

Risemantizzazioni e neoformazioni tra linguistica e informatica: *token*, *tokenizzazione* e *NFT*

Lucia Francalanci

PUBBLICATO: 21 MARZO 2023

I termini *token* e *tokenizzazione* sono in circolazione in ambiti specialistici della nostra lingua già dal secolo scorso (*token* a partire dagli anni Trenta e *tokenizzazione* dagli anni Ottanta), ma negli ultimi anni si è assistito a un rilancio delle due forme, dovuto in parte alla diffusione dei concetti di *blockchain*, *criptovaluta* e *metaverso*, sui quali torneremo più avanti. Parallelamente all'aumento della frequenza, i due vocaboli hanno subito anche un processo di risemantizzazione, diventando così neologismi semantici. Contemporaneamente a *token* e *tokenizzazione* si è affermato anche l'acronimo NFT (*non fungible token* 'token non fungibile'), che, come vedremo, è invece una neoformazione a tutti gli effetti.

La prima parte di questa trattazione è dedicata alla discussione del significato dei termini nei diversi ambiti d'uso, mentre la seconda ne ripercorre l'evoluzione nel tempo.

Token

Il sostantivo maschile invariabile *token*, prestito integrale dall'inglese, è usato in italiano con diverse accezioni, soprattutto in ambiti specialistici, come la linguistica e l'informatica. I principali dizionari inglesi online (*Cambridge Dictionary*, *Collins Dictionary*, *Merriam-Webster*) registrano vari significati: nell'uso comune il termine può indicare 'un segno, una prova', cioè 'qualcosa che serve a rappresentare o a indicare un fatto, un evento' (*black is a token of mourning* 'il nero è un segno di lutto'), o 'qualcosa che serve a esprimere un sentimento, un'intenzione' (*as a token of goodwill, I'm going to write another letter* 'come prova di buona volontà, scriverò un'altra lettera'); 'un ricordo, un souvenir' (*the seashell was a token of their trip* 'la conchiglia era un ricordo del loro viaggio'); 'un emblema, un distintivo, un simbolo' (*a white flag is a token of surrender* 'una bandiera bianca è un segno di resa'). Inoltre, come sinonimo di *coin* 'moneta', può fare riferimento a un 'gettone', ovvero a un 'dischetto di metallo o di plastica usato come mezzo di pagamento al posto del denaro', ma anche a un 'buono, coupon, tagliando', cioè a 'un documento, per lo più cartaceo, che può essere scambiato con merci, servizi o denaro di pari valore'.

Di queste accezioni, l'unica che si attesta nell'uso corrente anche nell'italiano comune (un discorso a parte va fatto per gli ambiti specialistici) è quella di *token* come 'gettone' o 'moneta': si definiscono *token* i gettoni usati nel gioco d'azzardo o nelle sale giochi, nei distributori automatici, nelle lavanderie a gettone, negli autolavaggi, ecc., ma anche le monete usate per i carrelli della spesa, le monete commemorative, i buoni pasto, i buoni d'acquisto e così via. Sono così chiamati anche i gettoni che vengono usati durante concerti, manifestazioni ed eventi in sostituzione del denaro, generalmente per acquistare cibo o bevande; il loro valore è stabilito dall'organizzatore dell'evento e spesso non sono

rimborsabili (cioè non possono essere riconvertiti in moneta).

Token in ambito linguistico

In linguistica e nella semiotica, il termine *token* fa riferimento alla ‘realizzazione concreta di un’unità linguistica in un testo o in un discorso’. Talvolta tradotto con *replica*, si trova spesso in contrapposizione alla voce *type* ‘tipo, modello’; i due termini provengono dalla statistica e sono usati in linguistica per designare ‘un elemento linguistico concreto, osservabile, un’occorrenza’ (*token*) e ‘la classe astratta alla quale può essere ricondotto’ (*type*), come un fonema, una parola, una costruzione, ecc. In italiano, tale accezione è registrata esclusivamente nel *Supplemento* 2009 del **GDLI** e nel **GRADIT**, mentre nel *Vocabolario Treccani online* il riferimento a *token* è all’interno del lemma *replica* (per quanto riguarda i dizionari inglesi, gli unici a menzionare tale significato sono il *Collins Dictionary* e il *Merriam-Webster*).

Nella linguistica computazionale e nella linguistica dei corpora, specialmente nell’analisi automatica dei testi, il termine *token* viene spesso reso con *occorrenza* e rappresenta ‘l’unità minima in cui è diviso il testo elettronico’. La prima definizione di *token* in tale senso risale a Charles Sanders Peirce, nei suoi *Prolegomena to an Apology for Pragmaticism* del 1906, di cui riportiamo la traduzione di Massimo Bonfantini (il passo è tratto da Barbera 2013, p. 24, in cui è presente anche il testo originale):

Un modo corrente per giudicare della quantità della materia contenuta in un manoscritto o in un libro stampato è contare il numero delle parole, seguendo il metodo messo in uso dal dottor Edward Eggleston. Di solito ci saranno una decina di *il* in una pagina, e naturalmente conteranno per dieci parole. Ma in un altro senso della parola “parola” c’è solamente una parola “*il*” nella lingua [...]. Una tale Forma definitivamente significativa propongo di chiamarla *Type*. [...] Un *Type* per poter essere usato deve essere reso attuale in un **Token**, che sarà un segno del *Type* e perciò dell’oggetto che il *Type* significa. Propongo di chiamare un tale **Token** di un *Type* *Occorrenza* del *Type*. Così, in una pagina ci potranno essere dieci *Occorrenze* del *Type* “*il*”. (Charles Sanders Peirce, *Semiotica*, testi scelti e introdotti da Massimo A. Bonfantini, Letizia Grassi, Roberto Grazia, Torino, Einaudi, 1980, p. 230)

Il *token* è quindi un’entità piuttosto eterogenea, che comprende non solo le parole tradizionali (nel trattamento elettronico dei testi si pone anche il problema della distinzione tra *parola*, convenzionalmente definita come ‘una qualsiasi sequenza di caratteri delimitata da spazi’ e *lessema* ‘unità minima del lessico’, per cui in una sequenza come *chiodo scaccia chiodo* abbiamo tre parole e due lessemi), ma anche numeri, formule, date, sigle, acronimi, abbreviazioni, segni di punteggiatura. Altre entità che possono essere trattate come un unico *token* sono le polirematiche (*di rado*, *ad hoc*, *fuori servizio*), i nomi propri, le strutture alfanumeriche (*50 euro*, *130 kg*, *2 agosto 2022*).

Il numero totale delle occorrenze di un testo (*token*) corrisponde all’*estensione* o *lunghezza* di un testo (N), cioè alla sua “dimensione”. Il numero delle forme distinte di un testo, ovvero l’insieme delle parole diverse (*type*, chiamate anche *forme grafiche* o *parole tipo* o *tipi di parole*) individua il vocabolario di un testo (V), o meglio, l’*ampiezza* del vocabolario. Il rapporto *tipo/unità*, cioè tra i tipi di parole e le occorrenze testuali (*type/token ratio*: V/N) può essere interpretato come indice della *ricchezza lessicale*

di un testo. I valori di questo rapporto oscillano tra 0 e 1: maggiore è il valore, maggiore è la varietà del vocabolario presente nel testo. Ad esempio, se in un testo ci sono 1.000 occorrenze (N) e i tipi di parole sono solo 50 (V), il rapporto *type/token* sarà uguale a 0,05.

Token in ambito informatico

Per quanto riguarda l'ambito informatico, vi è una sola accezione di *token* registrata dai dizionari sincronici italiani (e soltanto dallo Zingarelli, a partire dal 2016, e dal Devoto-Oli 2023 online, che non riportano invece quella linguistica), cioè quella di 'dispositivo elettronico di piccole dimensioni che genera automaticamente codici di sicurezza temporanei da utilizzare nelle operazioni di home banking'. Il token viene utilizzato sia per l'autenticazione e l'accesso remoto di un utente al proprio conto corrente e ai servizi bancari, sia per l'autorizzazione a effettuare operazioni online (vengono usate con lo stesso significato anche le espressioni *codice token*, *token bancario*, *chiavetta token*, ecc.). Tale dispositivo può essere sia fisico (la classica "chiavetta"; si parla in questo caso di *hardware token*, anche se nella nostra lingua è maggiormente usata la sequenza *token hardware*, in cui le parole sono disposte nell'ordine tipico dell'italiano), sia virtuale (chiamato *software token*; in italiano sono più frequenti forme come *token software*, *token virtuale* o *token digitale*).

