



SINTAGMA: Dall'informazione alla conoscenza

di

Joaquín Bastos Amigo

Traduzione italiana, a cura di Giuseppe Prencipe, dell'articolo
"Sintagma, from Information to Knowledge"
pubblicato sul vol. III, No. 1, Febbraio 2002
della rivista online UPGrade, a cura del CEPIS

Riassunto

Sintagma è una piattaforma di gestione integrale dell'informazione. È in grado di completare l'intero ciclo della informazione in una soluzione software, sebbene abbia interfacce per utilizzare o per essere utilizzata con o da ogni altro sistema esterno. Ha ambienti di sviluppo, controllori generici interattivi, possibilità di generare pagine HTML dinamiche e di progettare strutture, ha un proprio motore di ricerca dati, ecc. Sintagma è basato su un principio diverso da quello che sovrintende alla maggior parte dei sistemi informatici di gestione delle informazioni. È basato su nodi e relazioni e segue un modello che può dirsi neurale o cognitivo.

Parole chiave: Knowledge Management, modelli cognitivi

1. Introduzione

La domanda da parte di enti e aziende di sempre maggiori e più efficaci informazioni indica chiaramente che c'è un divario oggi tra gli strumenti software a disposizione e ciò che ci si aspetta da loro.

La maggior parte delle soluzioni associate alla "new economy", o piuttosto alla "infonomia"¹ (come direbbe il mio amico Alfons Cornellá), ha un costo molto alto in tempi e risorse. La necessità di trattare grandi quantità di informazioni altamente strutturate o non strutturate rivela in modo chiaro l'inadeguatezza dei tradizionali sistemi bidimensionali (database, fogli di calcolo, elaboratori di testo, compilatori di indici, ecc.) a rispondere alle nuove esigenze. Tuttavia, l'alto livello di standardizzazione di questi sistemi e la loro ampia diffusione rendono difficile prendere in considerazione soluzioni basate su sistemi alternativi. Questa resistenza al cambiamento, tipica dei sistemi critici o strategici, diminuisce con l'avvento di nuove tecnologie e di nuovi standard (XLM/SGML).

Sintagma è uno di questi strumenti di nuova generazione che rompe con il modello bidimensionale. Offre uno strumento abbastanza semplice per creare complessi sistemi di gestione delle informazioni, in grado di rispondere alle più difficili sfide che nascono dallo sviluppo di soluzioni adeguate ai nuovi bisogni del mondo d'oggi.

¹ <<http://www.infonomia.com>>

2. Primo approccio al problema della conoscenza

Il primo tema che si presenta quando ci si avvicina al problema della gestione della conoscenza è vedere in che modo la conoscenza differisce dalla informazione. Non basta modificare l'oggetto dell'informazione (ciò di cui l'informazione parla), ma si deve anche sapere come trattare questo oggetto per poter fare un uso appropriato della conoscenza. Sintagma è stato creato come un sistema cognitivo di trattamento delle informazioni, esso cioè mira a elaborare le informazioni così come fa il nostro cervello e a volgerle poi in conoscenza. In sostanza, Sintagma non è stato creato con lo scopo, forse troppo pretenzioso, di gestire la conoscenza, ma esso è piuttosto il risultato finale dello studio del modello cognitivo proprio dell'uomo.

2.1 Interpretazione del modello di conoscenza dell'uomo

Da un punto di vista strutturale, la conoscenza può essere vista come il tessuto delle relazioni logiche create dalle informazioni. È questo che ci permette di ricavare dalle informazioni, relazioni, astrazioni, generalizzazioni o deduzioni. La struttura di Sintagma si basa su una serie di relazioni orizzontali e verticali che sviluppano le informazioni e vanno a formare "mappe di conoscenze" diverse per ogni caso.

Gli studi fatti precedentemente alla creazione di Sintagma ci convinsero sul fatto che i migliori alleati per l'analisi di questo settore fossero la semantica e la linguistica. Il fatto che la lingua è la forma più comune per trasmettere e memorizzare conoscenza, essendo essa più tangibile della conoscenza stessa, ci spinsero a indagare sulle modalità con cui il linguaggio conduce alla conoscenza. Il dizionario o l'enciclopedia, con l'ausilio della logica computazionale, erano il modello da imitare. Così la parola divenne l'unità di base dell'informazione e il suo significato il contenuto.

