The connection of compliance criteria into FLO standard is represented by the number of section of the FLO standard. For instance, compliance criterion of "not employed children below the age of 15" correspond to FLO standard sub-section no. 3.3.7. However, Flo-cert is also allowed to add non-core requirements into the compliance criteria which are not correspond to the standard

b. ***Applicable for*** – this column represents to whom the standard/criteria applied for. There are two level of applicability, a) organization level and b) product types.

   a) There are three organization levels (1st grade, 2nd grade, 3rd grade) in which 3rd grade is combination of 2nd grade and 2nd grade combination of 1st grade. 1st grade represents Small Producers Organization (SPO) which is an organization consisting of small farmers/producers.
   b) Fairtrade is applicable to at least 17 different products, including coffee, tea, sugar, cocoa, nuts, honey and other. Some standard/compliance criteria are only applicable to certain products or product types

c. ***Compliance criteria*** – indicate the translation of standard into verifiable control points. In this sense, one particular standard (e.g. no child labor) could be translated into several criteria (e.g. no children below 15; no children below 18 in dangerous work). On the other hand, each criterion correlates only to one particular standard.

d. ***Rank*** – this column represents the ranks for measuring the level of compliance to the criteria. In Flo-Cert system there are five ranks in sequence, each signifying a different level of compliance to the criterion. This column correlates with time of certification and criteria types. The measurement schemes of the criteria differ according to the criteria types and time of certification. For criteria with criteria type "core (C)", it will be measured by pass/fail. Considering there are five ranks, the threshold for assigning pass/fail decision depends at least on rank 3. If applicant cannot comply with rank 3, they will fail. On the other hand, for criteria type "Development (D)", measurement is based on the average of 3. Meaning the applicant has to comply with rank 3 in average.

e. ***Time*** – indicates time in the audit cycle. The audit cycle for Fairtrade certification is 6-years and four audits are conducted during this cycle. These four audits are for initial (0), year one (1), year three (3) and year six (6). Time also corresponds to the criteria type. Time 0 and 1 always correlate with criteria "C". Time 3 correlate with both criteria "C" and "D". Time 6 always correlate with criteria "D".

f. ***Criteria Type*** – there are two categories for criteria type, a) criteria C or D and b) criteria to indicate the applicability of the compliance. There are two categories for the second types, which are O (indicating that the criteria applied to organization only) and M (indicating that the criteria applied to member of organization only, qua farmers).

Figure 3. Example of Fairtrade (FLO) certification standard

FLO Standard [FLO e.V.] or other	Applicable for:	CC No.	FLO-CERT Compliance Criteria	Rank 1	Rank 2	Rank 3	Rank 4	Rank 5	Time	Criteria Type
		3.3.3	Child labour and child protection							
3.3.7	1st grade, 2nd grade, 3rd grade Not applicable for: Shrimps, Timber,	3.3.3.01	You and the members of your organisation have not employed children below the age of 15.	There are indications of this being common amongst members as employers or at organisation as employer.	There are indications of this happening on rare occasions but the organisation took no actions OR members hire labourers that come along with their children and work under parent's supervision.	The auditor found no indication (through workers' and employers interviews, and local knowledge) that this is present OR if there are indications of rare cases the organisation can demonstrate it took action.	There is a written policy approved by the General Assembly and made known to members.	= RANK 4 AND ICS that controls this at member level.	0	M

5. I-Choose under GOEF

Function:
Enable retrieval of specific criteria evaluations that occurred during an evaluation process of a particular product.

Design objective:
Initial system: Satisfy consensus user criteria pre-determined by survey research

Semantic components:

Compliance Criteria for LaborConditions:
a) Freedom of Association
b) Freedom of Labor
c) Freedom from Discrimination
d) Child Labor & Child Protection
   1) No children under 15
   2) Under 18 not under harmful conditions
   3) Preventive measures child safety

Certification Body
a) Flo-Cert
b) Certified private inspectors
c) Small Farmers Organization (SPO).

Standard Setter
a) FairTrade International
b) UTZ Good Inside
c) Rainforest Action Network
d) Common Code of Conduct for Coffee (4C)

Product
a) Coffee
b) Sugar Cane
c) Fruit, and 14 other

Evaluation Metrics:

Correctness
• General logical/syntactical validation
• Are the right terms used (compliance criteria vs. guidelines vs. standards)
• Match information provided in the ontology to information consensus user wants (surveyed).

Completeness

• Calculate % coverage of minimum terms
• All “severe” pesticides listed (certain %)
• All pesticides prohibited by U.S. EPA. Listed

Utility
• Validate against known test sets
• Consumer Consensus Questions Satisfied

6. Final remarks

There are two issues presented here. One is on how to evaluate the I-Choose ontology. A method is presented to carry this out, namely the GOEF approach. The second issue, however, concerns the evaluation of this approach itself. We hope that the Hackathon Ontology Clinic will assist both the developers of the GOEF approach as well as the developers of the I-Choose ontology to improve the evaluation framework and/or assist in the evaluation of the I-Choose ontology.