Nei dizionari specialistici di informatica è presente un'altra accezione tecnica di *token* (che è reso talvolta in italiano con *gettone*, talvolta con *segnale*, ma si trovano come traducenti anche *marcatore* e *contrassegno*), assieme ad alcune locuzioni formate con il sostantivo (ad esempio, sono lemmatizzate, come prestiti integrali dall'inglese, le espressioni *token bus*, *token passing* e *token ring*). Riportiamo la definizione presente sul volume di Paolo Luigi Monti (*Informatica: dizionario enciclopedico*, Milano, Jackson, 1987), il primo dizionario specialistico, tra gli strumenti consultati, a lemmatizzare la voce *token*:

TOKEN *Gettone*. Termine utilizzato per indicare una sorta di contrassegno che viene passato tra le varie stazioni di una rete per dare la possibilità a ciascuna di esse di trasmettere dati.

In una rete di calcolatori, il token è quindi un segnale, tipicamente costituito da una sequenza di bit, che indica a una stazione la possibilità di trasmettere sulla rete (cioè la stazione è autorizzata a immettere dati senza che si verifichino delle contese).

Accanto ai significati di *token* registrati dalla lessicografia, troviamo un'ulteriore accezione informatica, più recente e collegata ai concetti di *blockchain* e *criptovaluta*. Questo nuovo uso è segnalato soltanto dall'Enciclopedia Treccani online, in cui sono lemmatizzati sia *token* sia il NFT, acronimo di *non fungible token* 'token non fungibile', già ricordato all'inizio, di cui parleremo più avanti.

Partiamo dal testo della **voce** *token* dell'Enciclopedia Treccani per la discussione del significato:

In informatica, termine con cui si designa un indicatore univoco registrato in una *blockchain* (registro condiviso), con funzione di rappresentare un oggetto digitale, di certificare la proprietà di un bene o di consentire l'accesso a un servizio. Costituiti da una sequenza di informazioni digitali protette, i **token** si

distinguono in due categorie principali: **token fungibili** (sostituibili con altri dalle caratteristiche identiche, come nel caso della moneta elettronica), utilizzati come mezzo di pagamento per l'acquisto di beni o la fruizione di servizi all'interno di un ambiente digitale, o garantiti da *commodities* quali preziosi, oro o argento; e **token non fungibili** (dotati di un codice di identificazione che ne attesta l'unicità e la proprietà), utilizzati per la gestione dell'identità digitale e la certificazione della proprietà e dell'autenticità di un bene digitale, quali ad esempio gli **NFT**.

Cerchiamo di definire in primis il termine e il concetto di *blockchain*. Questa voce inglese, composta da *block* 'blocco' e *chain* 'catena', letteralmente 'catena di blocchi', è registrata, come sostantivo femminile invariabile, soltanto dallo **Zingarelli** (dal 2020) e dal **Devoto-Oli** online (che rimanda al lemma *catena*), ma è presente anche nell'**Enciclopedia Treccani** e nel repertorio **Treccani Neologismi 2018**. Indica un "database le cui informazioni, strutturate in pagine (dette 'blocchi'), criptate con rigidi criteri di sicurezza e decentralizzate in vari nodi della rete, sono accessibili a chi ne ha titolo, ma modificabili solo col consenso di tutti" (Zingarelli 2023). Ogni blocco contiene sequenze di dati (dette *transazioni*), protetti da crittografia e retrospettivamente non modificabili. Ogni volta che avviene una transazione, questa viene registrata, ovvero viene inviata alla rete con tutte le informazioni correlate alla transazione: le stesse informazioni saranno presenti su tutti i nodi, per cui sono immutabili se non tramite l'approvazione della maggioranza dei nodi della rete. Collegando queste sequenze si crea appunto una "catena di blocchi": in pratica, ogni nuovo gruppo di dati forma un nuovo blocco che si concatena a quello precedente.

Un primo prototipo della tecnologia blockchain viene descritto nel 1991 dai ricercatori Stuart Haber e W. Scott Stornetta. Alla fine del 2008, Satoshi Nakamoto (pseudonimo di un programmatore anonimo o di un gruppo di persone la cui identità è sconosciuta) pubblica un *white paper* (***Bitcoin: A Peer-to-Peer Electronic Cash System***), un documento tecnico che introduce un sistema di moneta elettronica peer-to-peer decentralizzata, chiamato Bitcoin. Bitcoin è quindi la prima criptovaluta a usare la blockchain come registro contabile delle transazioni effettuate: la prima transazione di bitcoin avviene il 12 gennaio 2009. Oggi la tecnologia blockchain ha molti ambiti di applicazione, non necessariamente legati al mercato delle criptovalute.

Tornando al token, una definizione esauriente è quella proposta nel blog degli Osservatori Digital Innovation del Politecnico di Milano:

Un **token** su Blockchain consiste in un'informazione digitale, registrata su un registro distribuito, univocamente associata a uno e un solo specifico utente del sistema e rappresentativa di una qualche forma di diritto: la proprietà di un asset [dal GRADIT: "ciascun elemento dell'attivo di bilancio, come beni di proprietà, liquidità, crediti e simili", ndr], l'accesso a un servizio, la ricezione di un pagamento, e così via. (Valeria Portale, ***Initial Coin Offer (ICO) e Token: ecco cosa sono e quali sviluppi promettono per il futuro***, blog.osservatori.net, 2/1/2019)

Il token è quindi una sequenza di informazioni digitali, registrate in una blockchain (che è una forma di registro distribuito) e rappresentative di una qualche forma di valore o di diritto. Un *registro distribuito* è un archivio di informazioni condiviso, sincronizzato in rete, che ha un'architettura decentralizzata, anziché centralizzata; per indicare le tecnologie che si basano sui registri distribuiti, come la blockchain, si usa l'acronimo DLT, che sta per *distributed ledger technology* 'tecnologia dei

registri distribuiti’.

Il valore assunto da un token è stabilito da regole di mercato ed è fissato dall’ente o dall’organizzazione che emette tale token; è necessario specificare che un token ha un valore determinato solo ed esclusivamente se inserito all’interno di un contesto. Si pensi ad esempio ai bollini o ai punti che si ricevono ai supermercati a ogni spesa effettuata, grazie ai quali si possono ottenere premi o sconti; di per sé tali oggetti non avrebbero alcun valore, ma viene attribuito loro un valore riconosciuto esclusivamente all’interno del contesto in cui vengono scambiati. Il loro valore è deciso arbitrariamente dal supermercato che li emette.

Le transazioni di token sono regolate da *smart contract* (‘contratti intelligenti’), contratti digitali inseriti in una blockchain (quindi immutabili e trasparenti) in cui vengono definiti i limiti e le proprietà dei token: il contratto contiene ad esempio un elenco di indirizzi che consente l’individuazione di chi può disporre dei token (i cosiddetti *token holder*); uno sviluppatore (chi crea token e applicazioni basate sulla blockchain) può anche imporre un limite specifico alla quantità di token in circolazione e applicare una proprietà speciale che ne determina la rarità (ad esempio specificando che in una collezione di 1.000 token vi sono solo 10 rarità).

Token fungibili e token non fungibili (NFT)

I token sono suddivisibili in due categorie: i token fungibili e quelli non fungibili. Nel diritto e nell’economia, un bene è detto *fungibile* quando è privo di una sua individualità specifica e può essere sostituito con un altro bene dello stesso genere, ovvero quando è intercambiabile; si pensi ad esempio al denaro, all’oro, ma anche al grano, all’olio, al petrolio. Se scambio una banconota da 50 euro con un’altra da 50 euro oppure con 5 banconote da 10 euro il valore è lo stesso. Un bene è invece infungibile, o non fungibile, quando è dotato di una propria individualità socio-economica e non è sostituibile con un altro simile.