Si poté quindi parzialmente costruire una mappa piramidale (comunemente nota come "albero semantico") in cui ogni punto o nodo aveva un'etichetta arbitraria (la parola) e un gran numero di rimandi o relazioni che conducevano a quel nodo o che da esso nascevano. Dal momento che l'etichetta è arbitraria (varia da lingua a lingua), la conoscenza deve risiedere nelle relazioni create tra i nodi o tra i punti delle informazioni.

2.2 Relazioni di base

L'analisi delle relazioni ci permise di differenziare tra relazioni verticali e relazioni orizzontali. Quelle verticali si riferivano alla classe delle informazioni ("è un...") e quelle orizzontali alle caratteristiche secondarie ("è caratterizzato da...").

Per poter trasformare le informazioni in una qualche forma di conoscenza, si rese necessaria una ulteriore analisi che individuasse le condizioni a cui le relazioni dovevano rispondere. Prima di tutto si vide che esse dovevano essere bidimensionali (se A è in relazione con B si verifica necessariamente una relazione reciproca per cui B si rapporta ad A) e transitive (se A si rapporta a B e B a C ne viene di conseguenza che A si rapporta a C). A queste due caratteristiche si applicano sia il concetto di ereditarietà che quello di relazione multipla. Si prenda, ad esempio, il fatto che un gatto è coperto di pelo, perché è un felino e i felini sono mammiferi e tutti i mammiferi hanno questa caratteristica, si vede che il processo di ereditarietà ci evita il peso di dover immagazzinare nel cervello tutte le informazioni proprie del gruppo a cui il concetto appartiene (relazione verticale). Si ricorre poi ad una relazione orizzontale, ereditata verticalmente, per collegare due concetti (mammiferi – pelo) che sono in categorie diverse. Il concetto "pelo", che di per se stesso è un'entità, quando viene applicato al concetto "mammifero", ne diventa una caratteristica o un attributo.

L'analisi, se pur breve, dell'architettura di questo modello di conoscenza, ci porta ad esaminare ora come Sintagma usi questi elementi logici per offrire una soluzione semplice, ma assai efficace, al problema della gestione dell'informazione.

3. Il modello Sintagma

3.1 Il modello fisico

Le unità di base delle informazioni di Sintagma sono nodi e relazioni (figura 1). I nodi contengono le informazioni esplicite e sono anche parte della rete su cui si formano e si trovano le relazioni. Questo livello fisico è gestito dal sistema NEUROMAT, responsabile della gestione della persistenza di queste unità di base. La struttura del nodo è estremamente semplice: un nodo ha un nome e un contenuto. Sia il nome che il contenuto possono contenere l'informazione o anche un rimando o una relazione con un altro nodo. Il contenuto di un nodo può anche essere una formula (come avviene per la cella di un foglio elettronico). La posizione all'interno dell'intera mappa delle informazioni è rivelata dal rimando a un nodo "padre". A livello fisico, abbiamo tre tipi di nodi: il nodo propriamente detto, l'attributo e l'atomo.