I token fungibili sono standardizzati e intercambiabili. A tale categoria appartengono le criptovalute o *coin* (come Bitcoin (BTC) o Ether (ETH), la valuta utilizzata per le transazioni sulla piattaforma Ethereum) e tutti quei token le cui caratteristiche sono assimilabili a quelle di una moneta elettronica. Sono utilizzati per l’acquisto di beni o la fruizione di servizi (ad esempio il biglietto per un concerto o per un volo aereo); si distinguono a loro volta in *utility token* (spendibili in un dato ambiente digitale, come un sito o un **marketplace**, per l’acquisto di beni o servizi, per usufruire di sconti o effettuare determinate azioni), *payment token* (utilizzati esclusivamente come forma di pagamento; rientrano in tale categoria gli *stablecoin*, criptovalute dal valore stabile legato a valute reali, come il dollaro americano), *commodity token* (garantiti da *commodities* quali preziosi, oro o argento), *security token* (token emessi sulla blockchain, che rappresentano quote di una società o di un qualche bene; costituiscono l’equivalente dei titoli finanziari tradizionali, come azioni e obbligazioni, ma beneficiano delle proprietà della blockchain, oltre che di misure di regolamentazione per prevenire le frodi); alla categoria dei *security token* appartengono gli *equity token*, che rappresentano il possesso di una attività, come le azioni di una società.

I token non fungibili (NFT, acronimo, come si è detto più volte, di *non fungible token*) sono invece definiti dalla loro unicità: sono originali e rappresentano oggetti unici, non divisibili, non interscambiabili, non ripetibili. Contengono dei metadati identificativi e descrittivi, che attestano l'unicità di un bene e la sua proprietà. In pratica gli (sull'uso di questo articolo invece di *i* si tornerà più oltre) NFT sono certificati di autenticità e di proprietà di beni, fisici o digitali, registrati su una blockchain e dunque non modificabili. Il termine NFT è messo a lemma soltanto dallo Zingarelli 2023, che registra anche, all'interno del lemma *token*, le due polirematiche *non-fungible token* e *fan token*.

Gli NFT sono utilizzati per la gestione dell'identità digitale, per la gestione della tracciabilità e dell'automazione dei processi di *supply chain* ('catena degli approvvigionamenti', ovvero il processo che permette di portare sul mercato un prodotto o servizio, trasferendolo dal fornitore fino al cliente), per i sistemi di voto elettronico, ma trovano applicazione anche e soprattutto nell'ambito del collezionismo digitale (per indicare gli oggetti da collezione digitali si usa solitamente il prestito integrale *collectibles*), in particolare nei giochi online (*gaming online*) e nella *criptoarte*, termine che fa riferimento a opere d'arte digitali caricate su una piattaforma e registrate su una blockchain; rientrano in tale categoria immagini, foto, video, meme, GIF, ma anche *tweet* (numerose le varianti formali attestate: *criptoart*, *cripto art*, *cripto arte*, *cripto-arte*, *cryptoarte*, *cryptoart*, *crypto art*, *crypto arte*, *crypto-arte*). Il loro impiego è comunque in continua espansione anche in altri campi, come la *musica*, la *letteratura*, lo sport (ad esempio, nel mondo del calcio, il sistema dei *fan token* per coinvolgere i tifosi), la *moda*, il *food & beverage* ('cibo e bevande'; si parla a tale proposito di *eatable token*).

Gli NFT sono divenuti popolari nel 2017 grazie a CryptoKitties, videogioco sulla blockchain di Ethereum in cui è possibile acquistare, collezionare o rivendere gattini digitali; ogni CryptoKitty è rappresentato da un NFT e può essere acquistato con la criptovaluta Ether (all'inizio il costo di un gattino era di 0,008 ETH, che corrispondeva a qualche dollaro; il gioco è divenuto virale in poco tempo e alcuni gattini sono arrivati a un valore di centomila dollari). È però dal 2021 che l'interesse per gli NFT è decollato, soprattutto nel mondo dell'arte. "Con un mercato globale (comprendente, tra gli altri, videogiochi, avatar, *utilities*, musica, proprietà immobiliari virtuali e opere cinematografiche) attestato intorno ai 10,7 miliardi di dollari nel terzo trimestre del 2021 e in crescita costante, gli NTF si sono imposti con amplissima risonanza nell'arte digitale, nel cui ambito consentono di decentralizzare il processo di esposizione e commercializzazione, fornendo massime garanzie alla tutela del diritto d'autore e garantendo l'accesso a mercati di avanguardia" (dalla voce *non fungible token* (NFT) dell'Enciclopedia Treccani online). Un esempio che bene illustra la portata del fenomeno è il seguente: a marzo 2021 l'artista digitale americano Mike Winkelmann, noto con il nome d'arte Beeple, ha venduto la sua opera *Everydays: The first 5000 days*, un collage digitale di 5.000 jpeg, per 69,3 milioni di dollari: è il token non fungibile più costoso venduto finora.

L'espansione del mercato degli NFT è strettamente collegata al fenomeno del *metaverso* (insieme di ambienti virtuali tridimensionali in cui le persone possono interagire tra loro attraverso avatar personalizzati). Il metaverso, anzi, i vari metaversi operano sulla blockchain e possiedono una vera e propria economia funzionante (detta appunto *meta-economia*), basata sulle criptovalute: gli utenti possono creare, acquistare o scambiare i propri beni digitali sotto forma di NFT. "Sta già cambiando il modo in cui compriamo, vendiamo e persino il modo in cui pensiamo al concetto di denaro stesso. [...] Gli NFT, le criptovalute e la tokenizzazione sono, per molti versi, le caratteristiche di un mondo

completamente nuovo. Che è, chiaramente, proprio quello che il Metaverso si propone di essere” (dal discorso di Giuliano D’Acunti, responsabile per l’Italia di *Invesco*, riportato sul sito *financialounge.com*: Leo Campagna, *Transazioni nel Metaverso, la spinta da blockchain e token non fungibili (NFT)*, 22/6/22).

Questa citazione ci consente di fare un ulteriore passo avanti e introdurre anche il concetto di tokenizzazione.

Tokenizzazione

In linguistica, il processo di segmentazione di un testo in unità minime di analisi, cioè di suddivisione in token, è detto *tokenizzazione*; i token rappresentano l’unità di base per i successivi livelli di elaborazione (analisi morfo-sintattica del testo, lemmatizzazione, ecc.). La tokenizzazione, intesa come identificazione delle occorrenze di un testo, non è un’operazione semplice e non corrisponde all’identificazione delle parole: ad esempio, sono considerati token singoli *diglielo*, *portami*, *andandosene*, *darglielo*, anche se dal punto di vista morfo-sintattico sono composti da unità distinte (per approfondimenti sui token e sui criteri di tokenizzazione si rimanda a Barbera, Corino, Onesti 2007; Bolasco 2013; Chiari 2007; Lenci, Montemagni, Pirrelli 2005).

In ambito informatico, invece, col sostantivo *tokenizzazione* si indica il processo di ‘assegnazione di un token a un dato asset’ (il verbo corrispondente è *tokenizzare*): in pratica, consiste nel generare un token e collegarlo a un dato bene mediante uno smart contract, registrando tutte le sue informazioni (il valore, la proprietà o i diritti di accesso) in una blockchain. Potenzialmente, tutto può essere “tokenizzato”: opere d’arte, letteratura, abbigliamento, accessori, cibo, squadre sportive, ma anche immobili, quote di azioni, identità, materie prime, ecc.

In un *documento* del 2020, l’Organizzazione per la cooperazione e lo sviluppo economico (OCSE; in inglese OECD = Organization for Economic Co-operation and Development), riprendendo le definizioni proposte da Hileman e Rauchs (2017) e dal Financial Stability Board (FSB 2019), spiega che la tokenizzazione “comporta la rappresentazione digitale di asset reali (o fisici) su un registro distribuito oppure l’emissione di classi di asset tradizionali sotto forma di token” (p. 7, trad. mia). Nel primo caso, il valore economico e i diritti derivanti dagli asset reali preesistenti sono collegati o incorporati in token basati sulla DLT: i token emessi esistono sulla blockchain e condividono gli stessi diritti degli asset che rappresentano, fungendo da riserva di valore; gli asset reali su cui sono emessi i token continuano ad esistere *off-chain* (‘fuori dalla catena’, ovvero fuori dalla blockchain). Esempi di beni reali che esistono *off-chain* sono le proprietà immobiliari, le *commodities*, le opere d’arte, i *collectibles*, ecc.