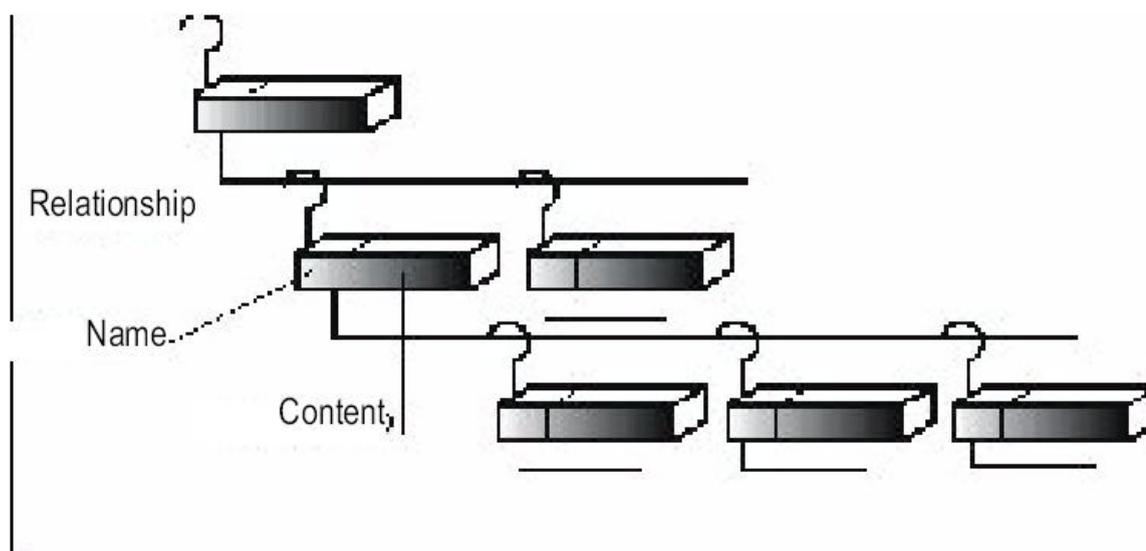


Figura 1 – Il modello Sintagma

Il nodo lo abbiamo già descritto. L'attributo ha le stesse caratteristiche del nodo, ma è dipendente da un altro nodo e ad esso appartiene (l'attributo ha un nome che ne indica le caratteristiche e un contenuto che ne specifica il valore). Infine c'è l'atomo che differisce dagli altri nodi non soltanto per il fatto che occupa una particolare posizione nell'ambito della struttura gerarchica, ma anche perché esso si trova, con proprie caratteristiche che lo identificano in maniera precisa, in uno specifico punto all'interno del contenuto del nodo che lo precede gerarchicamente (per esempio un collegamento ipertestuale nel mezzo di un testo). Ciascuno di questi nodi genera un diverso tipo di relazione dal suo predecessore (padre-figlio, nodo-attributo, nodo-atomo)

Vediamo ora come è ricavata da questa architettura fisica (livello NEUROMAT) la struttura logica di Sintagma.

3.2 Il modello logico

Tutte le informazioni trattate simultaneamente da Sintagma concorrono a formare un "universo" Sintagma. Ogni "universo" ha un nodo radice da cui dipendono gerarchicamente i seguenti nodi dipendenti: Metodo, Account, Applicazione, Programma, Database, Funzioni automatizzate (Auto), Deposito e Memoria.

I nodi dipendenti da altri nodi sono detti "discendenti" e vanno a formare un "ramo". Ogni nodo di un universo Sintagma ha una funzione diversa a seconda del ramo a cui appartiene.

Account: informazioni su utenti e permessi d'accesso.

Programmi: nodi eseguibili.

Database: il complesso delle informazioni (è il ramo più grande).

Applicazione: la struttura che consente di separare i dati dalla loro implementazione.

Funzioni automatizzate: informazioni diverse come utility per utenti, cataloghi di stile, filtri in entrata e in uscita, guida per un download "intelligente" di documenti, routine di avvio della applicazione, ecc.

Deposito: contenitore di riciclaggio per nodi cancellati.

Memoria: nodi non persistenti.

Metodo

Abbiamo intenzionalmente tralasciato di fare un accenno al Metodo, poiché esso richiede, per la sua importanza, una trattazione più ampia.

In Metodo vengono ripresi i rami di primo livello, ma, in questo caso, essi hanno il compito di definire comportamenti e caratteristiche.

Le caratteristiche sono, in sostanza, gli attributi che le future istanze dei nodi avranno. Se un attributo implica una relazione con un'altra area di dati, il suo nome sarà importato dal nodo originario di quei dati. Se si vogliono, cioè, assegnare le caratteristiche dei nodi del tipo Rapporti a quelli di Azienda, si associa, sul ramo Database, il nome dell'attributo Rapporti al nodo Azienda. Questa associazione dirà a Sintagma che c'è una relazione logica tra Rapporti e Azienda ed anche i valori che si assegneranno alle istanze del nodo attributo Rapporti saranno ereditati dal nodo Azienda. (Vedi Figura 2).

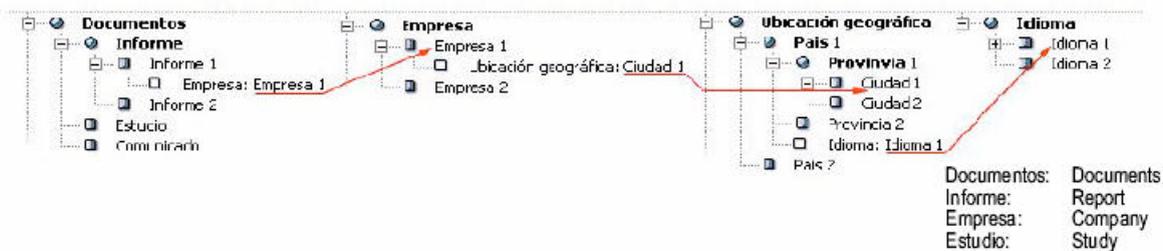


Figura 2 – Esempio di nodo Metodo

In aggiunta, nel definire i diversi tipi di nodi si può utilizzare il valore di default di un attributo. Così quando si renderà necessario dare un valore ad un attributo non classificato, il sistema gli assegnerà automaticamente il valore predefinito. E poiché Sintagma è in grado di definire il contenuto attraverso formule, il valore assegnato può anche essere il risultato di una formula applicata a quello specifico nodo. È così possibile definire l'attributo Indirizzo come la somma degli attributi Via, C.A.P., Città, Nazione.

Infine, si può assegnare ad un attributo la caratteristica della ereditarietà, nel senso che, quando si richiede un valore per un attributo non ancora classificato, il sistema salirà gerarchicamente nella ricerca fino al nodo più vicino con quel valore. Per esempio, se l'attributo Lingua del nodo del tipo Posizione geografica ha la capacità di trasmettere la propria caratteristica per via ereditaria, quando si assegna una lingua a livello di Nazione, questa sarà immediatamente attribuita a tutte le Province, Città, ecc. (con un tipo di procedimento che abbiamo visto nell'esempio di "gatto"). Nel caso che una Provincia usi un'altra lingua, questa sarebbe assegnata solo a livello di provincia e sarebbe di conseguenza applicata a tutti gli attributi discendenti di Provincia.

Se si estrapola un caso e si indica che la Lingua di default di una azienda è quella della sua Posizione geografica e, allo stesso tempo, si indica che per Rapporti la Lingua predefinita è quella dell'azienda che ha prodotto i rapporti, si vedrà che tutti i rapporti hanno una lingua di default, che è quella della stessa azienda, e che è, a sua volta, quella della Città, della Provincia e, infine, della Nazione. Nel caso di un rapporto o di un azienda con un attributo diverso dal normale valore, sarà necessario assegnare un valore specifico, ma non dovremmo farlo in tutti gli altri casi in cui si applica la regola generale.