Nel secondo caso, si parla invece di emissione di *token nativi*, creati direttamente *on-chain* (‘sulla catena’, cioè sulla blockchain) ed esistenti esclusivamente sul registro distribuito. Esempi di token nativi sono le criptovalute, i *payment token* e le ICO (*Initial Coin Offer*; si tratta di una forma di finanziamento di un dato progetto realizzato tramite blockchain, che prevede la creazione di token, che sono poi venduti, in cambio di un corrispettivo, ai finanziatori del progetto. Per approfondimenti

si rimanda agli articoli presenti sui siti *Osservatori Digital Innovation* del Politecnico di Milano e *Network Digital 360*).

Riassumendo, possiamo dire che la lessicografia italiana registra esclusivamente l'accezione linguistica di *token* come 'occorrenza', in opposizione alla voce *type* (Supplemento 2009 del GDLI e GRADIT) o quella informatica di 'dispositivo di sicurezza' (Zingarelli dal 2016 e Devoto-Oli 2023). Il significato più recente di *token* come 'sequenza di informazioni digitali registrate in una blockchain' non è (ancora) segnalato dai dizionari, ma soltanto dall'Enciclopedia Treccani online, in cui sono lemmatizzati sia *token* sia NFT. Risultano assenti dai dizionari italiani anche il sostantivo *tokenizzazione*, in entrambe le accezioni (linguistica e informatica) e il verbo *tokenizzare*; l'unico dizionario a lemmatizzare l'acronimo NFT è lo Zingarelli 2023, che registra anche le polirematiche *non-fungible token* e *fan token* all'interno del lemma *token*.

Per quanto riguarda i dizionari inglesi, il *Merriam-Webster* è l'unico a riportare l'accezione di *token* come *cryptocurrency* 'criptovaluta' e a lemmatizzare la polirematica *utility token*; sono invece censiti sia il sostantivo *non-fungible token* (*Cambridge Dictionary*, *Collins Dictionary*, *Merriam-Webster*), nella forma con il trattino, sia il suo acronimo NFT (*Cambridge Dictionary*, *Collins Dictionary*, *Merriam-Webster*), datati, dal *Merriam-Webster*, 2017. Il sostantivo *tokenisation/tokenization* non è registrato dalla lessicografia inglese; sul sito *Dictionary.com* è tuttavia presente il verbo *tokenize* con vari significati, tra cui quello, marcato come specialistico della tecnologia digitale, di 'dividere (un bene materiale o immateriale) in token virtuali che possono essere venduti o scambiati' (trad. mia).

L'evoluzione dei termini *token* e *tokenizzazione*

La storia di *token*: le prime attestazioni

Per cercare di comprendere quando e in che modo sia avvenuta la rideterminazione semantica delle forme *token* e *tokenizzazione*, è necessario ripercorrere le tappe della storia dei due termini. La prima attestazione rintracciata in italiano del sostantivo *token* è del 1933, in un articolo della "Stampa della Sera", e il senso è quello di 'buono usato in sostituzione del denaro' (che è uno dei significati di uso comune della voce inglese): nel testo si fa riferimento ai buoni del valore di 1 dollaro emessi dal governo degli Stati Uniti durante il periodo della Grande depressione; gli Stati Uniti stavano infatti affrontando una grave crisi economica e finanziaria, che portò nell'aprile del 1933 all'emanazione da parte del presidente Roosevelt dell'Ordine esecutivo 6102, con il quale si proibiva il possesso dell'oro in qualsiasi forma (monete, lingotti, certificati aurei) da parte di ogni individuo, associazione e società residente negli USA.

Il "TOKEN": Una eccezionale documentazione fotografica dei provvedimenti adottati dagli S. U. per fronteggiare la situazione finanziaria: Buoni - chiamati "token" corrono ora in sostituzione del dollaro carta di cui non possono durante la «vacanza bancaria» essere fatti prelievi presso le banche dal momento che il dollaro-banconota resta moneta cambiabile in oro. La Stampasera è in grado di pubblicare uno di questi «token» la cui fotografia è stata trasmessa per radio a Londra e di là per postaerea a Torino. (*Il Labaro del P.N.F.* [Partito Nazionale Fascista] *issato stamane a palazzo Venezia*, "La

Stampa della Sera”, 9/3/1933, p. 1)

Tra il 1933 e il 1937, sempre sulla “Stampa” o sulla “Stampa della Sera”, il termine compare in riferimento al debito di guerra contratto dagli inglesi nei confronti degli Stati Uniti; il token è un versamento simbolico, pro forma:

Si prevede che le discussioni saranno lunghe e, forse, faticose, al punto che in alcuni ambienti di Washington non si nasconde la probabilità che si giunga al 15 dicembre, data di scadenza del prossimo pagamento, senza una intesa su una soluzione, magari provvisoria, del problema; in questo caso si ricorrerebbe, ancora una volta, al pagamento del «token», ossia del versamento di una somma, da convenirsi, mirante nient'altro che a salvaguardare lo spirito dell'accordo Baldwin-Mellon ed a provare la sua sopravvivenza all'uragano finanziario ed economico degli ultimi anni. (R. P., *I debiti e la stabilità della moneta nelle imminenti conversazioni anglo-americane*, “La Stampa”, 29/9/1933, p. 7)

Il termine sembra poi sparire dai quotidiani nazionali e ricompare solo nel 1958, stavolta in un testo a stampa (Carlo Ravasini, *Documenti sanitari: bolli e suggelli di disinfezione nel passato*, Torino, Minerva medica) con il significato di “marca monetata emessa in epoca di emergenza da privati o da banche” (p. 38), usata in sostituzione del *farthing*, antica moneta britannica dal valore di un quarto di penny, corrispondente al nostro *quattrino*, coniata fino al 1956.

In queste prime attestazioni, il *token* viene semplicemente citato come voce straniera (tra virgolette), usata esclusivamente in riferimento a contesti anglosassoni: non ha, cioè, ancora assunto lo status di prestito.

La storia di *token*: prime occorrenze nei linguaggi specialistici

Le attestazioni successive sono del 1971; una si trova nel volume di Umberto Eco *Le forme del contenuto*:

In questo senso il referente viene semiotizzato: esso non viene preso come *token*, come individuo, ma viene arbitrariamente reso il *type* di una serie più vasta di oggetti di uno dei quali esso faceva parte. (Umberto Eco, *Le forme del contenuto*, Milano, Bompiani, 1971, p. 39)

Si ha quindi una prima occorrenza di *token* nella sua accezione semiotica (in realtà il GRADIT indica come data di prima attestazione il 1962, con un rimando al volume di Eco *Opera aperta*: la ricerca della forma all'interno del testo non ha però dato riscontri. Il GDLI, invece, data il sostantivo 2002). Nel libro, Eco fa esplicitamente riferimento alla terminologia peirciana, in particolare alla raccolta *The Collected Papers* (Harvard University Press, 1931-1935): sembrerebbe quindi che tale significato specialistico del sostantivo inglese sia penetrato nella nostra lingua grazie agli studi italiani delle opere di Pierce in lingua originale (la prima traduzione del volume *Prolegomena to an Apology for Pragmaticism* del 1906, in cui – lo ricordiamo – troviamo la prima definizione di *token*, è infatti solo del 1980).

L'altra occorrenza si trova nell'ambito della statistica linguistica, nella prefazione al volume di Gustav

Herdan *Linguistica quantitativa*:

Si osservi dapprima che i termini «particella», «stringa», «campo» sono virtualmente geometrici e corrispondono rispettivamente a «punto», «retta» e «piano», ai quali, come ho fatto vedere nell'area del linguaggio corrispondono i termini «type», «**token**» e la «relazione di dualità **type-token**» analoga a quella fra punti e rette nel piano. (Gustav Herdan, *Linguistica quantitativa*, Bologna, Il Mulino, 1971, p. 5)

Questi usi specifici di *token* rimangono comunque circoscritti ad ambiti prettamente specialistici: se ne trovano molte occorrenze in rete, in volumi dedicati all'argomento e nelle riviste specializzate, ma la forma (in tali significati) è quasi del tutto assente nella stampa quotidiana. L'unica attestazione rintracciata in tale senso (e ben 40 anni più tardi) è nella rubrica *Lessico e Nuove* di Stefano Bartezzaghi:

I giochi di riduzione alfabetica non vanno confusi con gli anagrammi, totali o parziali, a cui pure assomigliano. Mentre gli anagrammi danno un inventario chiuso di lettere, da usare per un numero stabilito di volte, nei giochi di riduzione alfabetica ogni lettera può ricorrere tante volte quante si vuole (per gli adepti: non si gioca con l'occorrenza, o **token**, ma con il tipo, type). Il più famoso gioco di riduzione alfabetica è l'antico lipogramma, riportato alla gloria nel secolo scorso da Georges Perec: si usano tutte le lettere dell'alfabeto tranne una. (Stefano Bartezzaghi, *Il Tuscolano e l'amore*, "la Repubblica", 9/12/2005)

Proseguendo con la rassegna, troviamo nel 1977 un'occorrenza della locuzione *book token* 'buono per l'acquisto di libri' nella "Stampa" (la successiva è del 1992) e, a partire dal 1982, alcune attestazioni dell'espressione *codice token* 'codice a un solo byte' in "MC microcomputer", una delle prime riviste di informatica in Italia, edita da Technimedia dal 1981 (nella rivista troviamo anche la voce singola *token*, usata con lo stesso significato). Ancora in riviste specializzate è attestata nel 1987 la locuzione *token ring* (che nello stesso anno viene lemmatizzata nel già citato dizionario di informatica di Monti); è però plausibile che fosse già in circolazione anche negli anni precedenti: l'architettura *token ring* è stata infatti introdotta da IBM nel 1984 e successivamente standardizzata nel 1989 come IEEE 802.5.