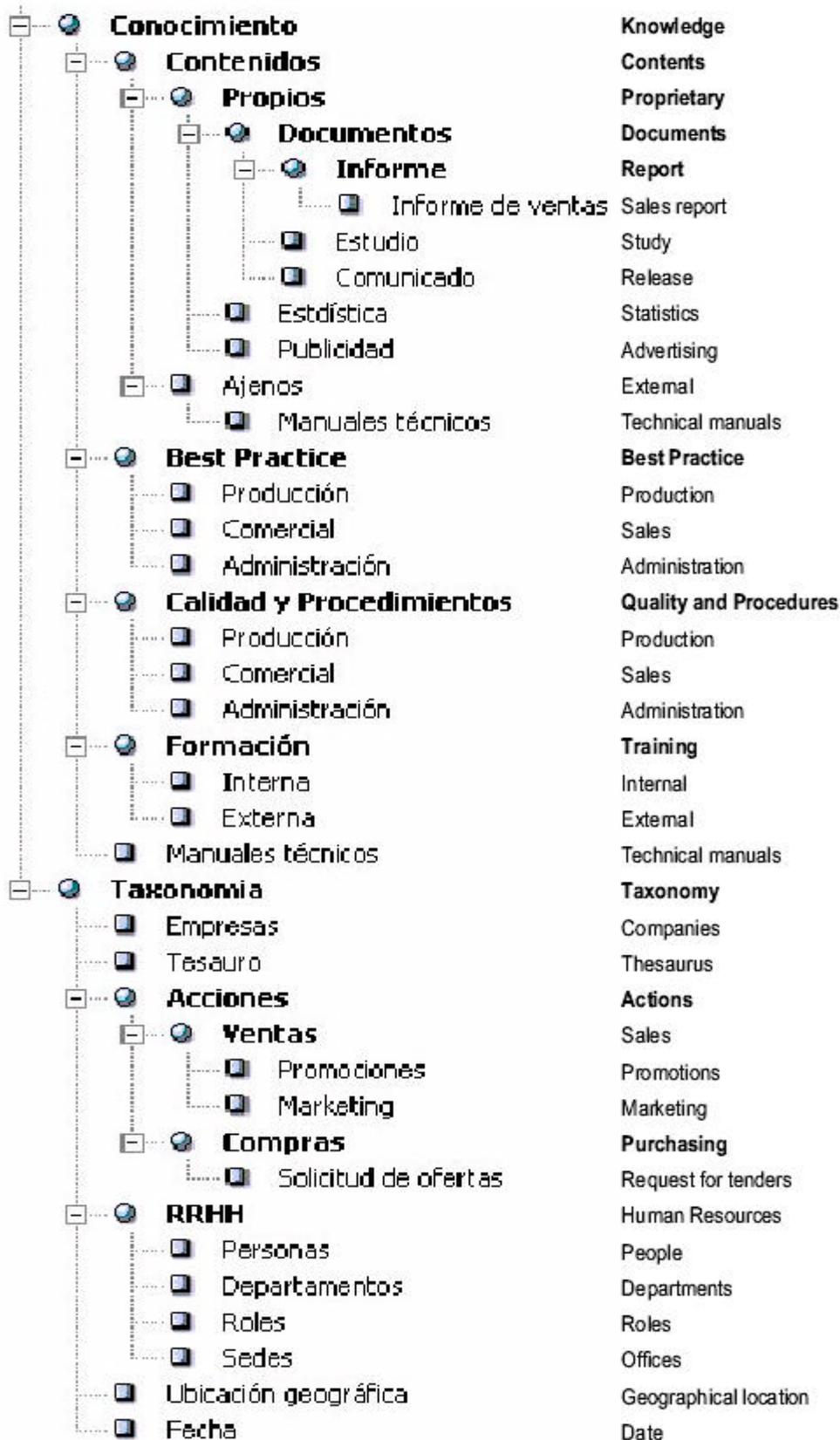


Figura 3 – Il ramo Database

Sotto questo aspetto Sintagma si comporta come un sistema cognitivo, dal momento che ricava deduttivamente dal tipo di nodo definito in Metodo le informazioni prevedibili o normalizzate.

Da ultimo dovrebbe essere evidente che in Sintagma ci sono dei nodi che si inseriscono in Metodo ed altri che non lo fanno per cui non ne vengono specificati uso e caratteristiche. Ciò rende in alcuni casi il sistema più potente.

Database

Come si è detto, la massa delle informazioni è posta nel ramo Database. I diversi livelli di classificazione delle informazioni sono strutturati in maniera discendente dall'alto in basso dove si trovano i dati e i contenuti. In un esempio semplice di gestione della conoscenza troveremmo un ramo Database con i seguenti nodi discendenti (figura 3).

Sebbene l'esempio sia forse troppo semplice, dà comunque un'idea di come Sintagma suddivide le informazioni in successive classificazioni con un numero illimitato di rami e livelli adattabili a ogni caso particolare. Ciò significa che le organizzazioni possono costruire l'albero o le suddivisioni come meglio richiede l'informazione che deve essere trattata. L'unica cosa che un'organizzazione deve avere chiara è stabilire quali rami conterranno le informazioni o le conoscenze che si richiederanno in seguito e quali rami saranno usati come criteri di ricerca (ma ciò non esclude che quei criteri siano essi stessi il punto di riferimento di ricerche o consultazioni). I primi saranno l'oggetto della ricerca, i secondi ne definiranno la tassonomia. In Sintagma è necessario assegnare attributi relazionali ai nodi oggetto che "puntano" ai nodi di tipo classificatorio. Ad esempio, i nodi del tipo Rapporti avranno nell'ambito del contenuto il testo dei rapporti stessi e gli attributi che si riferiscono ad Azienda, Autore, Data, Thesaurus, ecc. L'assegnazione di questi attributi nella forma di relazioni formerà un tessuto di conoscenze che potrà essere sfruttato in tutte le sua potenzialità.

4. L'utilizzo della conoscenza.

La possibilità di sfruttare al massimo livello le informazioni-conoscenze presenti in Sintagma è resa possibile dalla creazione delle interfacce utente con il linguaggio incorporato HTML (ma si può anche usare un altro tipo di linguaggio attraverso proprie API). Questo fa in modo che l'interfaccia utente sia del tutto adattabile ai bisogni dell'organizzazione. Ecco le azioni che si eseguono normalmente sulle interfacce dell'applicazione del Knowledge Management. Esse fanno parte degli strumenti di base di ricerca dei contenuti.

Ricerca principale: consente, partendo dai livelli più alti (Contenuti) e scendendo fino al livello di specificazione desiderato (Rapporti di vendita), di avvicinarsi nella maniera più approssimata possibile al tipo di informazione che si cerca.

Ricerca secondaria: consente, una volta selezionato il tipo di informazione e a partire dai risultati ottenuti, di perfezionare la nostra ricerca nell'ambito dei gruppi e sottogruppi delle caratteristiche collegate agli attributi (uso della tassonomia). Attraverso i criteri possiamo ampliare o restringere il campo della ricerca a piacimento.

Ricerca a tutto testo: consente di specificare la parola o il testo che deve apparire nel risultato.

Consultazione dei risultati

Si supponga, ad esempio, di ricercare rapporti di vendita che contengano il testo "prodotti a rischio" di aziende di Barcellona e ad iniziare dal mese di febbraio del 2000. Il sistema potrebbe rispondere che non ci sono risultati che soddisfano letteralmente le condizioni di ricerca, ma potrebbe anche indicare che ci sono riscontri geograficamente vicini (non in Barcellona, ma nelle altre città della Catalogna) o vicini nel tempo (non in febbraio, ma in marzo e aprile dello stesso anno).

La standardizzazione dei browser e l'ulteriore vantaggio di poter navigare attraverso gli alberi di classificazione (evitando così di dover porre condizioni) rendono l'interfaccia operativa del Knowledge Management di Sintagma molto ben accolta dagli utenti delle organizzazioni. Un dato, di certo, di fondamentale importanza in questo tipo di applicazioni .

5. Conclusioni

Sintagma consente all'utente di creare tutti gli "oggetti di conoscenza" che vuole, secondo sottolivelli e classificazioni necessari. Gli permette di creare a piacimento, sempre attraverso sottolivelli e un sistema di gestione del flusso di lavoro che tiene traccia del percorso dei documenti, criteri alternativi di classificazione. L'interfaccia dell'applicazione sarà esattamente ciò che l'organizzazione richiede, poiché le uniche limitazioni sono quelle imposte dagli utenti stessi. Infine, e questa è la funzionalità più interessante, consente agli utenti di modificare o aggiungere caratteristiche e di cambiare (con la funzione "drag and drop") la posizione del nodo con i corrispondenti nodi discendenti lasciando inalterata la coerenza del sistema. Queste prerogative fanno di Sintagma una ideale piattaforma per soluzioni su misura nella gestione delle conoscenze.

Diversamente da quanto avviene per altre soluzioni software in cui l'oggetto è semplicemente modificato dall'applicazione, Sintagma considera la conoscenza come un elemento di una struttura o come il risultato di uno specifico trattamento delle informazioni. È questo originale concetto che consente all'applicazione di risolvere in pratica la gran parte dei problemi posti dalla gestione della conoscenza.

L'estrema semplicità di una tale architettura a nodi e di definizioni che scaturiscono da una procedura prefissata è, paradossalmente, ciò che dà al sistema la capacità di raggiungere l'alto livello di complessità richiesto da una moderna gestione della conoscenza.

Joaquín Bastos Amigo è amministratore delegato e direttore del Dipartimento R&D della Carrot SL. Ha iniziato a lavorare nel mondo dell'IT nel settore dei software commerciali, continuando a coltivare l'interesse per la storia e la linguistica. Nel 1998 è entrato alla Carrot SL, un'azienda volta principalmente alla ricerca nel campo della gestione delle informazioni. Ha contribuito a sviluppare il primo sistema Carrot basato sulle pagine "intelligenti" e il secondo sistema, Dconcept, che combinava la gestione logica dei concetti con database relazionali volti al trattamento dei documenti. L'esperienza acquisita nello sviluppo di questi primi sistemi, insieme agli studi sui processi di memorizzazione del cervello umano e sui meccanismi di apprendimento delle lingue naturali, lo hanno portato allo sviluppo di Sintagma. <joaquinbastos@e-carrot.net>

Giuseppe Prencipe, insegnante e traduttore freelance. Si è laureato in Lettere all'Università Cattolica "S. Cuore" di Milano e ha approfondito le tematiche relative alla documentazione, alla biblioteconomia e alla gestione delle conoscenze conseguendo il diploma in Archivistica presso l'Archivio di Stato di Bari e il diploma di perfezionamento in Storia moderna presso l'Università di Urbino.