A partire dal 1990 riappare l'accezione più comune di *token* come 'gettone', in particolare in riferimento a quello della metropolitana usato negli Stati Uniti:

A differenza di Londra o Parigi, il "subway" viaggia tutta la notte [...]. Ogni giorno il "Sistema" è utilizzato da 3,7 milioni di persone, più di 1 miliardo all'anno. La maggioranza entra con il **token**, il gettone da un dollaro e 25 centesimi [...]. (Arturo Zampaglione, *Un treno di desideri negli 'inferi' della Mela*, "la Repubblica", 2/1/1992, p. 8)

Come si può notare il termine non è più tra virgolette (e così anche negli altri esempi che non riportiamo), ma è comunque ancora strettamente legato al contesto anglosassone (si fa infatti riferimento a notizie della cronaca estera: la notte degli Oscar, la metropolitana di New York, ecc.).

Di *token* inteso come 'gettone della metro' si trovano occorrenze nei quotidiani almeno fino al 2011, a cui si affiancano nel corso degli anni Duemila anche altri usi del sostantivo nel suo significato più comune di 'gettone' o 'buono' (nei vari articoli si parla di *gettone del casinò*, *gettone del monopolio*, *buono spesa*, e così via). L'impiego di *token* come 'tagliando per le consumazioni durante i concerti o altre

manifestazioni' è più recente (in rete se ne trova traccia a partire dal 2015, ma non è escluso che tale uso fosse già in circolazione negli anni precedenti).

La storia di *token*: le accezioni informatiche

Non è invece semplice rintracciare le prime occorrenze del sostantivo nell'accezione informatica di 'dispositivo di sicurezza' (che Devoto-Oli e Zingarelli datano, rispettivamente, 2006 e 2008). Nel 1997 e nel 2001 è possibile individuarne alcuni esempi in testi che si occupano di sicurezza in ambienti informatici, in cui si parla di *token di autenticazione*. La prima testimonianza si trova sulla rivista "AEI" (Automazione, energia, informazione: organo ufficiale dell'Associazione elettrotecnica ed elettronica italiana): si tratta della traduzione di un articolo in inglese scritto da due studiosi canadesi (Ping Lin, Lin Lin, *Security in Enterprise Networking: a Quick Tour*, in "IEEE Communications Magazine", gennaio 1996, p. 61) dedicato alla sicurezza nelle reti d'azienda. Nel glossario dei termini specialistici, *token*, che viene tradotto con 'gettone', è definito come un "oggetto di dati strutturati o messaggio univoco che circola continuamente tra i nodi di un token ring e descrive lo stato della rete in quel momento" (p. 87/495), ma più avanti nel testo si parla anche del metodo di autenticazione "a gettone", alternativo a quello basato sulle password:

I "gettoni d'autenticazione" denominati **token** sono tipicamente dei piccoli dispositivi, tenuti in prossimità o attaccati alla porta seriale o parallela di un pc oppure inseriti nell'unità di controllo dei dischetti per generare sequenze di dati non ripetibili, da usare al posto di parole d'ordine a lungo termine. (Ping Lin, Lin Lin, *La sicurezza nelle reti d'azienda*, "AEI", vol. 84, n. 5, 1997, pp. 85/493-92/500)

L'esempio del 2001 è invece tratto da un manuale rivolto a specialisti del settore e dunque piuttosto tecnico:

Le entità di SPX utilizzano degli **authentication token** per autenticarsi gli uni con gli altri. In modo simile al ticket di Kerberos, i **token** vengono scambiati in una sessione protetta da crittografia DES [...]. (Mariagrazia Fugini, Fabrizio Maio, Pierluigi Plebani, *Sicurezza dei sistemi informatici*, Milano, Apogeo, 2001, p. 35).

È del 2002 la prima attestazione al di fuori dell'ambito prettamente specialistico, in un articolo pubblicato sul "Corriere della Sera":

Ma come vengono riconosciute le identità digitali sicure? «Ci vuole un'autenticazione forte, basata su due fattori – spiega Schnell – . Qualcosa che l'utente conosce, come un codice Pin, e qualcosa che possiede, o che gli è stato attribuito, come una smart card o un dispositivo **Token**. Attualmente il sistema a minor costo è l'abbinamento di un Pin e di un **Token**». Di che si tratta? **Token** significa gettone. Questo dispositivo, che spesso ha la forma di un portachiavi, mostra su un piccolo schermo una sequenza di numeri che cambia ogni minuto. (Chiara Sottocorona, *Una carta d'identità digitale per chi naviga*, "Corriere della Sera", 21/10/2002, p. 23)

In questo esempio, la giornalista riporta le parole di Scott Schnell, vicepresidente dell'azienda americana Rsa Security, ma sente il bisogno di chiarire cosa sia un token dato che, a differenza degli

Stati Uniti, in cui “già diverse banche, assicurazioni e grandi aziende hanno distribuito milioni di dispositivi Token ai propri clienti e dipendenti”, in Italia tale tecnologia è ancora poco conosciuta.

Nonostante provenga dall'ambito specialistico dell'informatica, questo significato prende comunque subito piede in italiano, tanto che negli anni successivi se ne trovano diverse occorrenze anche sui quotidiani (ma fino al 2005 gli articoli sono tutti firmati dalla stessa giornalista del “Corriere della Sera”, Chiara Sottocorona). In ogni caso, la presenza del vocabolo in tale accezione sembra essere piuttosto costante nel corso degli anni Duemila, anche se non raggiunge mai un alto tasso di frequenza. Oltre alla singola forma *token*, risultano attestate per indicare il medesimo concetto anche altre locuzioni, come *token di sicurezza*, *token di autenticazione*, *dispositivo token*, *chiavetta token*, *token usb*, ecc.

La storia di *token*: le accezioni più recenti

I primi esempi individuati in cui il sostantivo viene associato al concetto di ‘moneta virtuale’ sono del 2008 e del 2009; in questi contesti il token è un ‘gettone virtuale’ che può essere accumulato e speso in giochi online o in mondi virtuali ed è quindi usato come una valuta che assume un dato valore esclusivamente all'interno di uno specifico ecosistema; non si fa ancora riferimento alle monete elettroniche o alle criptovalute (come abbiamo visto, *Bitcoin* viene introdotto solo alla fine del 2008 e la prima transazione avviene soltanto l'anno successivo):

Una piattaforma di giochi, ma soprattutto una community. Questo è Fueps.com [...]. Il gioco resta, però, l'imprescindibile punto di partenza. Partecipare significa anche vincere e accumulare «**token**». Una moneta virtuale da spendere su Fueps plus, sezione del sito dedicata ai concorsi. (Davide Milosa, *Febbre da «casual game»*. *L'antistress sulla Rete*, “Corriere della Sera”, 19/6/2008, p. 31)

Per partecipare ai tornei, si dispone di una somma di partenza costituita da una moneta virtuale: i «**token**» (gettoni). Più si gioca e più, chiaramente, si guadagnano posizioni in classifica. I **token**, poi, possono essere utilizzati per sfidare gli amici e partecipare alle aste on-line. (Simona De Santis, *Facebook alla romana, per gioco*, “Corriere della Sera”, 22/1/2009, p. 16)

Le prime attestazioni di *token* nell'accezione più recente, cioè quella strettamente collegata al concetto di blockchain, si hanno solo a partire dal 2017; guardando nel dettaglio i contesti delle occorrenze, anche negli anni successivi, si nota che in molti casi c'è una certa sovrapposizione del concetto di token con quello di criptovaluta o comunque con quello di moneta virtuale, identificazione che ancora oggi non sembra essere del tutto superata.

Riportiamo alcuni dei primi esempi rintracciati, che provengono dai siti web della rivista internazionale “Wired” e del quotidiano nazionale “La Stampa”; si nota che in entrambi i testi si fa riferimento alle ICO (*Initial Coin Offer*) e alle criptovalute:

Nel mondo della finanza si parla di Ipo, l'offerta pubblica iniziale (dall'inglese: Initial public offering) con cui un'azienda vende azioni proprie per raccogliere capitali. Nel mondo delle criptovalute esiste invece uno strumento non dissimile, chiamato Ico (Initial coin offering), con cui un'azienda offre al

mercato dei **token**, con il medesimo scopo di reperire capitali. (Giuditta Mosca, *Come funziona Ico, la nuova forma di finanziamento a startup e imprese*, wired.it, 24/7/2017)

Le Icos sono un modo per una startup di procacciarsi fondi attraverso la vendita al pubblico di una sua specifica moneta virtuale, di un **token** digitale. Chi scommette sul prodotto dell'azienda si compra questa valuta come se fossero delle azioni, e per farlo di solito utilizza criptovalute più solide, come bitcoin. (Carola Frediani, *Stop alle Ico, "cugine" del Bitcoin. La Cina le ha messe al bando*, lastampa.it, 13/9/2017)

Nel corso del 2017 questo uso di *token* è ancora limitato a pochi esempi, ma a partire dal 2018 le occorrenze diventano piuttosto numerose: proprio nel 2018 iniziano a comparire in italiano anche le locuzioni *token non fungibile*, *non-fungible token* e l'acronimo *NFT*.

Le prime attestazioni di *token non fungibile*, *non-fungible token* e *NFT*

È sul web che si possono individuare queste prime attestazioni, che riportiamo. Per quanto riguarda l'espressione *token non fungibile*, se ne trova traccia nel 2018 sia in rete sia nei testi a stampa, mentre si deve attendere il 2020 perché emerga anche sui giornali (sul "Corriere della Sera", mentre le prime occorrenze sulla "Repubblica" e sulla "Stampa" sono del 2021). Anche il prestito non adattato *non-fungible token* e l'acronimo *NFT* sono attestati in rete a partire dal 2018, ma è del 2019 la prima presenza in un volume cartaceo (peraltro, lo stesso: Michael Juntao Yuan, *Sviluppare applicazioni Blockchain: Guida per creare sistemi decentralizzati su reti distribuite*, Milano, Apogeo, 2019); per la presenza sui quotidiani si deve invece aspettare il 2021.

Le proposte di evoluzione vanno dalle proposte infrastrutturali a quelle applicative, le Ethereum request dor Comment. Alcune tra le più importanti sono ERC-20, che rappresenta lo standard per la creazione di nuovi **token** (detti **fungibili**, poiché funzionano come moneta) e ERC-721 che rappresenta quello per la creazione di asset generici, come per esempio una casa o un'opera d'arte (chiamati anche **token non fungibili**, perché rappresentano ognuno un bene specifico e non sono interscambiabili). (Nicola Attico, *Blockchain. Guida all'ecosistema*, Milano, Guerini Next, 2018, consultato nella versione e-book)

L'avvento dei Bitcoin e della rete Ethereum ha aperto la strada all'idea della "scarsità digitale" e dei **token non fungibili** (NFT). Oggi HTC ha stretto una partnership con il primo e più famoso gioco NFT [si osservi qui l'uso aggettivale di NFT] al mondo sulla blockchain, Cryptokitties, per una sua distribuzione esclusiva su alcuni dispositivi HTC, a partire da HTC U12+. (*HTC Exodus, telefono con blockchain integrata*, atomtimes.com, 14/7/2018)

Un ecosistema a parte sono i **NFT**, ovvero i **non-fungible token** che, al contrario di bitcoin e delle criptovalute, non possono essere interscambiabili. Gli **NFT** sono anche definiti collectible token o crypto-game token. La più famosa *CryptoKitties* che utilizza **gli NFT** sulla blockchain di Ethereum. (Stefania Stimolo, *Classificazione token: le differenze tra crypto, stable coin, security, utility ed equity*, cryptonomist.ch, 1/12/2018)

Quest'ultimo esempio ci consente di evidenziare anche l'incertezza relativa alla scelta dell'articolo

associato all'acronimo NFT: l'autrice del brano sembra infatti indecisa e utilizza entrambe le possibilità, *gli NFT* (due volte) e *i NFT* (una volta, la prima). La regola generale (sull'argomento si veda *L'articolo con le sigle*, che riporta quanto scritto in Serianni 1988, cap. IV, par. 9-12) prevede diverse possibilità. Se le sigle sono pronunciate per lettere distinte, avremo *il* quando il nome della prima lettera comincia per consonante: *il CNR* (il ci-enne-erre); con le lettere il cui nome ha iniziale vocalica (come nel nostro caso: *enne*), l'uso è molto incerto. Sembrerebbe però prevalere l'uso degli articoli prevocalici (*l'*, *gli*): *l'SMS* (l'esse-emme-esse), *l'RNA* (l'erre-enne-a), *l'HTML* (l'acca-ti-emme-elle). La forma da prediligere sarebbe quindi *l'NFT* (l'enne-effe-ti) per il singolare e *gli NFT* per il plurale. Tale preferenza è confermata anche dalle ricerche effettuate su Google, in base alle quali si ottengono 65.700 risultati per "l'NFT" e 157.000 per "gli NFT", mentre soltanto 1.230 per "il NFT" e 6.320 per "i NFT".

La storia di *tokenizzazione* e di *tokenizzare*

Per quanto riguarda il sostantivo *tokenizzazione* e il verbo *tokenizzare*, le prime occorrenze rintracciate risalgono all'inizio degli anni Ottanta ed entrambe appartengono all'ambito informatico. Le troviamo nella già citata rivista "MC microcomputer", una delle prime riviste storiche di informatica in Italia, in articoli dedicati al linguaggio di programmazione Basic e ai primi *home computer*, come Acorn Atom e Commodore 64; compare per prima la voce verbale, nel 1982, con il significato di 'convertire le parole chiave [in questo caso, del Basic] in un codice a un solo byte (cioè in codice token)' e, nel 1984, il sostantivo, con il senso di 'trasformazione delle parole chiave in codici a un solo byte':

Una delle grosse differenze è il formato con cui vengono memorizzate le parole chiave: esse non vengono "**tokenizzate**", cioè convertite, al momento in cui viene inserita una linea, in un codice (token) di un solo byte, ma rimangono in memoria così come sono state introdotte. (Mauro Di Lazzaro, *Acorn Atom*, "MC microcomputer", 11, 1982, p. 33)

Inizia la fase di **tokenizzazione**: per risparmiare spazio in memoria le parole chiave del Basic sono trasformate in codici a un solo byte. (Andrea de Prisco, *16 K di ROM. Il Basic e il Sistema Operativo: come funzionano*, "MC microcomputer", 34, 1984, p. 117)

Ancora nel 1984 troviamo nella suddetta rivista anche la forma *detokenizzazione*, che fa riferimento al processo inverso, ovvero alla "riconversione in caratteri del comando Basic tokenizzato" (Andrea de Prisco, *L'ADP Basic: 51 nuovi comandi per le vostre periferiche*, "MC microcomputer", 36, 1984, p. 138).

Questa accezione particolare di *tokenizzazione* e *tokenizzare* sembrerebbe circoscritta a questi primi esempi e a pochi altri, tutti individuabili in riviste specialistiche (troviamo un altro caso nel 1984 sul periodico di informatica "Applicando"). Non si rintracciano altre attestazioni delle due forme fino almeno ai primi anni Duemila, quando compaiono nell'accezione linguistica di 'suddivisione del testo in token'. È però verosimile che nell'ambito del trattamento automatico dei testi fossero già in circolazione da diverso tempo, essendo il sostantivo *token*, inteso come 'occorrenza', in uso fin dagli anni Settanta; si deve poi considerare il fatto che molti degli articoli o dei volumi usciti in questo trentennio e dedicati alla tokenizzazione (manuale o automatica) sono scritti da autori italiani in

inglese.

Se si osservano i contesti d'uso delle prime occorrenze trovate si può notare che il significato dei due lemmi è dato per già noto:

L'italiano presenta però una situazione molto più complessa riguardo la segmentazione di un testo in singole parole che talvolta pone dei problemi anche teorici di non facile risoluzione, come nel caso dei clitici e delle preposizioni articolate: la **tokenizzazione** automatica di un corpus non annotato né lemmatizzato comporta inevitabilmente che siano registrate come singole occorrenze anche *diglielo*, *andandosene*, *portaci*, o, nel testo precedente, *all*, *del*, *sulla*, *nel* ecc. (Stefania Spina, *Fare i conti con le parole*, Perugia, Guerra, 2001, p. 108)

La strategia più diffusa ed efficace adottata dai **tokenizzatori** [programmi che eseguono la tokenizzazione automatica, ndr] per gestire le abbreviazioni e gli acronimi è quella di combinare la consultazione di elenchi e glossari contenenti le espressioni più comuni, con l'uso di regole euristiche generali per individuare le abbreviazioni e gli acronimi non conosciuti. [...] Ciascuna di queste regole non è chiaramente in grado singolarmente di catturare tutte le abbreviazioni di un testo, ma il loro uso combinato, unitamente alla disponibilità di ricchi repertori, permette di arrivare anche a un'accuratezza del 99% di abbreviazioni e acronimi riconosciuti e **tokenizzati** correttamente. (Alessandro Lenci, Simonetta Montemagni, Vito Pirrelli, *Testo e computer*, Roma, Carocci, 2005, p. 109)

A partire dai primi anni Duemila, sembra che questa accezione di ambito linguistico si vada in parte a sovrapporre a quella relativa all'informatica, tanto che è possibile rintracciare in forum e siti di informatica un'altra sfumatura semantica del verbo *tokenizzare*, quella di 'scomporre una stringa di caratteri in unità linguistiche identificabili' (si trova spesso infatti nella forma tipica *tokenizzare la stringa*):

volevo solo provare altre strade, anche perché in questo modo mi tocca **tokenizzare la stringa** più volte quando potrebbe essere sufficiente farlo una volta sola (o al massimo due)....([java] *superclasse che richiama costruttore di una sottoclasse?*, dal forum di html.it, 26/1/2004)

Nel secondo decennio del Duemila compaiono anche le altre accezioni informatiche, ovvero quella relativa alla sicurezza e all'autenticazione (nel 2013 il sostantivo e nel 2015 la forma verbale; la tokenizzazione è il processo attraverso il quale il codice della carta di credito è sostituito da una sequenza di numeri casuali) e quella, legata alla blockchain, di 'assegnazione di un token a un dato asset' (entrambi nel 2018):

Dalla ricerca è emerso che il servizio di sicurezza basato su tecnologia Cloud più richiesto è quello per la sicurezza della posta elettronica, considerato da almeno il 74% del campione come uno dei servizi essenziali di security a cui non si può rinunciare. Il 27% degli intervistati sta considerando anche l'implementazione della **tokenizzazione** dei dati come servizio Cloud, vista la necessità di conformarsi a nuove regole come il Payment card industry data security standard (Pci Dss). (Vito de Ceglia, *Servizi di sicurezza IT: nel 2015 il 10% sarà Cloud*, Repubblica.it, 28/5/2013)

Come funziona? È sufficiente inserire la carta nella fessura del Bancomat, digitare il PIN e scegliere sul display interattivo il portafoglio virtuale su cui caricarla. Attraverso lo schermo il sistema fornisce le indicazioni per registrarla e, una volta ricevuta conferma, invia un SMS all'utente con il link per

scaricare la mobile app, mentre i dati della carta già crittografati e **tokenizzati** vengono trasferiti automaticamente dall'ATM sul Cloud dell'emettitore attraverso la tecnologia HCE (Host Card Emulation) di SimplyTapp, azienda specializzata in soluzioni di mobile payment. (*L'ATM di Diebold diventa uno strumento che virtualizza le carte di pagamento*, pagamentidigitali.it, 21/10/2015)

La **tokenizzazione** viene definita come un metodo per convertire i diritti di una risorsa in un gettone o token digitale. Espresso in altro modo, è il processo di conversione dei diritti di un bene in un token digitale all'interno di una catena di blocchi o blockchain.

In questo senso, **tokenizzare** una proprietà è fondamentale per generare un token e collegarlo ad un contratto intelligente o smart contract che permetta a questo token di avere una relazione intrinseca con un bene reale. Per semplificare il concetto, partendo dal modello base di crowdfunding, la proprietà è divisa in azioni e le "azioni" vengono **tokenizzate**. (Fabrizio Villani, *La blockchain e la tokenizzazione del mercato immobiliare*, Fintastico.com, 13/2/2018)

Una sintesi

In sintesi, le prime attestazioni in italiano della forma *token* si hanno a partire dagli anni Trenta del secolo scorso: inizialmente si fa riferimento a buoni che sostituiscono il denaro o a pagamenti simbolici usati nei paesi anglosassoni. All'inizio degli anni Settanta inizia a circolare anche l'accezione linguistica di 'occorrenza' e nei primi anni Ottanta quella informatica di 'codice a un byte'. Nel corso degli anni Novanta si attesta il significato più generico di 'gettone' o 'buono', uso che si consolida poi negli anni Duemila. Gli anni Duemila vedono anche l'affermazione dei significati specialistico-informatici di 'dispositivo di autenticazione/sicurezza' e di 'moneta virtuale'.

A partire dal 2017 si rintracciano anche le prime occorrenze di *token* in associazione al concetto di blockchain, la cui frequenza aumenta nel corso del 2018: è l'anno in cui iniziano a comparire in italiano anche le locuzioni *token non fungibile*, *non-fungible token*, l'acronimo *NFT* e le concernenti accezioni informatiche del sostantivo *tokenizzazione* e del verbo *tokenizzare*. Dal 2021 si registra un primo picco nelle attestazioni di tali forme, parallelamente all'aumento dell'interesse economico e socio-culturale nei confronti degli NFT, seguito da un'impennata ancora maggiore nei primi mesi del 2022.

La rivista "Wired Italia" ha inserito *NFT* tra le **dieci parole del 2021**, assieme a (seguendo l'ordine dell'articolo) *green pass*, *perseveranza*, *long Covid*, *Yolo*, *abilismo*, *azzurri*, *metaverso*, *vax* e *transizione*. L'acronimo risulta essere anche il termine più utilizzato in inglese nel 2021 secondo la **classifica** stilata dal *Collins Dictionary*. Per dare un'idea della diffusione e della crescita del vocabolo, riportiamo i dati raccolti interrogando le pagine in italiano di Google: in data 27/9/2022 la ricerca restituisce 4.560.000 risultati per *NFT*, al 27/11/2022 le occorrenze erano diventate 6.150.000 (con un aumento di circa 1 milione e mezzo in due mesi), mentre al 27/2/2023 i risultati sono 11.600.000: sono, quindi, quasi raddoppiati negli ultimi 3 mesi, con un incremento di circa 5 milioni e mezzo.

Segnaliamo anche i dati relativi alla circolazione delle altre voci: la ricerca su Google (sempre in data 27/2/2023) restituisce 6.930.000 risultati per *token* (a cui si aggiungono 1.480.000 risultati per la forma plurale *tokens*; nonostante l'anglismo sia registrato dalla lessicografia come invariabile, *i token*, è molto

frequente trovare la forma plurale declinata secondo le regole morfologiche dell'inglese), 55.100 per *token non fungibile/i*, 66.000 per *non-fungible token*, 25.740 per *tokenizzazione/i*, 272.510 per le forme verbali (il dato è relativo alle occorrenze del verbo nei modi infinito, participio passato e gerundio presente, cioè quelli che presentano un maggior numero di occorrenze). Si deve comunque tener conto del fatto che tali numeri sono solo indicativi, tanto più che una ricerca che permetta di distinguere tra le diverse accezioni di *token*, *tokenizzazione* e *tokenizzare* è praticamente impossibile.

Per quanto riguarda invece altri dati relativi alla diffusione dei termini, si nota che in altri contesti i numeri sono più contenuti: cercando nelle pagine in italiano di Google libri si rintracciano 35.300 occorrenze di *token*, 9.140 di *NFT*, 1.403 di *token non fungibile/i*, 1.340 di *non-fungible token*, 1.204 per *tokenizzazione/i*, 1.030 delle forme verbali. Sulla "Repubblica" sono presenti 554 risultati per *token*, 512 per *NFT*, 15 per *token non fungibile/i*, 147 per *non-fungible token*, 28 per *tokenizzazione/i*, 13 per le forme verbali (6 *tokenizzare*, 3 *tokenizzato*, 1 *tokenizzata*, 1 *tokenizzati*, 1 *tokenizzando*, 1 *tokenizzano*); sulla "Stampa" troviamo 203 occorrenze di *token* (a cui si aggiungono le 138 presenti nell'archivio, che arriva fino al 2006), 135 di *NFT*, 31 di *token non fungibile/i*, 41 di *non-fungible token*, 6 di *tokenizzazione* e 3 occorrenze delle forme del participio passato (1 *tokenizzata*, 1 *tokenizzate*, 1 *tokenizzati*). Infine, sul "Corriere della Sera" è possibile rintracciare 208 risultati per *token*, 498 per *NFT* (ma il rumore è decisamente alto, dato che la voce corrisponde a diverse sigle omonime), 11 per *token non fungibile/i*, 4 per *non-fungible token*, 5 per *tokenizzazione*, 7 per le forme del participio passato (2 *tokenizzato*, 2 *tokenizzate*, 3 *tokenizzati*).

Altre formazioni

Concludiamo con un accenno ad altre nuove formazioni (nessuna delle quali registrata dalla lessicografia), attestate più o meno frequentemente in italiano negli ultimi anni, costruite con il sostantivo *token* o derivate dal verbo *tokenizzare*: *tokenomics*, *criptotoken*, *tokenizzabile* e *tokenizzabilità* (che si aggiungono così alla "datata" *detokenizzazione* sopra citata).

La voce inglese *tokenomics*, giunta come prestito integrale, è una parola macedonia formata dai sostantivi *token* e *economics* 'economia' e fa riferimento a 'un'economia basata sui token'; è spesso usata come sinonimo di *criptoeconomia*, che indica un 'modello economico basato sulle criptovalute'. Oltre alla forma *tokenomics* (che conta 32.200 occorrenze nelle pagine in italiano di Google), sono usate in italiano, e con lo stesso significato, anche la locuzione *token economy* (47.400 risultati) e l'adattamento *tokenomica* (14.000 risultati; minoritaria la forma *tokenomia*, che ha soltanto 660 occorrenze).

Il sostantivo è segnalato nella sezione *Neologismi della settimana* (nella settimana che va dal 24 al 30 ottobre 2022) del portale Treccani, con un esempio risalente a settembre 2022; in rete è però possibile rintracciare un'attestazione già nel 2017 nella rivista "Wired":

Tokenomics e criptovalute sono argomenti caldi e in rapida evoluzione. A oggi più di un miliardo di dollari di investimenti in capitale di rischio sono andati a società che si occupano di sviluppo di tecnologie blockchain. (Ilaria Caielli, *Ora le università fanno corsi e master sulle criptovalute*, wired.it, 8/12/2017)

Il termine *criptotoken* è un adattamento grafico dell'inglese *crypto token*: è composto dal confisso *cripto-* e dal sostantivo *token*, sul modello di *criptovaluta* e *criptomoneta*. Usato come sinonimo di *token*, indica le monete virtuali che si basano su una valuta digitale preesistente. Il sostantivo non ha una grande diffusione in italiano (272 risultati per la forma unverbata e 612 per quella con la grafia separata), mentre risulta più frequente trovare il prestito integrale *criptotoken* (8.500 occorrenze). La prima occorrenza rintracciata è in un volume del 2018:

Con grande sollievo di Slock.it, la SEC decise di non portare avanti l'accusa, ma il comunicato stampa che spiegava le motivazioni di questa decisione era un avvertimento minaccioso: non solo chiariva che i sempre più numerosi produttori di **criptotoken** avrebbero dovuto mettere in conto delle azioni di regolamentazione, ma ricordava anche quanto fosse ampia la giurisdizione degli enti regolatori che hanno alle spalle il peso delle leggi negli Stati Uniti. (Michael Casey, Paul Vigna, *La macchina della verità*, Milano, Franco Angeli, 2018, traduzione di Stefano Ballerio; consultato in formato e-book)

L'aggettivo *tokenizzabile* è invece un derivato del verbo *tokenizzare*, con l'aggiunta del suffisso *-bile* e fa riferimento a qualcosa (specialmente un bene) 'che può essere tokenizzato'; a sua volta, il sostantivo *tokenizzabilità* è formato da *tokenizzabile* con il suffisso *-ità* e significa 'l'essere tokenizzabile'. La diffusione delle due forme è piuttosto limitata: su Google in italiano si hanno 340 occorrenze dell'aggettivo (219 per il singolare e 121 per il plurale) e soltanto 4 del sostantivo (ma due rimandano a Google libri). Riportiamo le prime attestazioni individuate, risalenti al 2018 (per l'aggettivo) e al 2019 (per il sostantivo):

Bisogna premettere che tutto ciò che è "**tokenizzabile**" sarà "tokenizzato" e tutto ciò che è decentralizzabile sarà decentralizzato. (Daniel Casarin, *Blockchain, marketing e vendite: tutto quello che c'è da sapere*, blog.advmedialab.com, 5/2/2018)

Non tutti i Business Model si prestano ad essere tokenizzati, dei metodi per valutare la **tokenizzabilità** di un business sono in corso di studio. [...] Il fatto che un Business Model sia **tokenizzabile** non significa che sia conveniente (Michele Mostarda, *Come Tokenizzare un Business Model*, hardest.medium.com, 21/8/2019)

Per quanto riguarda, infine, il vocabolo *tokenismo*, si tratta di un adattamento della voce inglese *tokenism*, derivato di *token*, inteso come 'simbolo', ma il significato non è in alcun modo collegato con quelli finora discussi: usato in senso spregiativo, indica "la pratica di fare concessioni formali a minoranze o a gruppi sottorappresentati (per dare l'impressione di equità e rispetto delle pari opportunità); gesto di concessione" (la definizione è tratta dalla rubrica *Parole del giorno* presente sul portale dizionari.piu.zanichelli.it ed è datata 2/2/2011; il termine risulta comunque attestato in italiano già dagli anni Settanta).

Nota bibliografica:

- Barbera 2013: Manuel Barbera, *Linguistica dei corpora e linguistica dei corpora italiana. Un'introduzione*, Milano, Qu.A.S.A.R., 2013.
- Barbera, Corino, Onesti 2007: Manuel Barbera, Elisa Corino, Cristina Onesti, *Cosa è un corpus?*

Per una definizione più rigorosa di corpus, token, markup, in *Iid.* (a cura di), *Corpora e linguistica in rete*, Perugia, Guerra Edizioni, 2007, pp. 25-88.

- Bolasco 2013: Sergio Bolasco, *L'analisi automatica dei testi: fare ricerca con il text mining*, Roma, Carocci, 2013.
- Chiari 2007: Isabella Chiari, *Introduzione alla linguistica computazionale*, Roma-Bari, Laterza, 2007.
- FSB 2019: Financial Stability Board, *Decentralised financial technologies: Report on financial stability, regulatory and governance implications*, 2019.
- Hileman e Rauchs 2017: Garrick Hileman, Michel Rauchs, *Global Blockchain Benchmarking Study*, 22 settembre 2017.
- Lenci, Montemagni, Pirrelli 2005: Alessandro Lenci, Simonetta Montemagni, Vito Pirrelli, *Testo e computer*, Roma, Carocci, 2005.
- OECD 2020: Organization for Economic Co-operation and Development (OECD), *The Tokenisation of Assets and Potential Implications for Financial Markets*, OECD Blockchain Policy Series, 2020.

Cita come:

Lucia Francalanci, *Risemantizzazioni e neoformazioni tra linguistica e informatica: token, tokenizzazione e NFT*, "Italiano digitale", XXIV, 2023/1 (gennaio-marzo)
DOI: 10.35948/2532-9006/2023.27949

Copyright 2023 Accademia della Crusca

Pubblicato con licenza creative commons CC BY-NC-ND